Modeling Reduction of Pandemic Influenza Using Pharmaceutical and Non-Pharmaceutical Interventions in a Heterogeneous Population

by

Anna Teytelman

B.S. Management & B.S. Mathematics with Computer Science
Massachusetts Institute of Technology, 2008

Submitted to the Sloan School of Management in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Operations Research

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2012

Signature of Author: _____

Anna Teytelman
Sloan School of Management
May 18, 2012

Certified by: _____

Richard C. Larson
Mitsui Professor of Engineering Systems Division
Director of Center for Engineering Systems Fundamentals
Thesis Supervisor

Accepted by: _____

Dimitris Bertsimas
Boeing Professor of Operations Research
Co-director, Operations Research Center

Modeling Reduction of Pandemic Influenza Using Pharmaceutical and Non-Pharmaceutical Interventions in a Heterogeneous Population

by

Anna Teytelman

Submitted to the Sloan School of Management
on May 18, 2012 in Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy in Operations Research

ABSTRACT

In an event of a pandemic influenza outbreak such as the great "Spanish Flu" of 1918 and the more recent 2009-2010 H1N1 "Swine Flu" scare, pharmaceutical as well as non-pharmaceutical resources are limited in availability and effectiveness. In this thesis we apply OR methods to evaluate the effectiveness of such resources and the strategies for reducing the number of infections resulting from an outbreak.
In the first half of this work, we focus on epidemiological analysis of influenza modeling in a heterogeneous population. The majority of existing epidemiological literature models influenza spread in a statistically homogeneous population, but the model-based inclusion of heterogeneity by contact rate, susceptibility, and infectivity introduces significant effects on disease progression. We introduce a new discrete-time influenza outbreak model for a heterogeneous population and use it to describe the changes in a population's flu-related characteristics over time. This information allows us to evaluate the effectiveness of different vaccine targeting techniques in achieving herd immunity, that is, the point at which there is no further growth in new infections.
In the second half of this work we switch to a practical application of OR methods in a pandemic situation. We evaluate the effectiveness of vaccines administered to US states during the 2009-2010 H1N1 pandemic. Since the US is geographically diverse and large, the outbreak progressed at different rates and started at different times in each individual state. We discuss dynamic, multi-regional, vaccine allocation schemes for large geographical entities that take into account the different conditions of the epidemic in each region and maximize the total effect of available vaccines. In addition, we discuss effective strategies for combining vaccines with non-pharmaceutical interventions such as hand-washing and public awareness campaigns to decrease the strain of an outbreak on the population.

Thesis Supervisor: Richard C Larson
Title: Mitsui Professor of Engineering Systems
Director of Center for Engineering Systems Fundamentals

Acknowledgements

The contributions of this thesis would have never seen the light of day if it wasn't for the support of many people throughout my undergraduate and graduate career. Most of all, I would like to thank my advisor, Dr Richard Larson, for the way he has supported me for the last four years. When I just entered the graduate program at the Operations Research Center, I had no idea what to expect from a research position or even how to approach academic work. It is thanks to Dr Larson that I started to see research problems in everyday life, and to appreciate the importance of understanding the "physics" of OR problems. His enthusiasm is contagious, and it's easy to get excited about work when discussing it with him. It is difficult to strike a balance between guiding a student in her research and letting her find her own research problems and do original work. I am very lucky to have worked with somebody who found this balance so well. I hope that 50 years at MIT is just the beginning, and he continues to help many more students in their academic careers.

I also want to thank my committee members Drs Stan Finkelstein, Steve Graves, and Asu Ozdaglar for taking the time to guide me in writing this thesis. It is an honor to have the support of such amazing people at MIT.

I also must mention the other members of the MIT "Flu Team" and their help with this work. Our research was truly a team effort. Dr Stan Finkelstein has been an invaluable advisor over the last four years. His education as a medical doctor and researcher helped me understand the policy needs of practitioners in an influenza pandemic. After all, the needs of healthcare officials are the motivation behind this work. Kallie Hedberg, Julia Hopkins, and Sahar Hashmi have also been working with us on flu research over the years. Collecting data for our analysis of 2009-2010 H1N1 pandemic was no small task, and this analysis only exists thanks to Kallie and Julia's efforts. Sahar and Abby Horn were also very kind to proofread and give feedback parts of this thesis.

I have also had multiple conversations with Denis Chebikin, Dimdim Katz, and Maks Maydanskiy as well as many other students at the ORC about my research. The open environment at the ORC is structured exactly to encourage brainstorming and forming new ideas with the help of people who might be working on things entirely unrelated to your research. I think this is fantastic, and I'm glad that I chose to continue studying OR here at MIT.

Of course I could have done none of this without the support of my family and my friends. Eight years at MIT is an exciting but exhausting endeavor. I could not have continued without knowing that some people will always support me in all my life decisions. My parents, Rebecca and Alex have always made time to call and visit me in Boston. I know how invested they are in anything that I do, and that makes me want to make them proud. I hope that I have been successful so far. It has also been great living in the same city as my brother Lev and my sister-in-law Anna. Knowing that family is not more than 20 minutes away has been a huge help. Of course, the greatest support came from my niece and nephew Dalia and Zyama, and I hope someday they will get a chance to read my words of gratitude. I also have to add my friends to this list. Abby, Angelique, Dasha, Denis, DimDim, Gabriel, Ilya, Jay, Lev, Maks, Matt, Mayank, Melis, Sawyer, and Sonya: thank you for all your support.

Finally, I want to thank the organizations and fellowships that made my work here possible. My research was primarily supported by U.S. Centers for Disease Control and Prevention (CDC), grant number 1 PO1 TP000307-01, "LAMPS (Linking Assessment and Measurement to Performance in PHEP Systems), awarded to the Harvard School of Public Health Center for Public Health Preparedness (HSPHCPHP) and the Massachusetts Institute of Technology (MIT), Center for Engineering Systems Fundamentals (CESF), with additional funds provided by the Sloan Foundation of New York under a grant entitled, "Decision-

Oriented Analysis of Pandemic Flu Preparedness & Response." I also extend my deepest gratitude to the MIT Ida M. Green fellowship which supported me for my first year at MIT.

Table of Contents

Chapter I: Introduction to pandemic influenza

According the US National Intelligence Council's 2020 Project, ''Mapping the Global Future," an influenza pandemic is the single most important threat to the global economy [1]. A potential new pandemic similar to the infamous "Spanish Flu" pandemic of 1918 or the "Asian Flu" outbreak of 1957 can have catastrophic consequences for the United States as well as the global economy and population. While medical and academic advances have been made since the first worldwide pandemic event in 1918, the global community still has very little power to stop or mitigate a pandemic in its devastating force. In addition to naturally evolving viruses, we now have to worry about a possible threat of a biologically engineered deadly strain of a vaccine. Just recently, a new, deadly version of an H5N1 virus has been produced in a lab setting [2]. If released to a susceptible population, such a virus can easily infect hundreds of millions of people throughout the world. A medical "cure" for influenza has not been identified. We can only fight an outbreak through various preventative measures such as vaccines, school closures, and good hygiene practices; and mitigating measures such as anti-viral drugs and self-isolation.

In this work we focus on using operations research techniques to improve decision-making in both preparation for a pandemic and reaction to an ongoing outbreak. We introduce a generalized influenza spread model and use this model to prove some fundamental results regarding intervention effectiveness. We then use the epidemic data from the recently occurring 2009-2010 H1N1 pandemic to demonstrate the importance of timing in implementing various interventions and suggest an effective vaccine allocation strategy for a large multi-regional entity such as the United States.

1. Seasonal and pandemic influenza

A new outbreak of influenza occurs when an already existing strain of a virus mutates to form a new strain, to which the majority of a population does not yet have a built up immunity. Once the new virus enters the community, if the strain is particularly transmissible it will pass from person to person spreading throughout the susceptible population. The virus can be airborne and pass through face-to-face contact or can be contracted by a susceptible person if she touches a surface contaminated by particles containing the virus.

Once a susceptible person becomes infected by virus particles, the virus replicates within the host until a critical mass is reached and the person becomes symptomatic [3]. That initial replication period is the latent period when the person is infected but not yet infections. After the latent period, the infected individual becomes infectious and starts shedding virus. Sometime during the infectious period, the infected person becomes symptomatic, and often retires from the population to treat the symptoms. See Figure 1 for stages of an individual flu case progression. Throughout this work, we will often refer to the entire timeline in Figure 1 as a "generation" of the flu. In our models, we will also use the term "day" to refer to a generation.

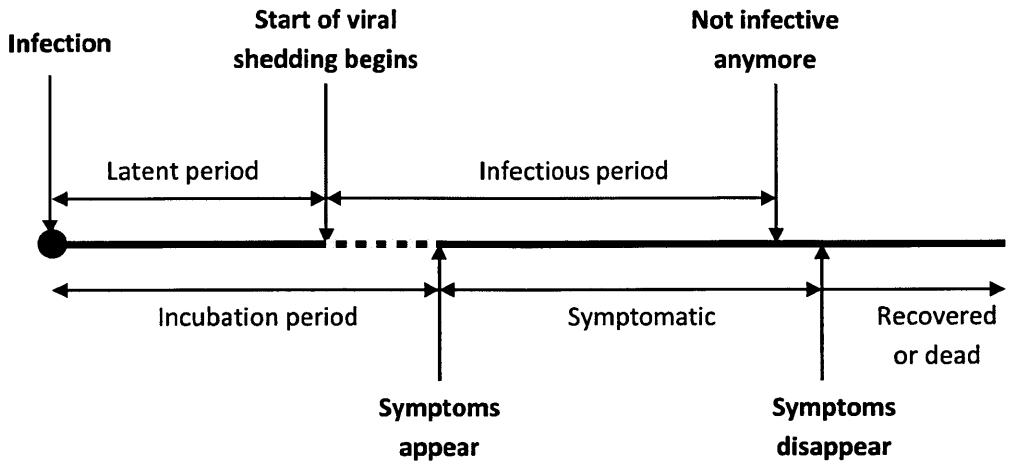


Figure 1: The temporal progression of the flu's symptoms and infectiousness. The highlighted segment between the start of viral shedding and appearance of symptoms is the time when people are particularly likely to spread the flu as they are unaware of their infectious status [3].

Seasonal influenza outbreaks occur almost every year. A seasonal virus is usually characterized by fairly low transmissibility and the typical demographics for a seasonal outbreak include the entire spectrum of a population. Although approximately 36,000 deaths [4] result from seasonal influenza each year in the US, the majority of these deaths occur due to complications in the at-risk segments of the population consisting of the elderly, the very young, and people with preexisting conditions. While seasonal influenza constitutes a burden on the population and the economy it is usually well managed on an annual basis due to its predictability and ample time to prepare for each yearly outbreak.

The seasonal influenza preparedness process begins with the identification of new strains of the influenza virus. Surveillance centers report on the new viruses found circulating in the general population and a governing body such as the World Health Organization or the Centers for Disease Control and Prevention in the US determine the few strains that are to be included in a particular season's vaccine that will be eventually delivered to healthcare providers. Once these strains are identified, vaccines are manufactured using chicken eggs to grow a weakened virus. Preparation of the environment and growing the virus in bulk quantities takes approximately six months. After the first few samples are created, the virus is clinically tested for safety and then ordered in bulk to be produced and delivered to the public (Figure 2).

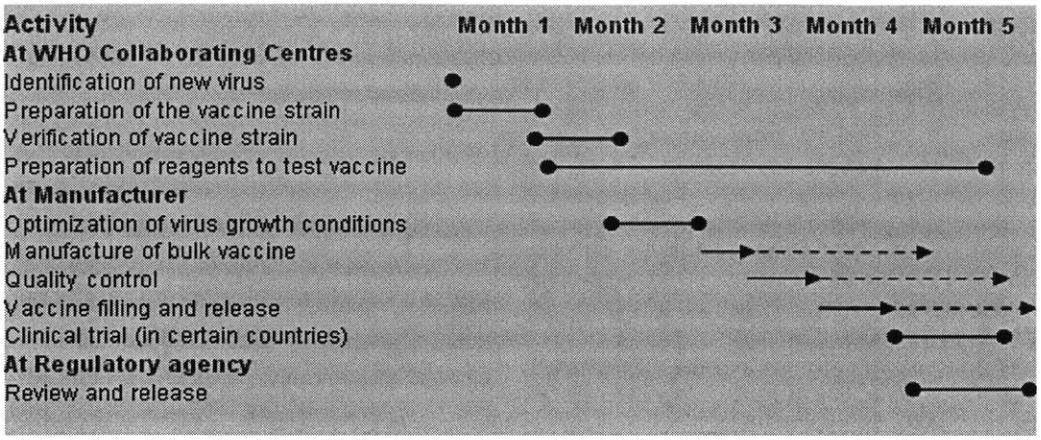


Figure 2: The arrows with dotted lines preceded by non-broken arrows indicate the time period required for the first time an activity is done (non-broken arrow line) that is then repeated (dotted arrow line). The solid lines signify that the activity takes place within a finite period [5].

Since the preparation for a seasonal flu virus begins relatively early, the six month lag time is not detrimental to managing the spread of influenza. Moreover, the people most at risk of serious complications such as pneumonia have been identified – elderly, infants, pregnant women, etc. - and are first to receive the protection offered by distributed vaccine [**6**]. In addition, recently developed methodology has provided a way to shorten the vaccine manufacturing process by 8 to 10 weeks [**7**, **8**]. The new methodology, which would use animal cells instead of eggs to cultivate the virus, will eliminate the majority of the lag time. One production plant using such methodology is on track to be opened in the United States [**9**].

Contrary to seasonal influenza, pandemic outbreaks have much more devastating effects on the world population. Pandemic viruses are usually a mutated variation of an animal strain and are much more deadly and more easily transmitted than its seasonal counterparts [**4**]. The general population usually has little built-up immunity to the novel virus. Consequently, groups that are usually not susceptible to strong effects from influenza also suffer infections. One of the defining characteristics of pandemic influenza is that the strongly effected demographic shifts to healthy individuals in their 20s or 30s.

In addition, since a pandemic arises suddenly without much time for preparation, governing bodies have significantly less time to commission vaccine production. When vaccines are finally manufactured, it is possible that they are no longer needed to the global population. This is particularly evident in our in-depth analysis of the 2009-2010 H1N1 "Swine Flu" pandemic in Chapters 4 and 5.

Experts agree that an emergence of a devastating new pandemic influenza strain is a matter of when not if. When such an event occurs, if the world is not meticulously prepared for the outbreak the results can cause devastating effects on the population, medical infrastructure, and economic situation of all affected countries. The first such event is referred to as the great "Spanish Flu," lasting from June 1918 until finally dying out in 1919. This outbreak was more virulent and deadly than any seen before, with nearly 500 million people infected, a third of the world's population at the time [**10**]. The infection rate was combined with a staggering 2.5% case-fatality rate, leaving anywhere from 20 to 40 million people dead throughout the world. The surprising nature of the outbreak was not limited to the high morbidity and mortality rates, but also included the demographics affected by the mortality rate. Consider the death rate

during this pandemic as compared to that of the seasonal influenza rates from a few years prior to 1918 in Figure 3.

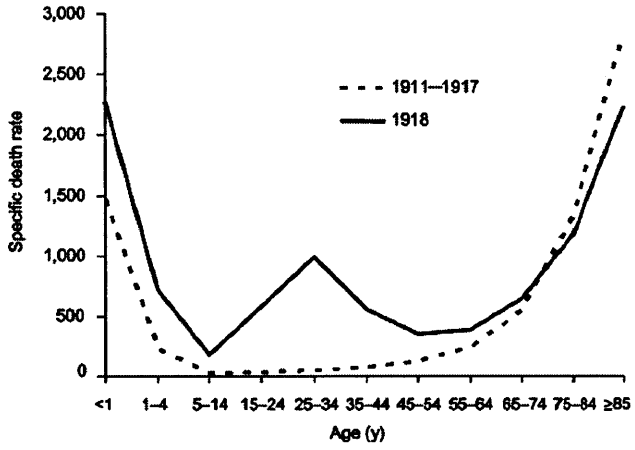


Figure 3: W-curve: the death rate by age for the 1918-1919 pandemic. The unique factor of this outbreak is that unlike most seasonal flu that is deadly for the very young and elderly, it killed an unusually high number of people in the 20-40 years old category [3, 10].

Two other major outbreaks occurred in the 20th century. The 1957 "Asian Flu" started in China but progressed to the other end of the world very quickly, infecting 10-35% of the world's population and killing about two million people [**11**]. The 1968 "Hong Kong Flu" also killed about one million people throughout the world [**12**].

The Lowy Institute for International Policy reported that even a very mild pandemic, like the Hong Kong or the Asian Flu would result in a global economic loss of $330 billion and a human loss of 1.4 million. A pandemic of 1918 proportions is projected to result in a cost of $4.4 trillion of global economic output and the deaths 140 million people [**13**].

While these 3 pandemics were the most devastating outbreaks experienced by the world in the last century, researchers did not yet have the resources to record and analyze the data from the outbreaks as extensively as we have been able to do in the last few decades. The most data-driven progress in modeling and evaluating pandemic response has come from outbreaks such as the SARS epidemic that lasted from 2002 to 2004 predominantly in China and the surrounding areas. This outbreak had a fairly low impact on a global scale, with approximately 1000 deaths worldwide, but it demonstrated the effectiveness of non-pharmaceutical interventions on mitigating the effect of a serious respiratory illness [**14**].

The most complete data available to us comes from a recent pandemic, the 2009-2010 H1N1 "Swine Flu" pandemic. While the outbreak was not as deadly as was at first feared, the information gathered by the United States during this pandemic is the key to the analysis we present in Chapters 4 and 5. The novel H1N1 influenza virus surfaced in San Diego in early April of 2009, very soon thereafter in Mexico and was announced as a pandemic by the World Health Organization in June 2009 [**15, 16**]. By early fall most of the world was experiencing the fall wave of the H1N1 pandemic [**17**]. The US Center for Disease Control (CDC) took the threat of a pandemic seriously, and by the middle of April 2009,

researchers began to develop vaccines matched to the new virus [**15**]. The inherent delays of egg-based vaccine production resulted in the first vaccines being shipped in early October 2009 [**18**]. The vaccine had varying ameliorative effects in different regions of the US. Due in part to early school openings in the Southeast [**19**], the major wave of the outbreak hit the southeastern states earlier than the northern ones. Nonetheless, vaccines as produced were delivered to states strictly on a per-capita basis [**20**], with each state receiving vaccine in direct proportion to its population. As a result, some states were able to vaccinate a significant portion of their population prior to the major flu wave hitting, while others did not receive vaccines until the major flu wave had already passed and interest in vaccinations had waned. Vaccine distribution within individual states was significantly more complicated as states had relative freedom in utilizing the vaccines shipped from the CDC [**21**].

Regardless of the classification of an influenza outbreak, there are a number of measures that can be taken to mitigate the effects of the disease on the population. There are three main categories of interventions.

1. **Pharmaceutical interventions:** Antiviral medication such as Oseltamivir (Tamiflu) can be used for treatment or prophylactic purposes. If taken within the first 48 hours of manifestation of influenza-like symptoms. Some of this medication has been shown to reduce symptoms and complications and shorten illness by one to two days [**22**]. Antiviral medication is generally not used for preventative purposes for anyone but members of the medical profession who are likely to come across many infectious individuals in the course of their daily activities. The medical intervention most used for massive distribution is vaccines. These are also not 100% effective, but they significantly lower the susceptibility and infectivity of an individual. Newly administered vaccines take up to proximately 2 weeks to take effect [**23**].

2. **Non-pharmaceutical interventions:**
    - **Individual**: a single person can make behavioral changes to help lower his or her chances of contracting influenza and to prevent the spread of the virus to others. These changes include improved hygiene practices such as increased hand-washing. Telecommuting to work for jobs involving multiple interactions with other people also prevents transmission. Changes such as these are largely up to the individual to implement, but with a high compliance rate they have been shown to be effective in curbing the spread of disease in an entire population [**14**].

    - **Government mandated**: policy-making bodies may also implement non-pharmaceutical interventions with significant effects. Such interventions include school closures and public awareness campaigns to encourage the public to use good hygiene practices.

Although it has been shown that individual behavioral changes can be very helpful in mitigating the effects of the flu, [**3, 14**] we will focus our work on the options available to policy makers such as the CDC in implementing both pharmaceutical and non-pharmaceutical interventions.

2. Modeling disease spread

2.1 Basic reproductive number R_0

The vast majority of all epidemiological literature modeling infectious disease includes estimates of the basic reproductive number, R_0. R_0 is the fundamental construct quantifying the transmissibility of a virus strain. It is defined as the expected number of secondary cases produced, in a completely susceptible population, by a typical infected individual during its entire period of infectiousness [**24**]. In formal terms, for each person in the population, there exists a random variable that corresponds to the number of secondary cases that person would cause if placed in a fully susceptible population. R_0 is the mean of this random variable for a randomly picked individual.

R_0 is the first parameter that is determined with the emergence of any new outbreak. It is considered a representation of the transmissibility of a particular strain of influenza and corresponds to the exponential growth factor of the epidemic curve in the beginning of the outbreak. Consider, for example, an epidemic in which $R_0 = 2$ and select the first infectious person, "patient zero", from the population. We assume that our population is large enough, that in the early stages of the epidemic nearly everyone is susceptible and our "completely susceptible population" criterion is satisfied. In that case, patient zero infects, on average, two people. Those two newly infected people in turn also infect about two people each, causing four new infections to emerge. As this goes on we experience a branching process where the number of infections increases by a factor of R_0 until the number of people who have already acquired immunity in the population becomes significant enough to stop exponential growth. See Table 1 for estimated values for R_0 for a few major pandemics in the last century.

Pandemic	Estimated R_0 range
1918 Spanish Flu	1.7-3.4 [**25**]
1957 Asian Flu	1.8 [**26**]
SARS	2-3.3 [**27**]
2010 Swine Flu	1.5-3.1 [**28**]

Table 1: R_0 values for major pandemics in the last century

R_0 is central to the idea of a threshold criterion. In general, the threshold criterion suggests that whenever $R_0 > 1$ the epidemic takes off and the epidemic curve grows while if $R_0 < 1$, the epidemic should die out.

While this is true in the average case, R_0 is an aggregate statistic and must be taken with a grain of salt. As the expectation of a random variable, it hides the underlying probability distribution. Consider two hypothetical homogeneous population examples in which R_0 is astronomically high.

- Scenario 1: Every person would infect exactly 90 others if placed in a fully susceptible population.
- Scenario 2: The number of people infected by any given member of the population in a fully susceptible population follows the following probability mass function $\begin{cases} 100\ w.p.\ 0.9 \\ 0\ w.p.\ 0.1 \end{cases}$

In both of these scenarios, $R_0 = 90$. However, in the first population, if the virus is introduced to the population, it will cause an epidemic with probability 1, as patient zero will infect exactly 90 people, In the second scenario, with probability 0.1 patient zero will infect no one at all, and the outbreak will die

out as she recovers despite the fact that $R_0 > 1$. The variance of the underlying population is hidden by R_0 but plays an important role in outbreak dynamics.

Another important aspect of the definition of R_0 is the use of the term "typical." The ambiguous part here is the selection mechanism used to choose patient zero in a heterogeneous population where not all individuals behave identically. One may be tempted (and some are) to select a person from the population uniformly at random. However, it is preferable to select patient zero using other methods. We will discuss in detail, in Chapter 2, the reasons why the uniform selection mechanism will likely provide an underestimate for the true R_0 in a heterogeneous population. A distribution that favors selecting the more active and susceptible people in the population is more appropriate.

As a final word of warning, one must also avoid the temptation of thinking of R_0 as a constant particular to a given outbreak. In fact, R_0 is an expectation of a random variable that consists of multiple components. One way to start breaking up R_0 is to think of it as the function, $R_0 = \lambda p$. Here, λ is the average number of contacts that a typical infectious person in the population has during the period of her infectiousness, and p is the expected probability that a single contact between an infectious and a susceptible person results in the susceptible person's infection. This is just one level of simplification, as p can be further deconstructed into components corresponding to the infectiousness of the infected person involved in a typical face-to-face contact and the susceptibility of the susceptible person. The key observation here, however, is that the distributions of λ and p vary not only with the strain of a flu virus, but also with the geographical, cultural, and other factors particular to any community.

Despite some of these concerns, R_0 is still a very good aggregate measure of a virus's spreading potential. When used cautiously, it should guide epidemiologists and healthcare officials in appropriately reacting to a new influenza outbreak.

2.2 Effective reproductive number $R(t)$

The effective reproductive number is closely related to R_0. Recall that R_0 is the expected number of infections a single infected person causes in a *fully susceptible* population, that is, at the very beginning of an outbreak. Similarly, $R(t)$, is a generalized concept of the expected number of infections caused by a single person. It is defined as the expected total number of secondary cases produced by a typical infectious individual at time t. If we were to take the infectious population at time t, and pick one person out, at random, and record all infections caused by that individual we would have an estimate of $R(t)$. In an entirely homogeneous population where every person is identical, the calculation for $R(t)$ is simple. Suppose that at time t the proportion of the population $K(t)$ is already immune to the infection due to having already been infected or through vaccination. In that case, we know that a typical infectious person would infect, on average, R_0 others if every other person were susceptible. However, only $\big(1 - K(t)\big)$ fraction of the population is susceptible, so if we had independent mixing only a fraction $\big(1 - K(t)\big)$ of an infectious person's usual contacts are susceptible. In other words, the contact rate parameter λ needs to be scaled by $\big(1 - K(t)\big)$ to reflect the number of contacts that occur with a susceptible person. Therefore, in a homogeneous population $R(t) = R_0\big(1 - K(t)\big)$. This analysis becomes more complicated in a heterogeneous population, where people behave in different ways causing the characteristics of a population to change with time.

Note that when $R(t) < 1$, the number of infectious people in the population declines, as each infectious person does not replace herself in the infectious population. Because of this, the goal of any intervention program is to lower $R(t)$ as much as possible through various means, ideally reaching a point where $R(t) < 1$.

2.3 Herd immunity

The fraction of the population that is immune at the time when $R(t)$ becomes equal to 1, is commonly referred to as herd immunity. Specifically, herd immunity is the fraction of the population that needs to be immune to the disease so that $(t) = 1$ [**29**]. Herd immunity coincides with the peak of infection. That is exactly the point at which the epidemic curve stops increasing and the outbreak begins to subside. Herd immunity can occur naturally without any intervention from the outside as people contract infection, recover, and acquire immunity to the outbreak. It is also possible to obtain herd immunity through vaccinating a portion of the population. In Chapter 3 we will discuss in detail how to achieve herd immunity using different targeting strategies.

Calculating herd immunity for a homogeneous population is once again relatively easy. Recall that in a homogeneous population $R(t) = R_0(1 - K(t))$. Herd immunity in this situation is the fraction of the population K, for which $R(t) = R_0(1 - K) = 1$. Solving for K, we get $K = 1 - \frac{1}{R_0}$. As with $R(t)$, herd immunity gets more complicated in a heterogeneous population.

2.4 Common modeling techniques

2.4.1 Homogeneous compartmental modeling

There is extensive literature on modeling pandemic influenza spread. The simplest and most common deterministic models for influenza spread follow some variant of the S-I-R compartmental approach where each person in the population is susceptible, infected, recovered, or dead. Such models are often used in a homogeneous setting, where all people in a single compartment behave identically and there is random mixing [**30, 31**]. Consider the illustration in Figure 4. Each person is either susceptible, infected or recovered. Within each compartment all people are identical. At any time, the susceptible, infected, and recovered compartments contain S(t), I(t), and R(t) individuals respectively. People move from the S compartment to the I compartment at rate $p\beta S(t)I(t)$ and from the I compartment at the rate $\alpha I(t)$.

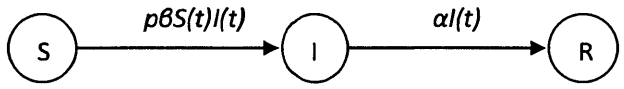


Figure 4: A typical S-I-R compartmental model. [3]

The values for *S(t)*, infectious *I(t)* and recovered *R(t)* are given by the solution to the following three differential equations:

$$\frac{dS}{dt} = -p\beta SI; \quad \frac{dI}{dt} = p\beta SI - \alpha I; \quad \frac{dR}{dt} = \alpha I$$

The assumptions that are embedded in this model are:

1. An average member of the population makes βN contacts per unit of time, assuming N is the total population size (mass action incidence).
2. The probability of infection spreading from an infected to a susceptible individual per contact is p.
3. Infected individuals leave the infected group at a rate of αI given I people in the infected compartment. To elaborate further, the duration of the infective period is exponentially distributed with mean $1/\alpha$.
4. There are no births or deaths from the overall model; there are no demographic effects.

The value of the basic reproductive number for this model is:

$$R_0 = \frac{\beta p S(0)}{\alpha}. \text{ [3]}$$

The S-I-R model is made more complex with the addition of the possibility of dying. Other compartments like the exposed state can also be added to reflect the reality of an influenza outbreak in which some people may be infectious without experiencing symptoms.

2.4.2 Heterogeneous compartmental modeling

This S-I-R approach may be considered incomplete because it ignores population heterogeneity. One visible form of heterogeneity is manifested as "super-spreaders". The extreme cases of super-spreaders such as Typhoid Mary in the case of typhoid fever in the US [**32**], or Liu Janlun during the SARS outbreak [**18, 33**] bring with them the possibility of one particularly active person sparking an outbreak even if the rest of the population is relatively weak at transmitting the infection [**34, 35**]. More typical are the populations where some members of the population are inherently more active at spreading the disease than others. Literature on heterogeneity consistently shows that this inherent heterogeneity causes a significant impact on the dynamics of the outbreak and even whether such an outbreak occurs at all [**3, 30, 36**]. Nigmatulina and Larson showed that people who have a large number of daily contacts are the "drivers" of the infection. These people tend to get infected early in the outbreak and drive the propagation of the flu in the early stages [**3, 36, 37**].

We will concentrate on three major dimensions of heterogeneity: *activity level, susceptibility, and infectivity.* Susceptibility is determined by a person's natural immune system, causing some to be more likely to contract influenza after a brief contact with an infectious person. Infectivity is a measure of how easily a person "sheds" virus particles when infected. Modeling heterogeneous activity levels has recently become plausible with the work of Fu[**38**], and Mossong *et al* [**39**]. Both of these works provided a summary of contact diaries filled out by participants, in which they enumerated all contacts that they had throughout a single day. In particular, Mossong *et al*, provided contact rate distributions from eight European countries with the specific goal of being used for influenza modeling. The availability of such data allows us to model infection spread accurately and more realistically than simple homogeneous modeling. In our research we intend to explore the ways in which this modeling can be used to effectively develop mitigation strategies in a heterogeneous population during an influenza outbreak. Susceptibility and infectiousness are much less visible constructs and there is still not enough reliable information about them to use for heterogeneity modeling. We hope that the influenza researches will continue working on quantifying these sources of heterogeneity.

Heterogeneous compartmental models are also developed in the literature. The general idea is to divide the population into a finite (usually small) number of discrete departments. People in each compartment are assumed to behave identically, with full intra-compartmental random mixing [**40, 41, 42**]. This approach usually involves a large number of differential equations with fixed Markovian assumptions, and needs to be solved numerically to compute the model-predicted average trends of the epidemic.

In discrete-time heterogeneous models, a next-generation matrix is a useful tool in tracking an epidemic [**43**]. The next-generation matrix takes into account the anticipated inter-group contact rates, and uses them to calculate the number of infected individuals in each heterogeneous class of the population as the outbreak progresses. It is used in modeling early-stage epidemic growth and in estimating flu parameters, such as R_0.

Employing an approach quite different from the traditional compartmental models, in Chapter 2 of this thesis, we will expand on the heterogeneous modeling by Larson [**36**]. Larson introduced simple discrete-time modeling that took into account population heterogeneity without the need for the data that forms the next-generation matrix. Instead, he used the idea of proportional mixing, where each person is likely to get infected in proportion to his or her contact rate, as opposed to a posited specific rate of infection between each pair of classes [**36, 41**]. In Chapter 2 we will generalize Larson's approach to eliminate the need for a finite number of discrete classes of statistically identical individuals. Before moving on to this model, we will mention another approach to influenza modeling even though we will not be focusing on this approach here.

2.4.3 Network models

One other way to include contact heterogeneity into a modeling framework is to use networks to represent human to human contacts. In such models, a community is represented as a graph, in which nodes represent individuals and connecting edges represent possible contacts between them. The weighting of each edge corresponds to the probability that the disease progresses from one node to another. Such models have been popular recently because they inherently incorporate heterogeneity and stochasticity. Random graphs may be used to generate representations of realistic communities and stochastic simulation on such graphs paints a picture of typical outcomes of an epidemic. Graph theory may be used to describe epidemic spread, identify high risk individuals, and evaluate various mitigation strategies [**44, 45**]. These models have been particularly helped by the development of technology that allows researchers to track real life contacts through high-tech means (e.g. wireless signals) to produce a realistic representation of a human social network [**46**].

The one shortcoming of these models is that they are inherently complex. They are quite useful for evaluating pandemic scenarios or possible interventions, but they are not ideal tools for drawing general insights that could guide decision making in a pandemic situation. Similarly, the few equations that characterize these models usually do not reflect much on how the population changes as the outbreak continues.

3. Research summary

In this work we will focus on the theoretical and applied aspects of pandemic influenza modeling. In Chapters 2 and 3 we will introduce a generalized version of the heterogeneous community spread model. The new model is more analytically tractable as well as more realistic than currently used influenza spread models. It replaces the need for many coupled differential equations currently used to model

influenza spread in a heterogeneous population. We will include insights about the changes in the population characteristics as an outbreak develops with time, as well as some theoretical results that can be proven about an outbreak developing in a heterogeneous population. In Chapter 3 we will discuss lower and upper bounds on vaccines needed to achieve various levels of $R(t)$ including herd immunity, and the different vaccination strategies that can be used to achieve those bounds in a heterogeneous population. In Chapter 4, we will switch to the more data-driven aspects of this thesis and discuss the effectiveness of vaccines administered during the 2009-2010 H1N1 pandemic. In particular, we will focus on the effect of time on how useful the vaccines have been in mitigating the epidemic in the US. In this chapter we used a simple, homogeneous model for motivation of the in depth analysis in Chapter 5. In Chapter 5 we introduce some dynamic heuristic algorithms to change vaccine allocation decision-making and improve the effectiveness of vaccines in a multi-region area such as the United States. Finally, in Chapter 6 we will zoom back to a smaller community such as a single US state. Prior to receiving vaccines, the community must utilize non-pharmaceutical interventions effectively to manage the outbreak until vaccines arrive. In its entirety, this work will provide an overview of new methodologies that we hope will significantly improve our ability to fight influenza spread on national and local levels.

References

1. Karesh W, Cook R. The Human-Animal Link. *Foreign Affairs*. August 2005. Available at: http://www.foreignaffairs.com/articles/60821/william-b-karesh-and-robert-a-cook/the-human-animal-link. Accessed January 22, 2012.
2. Malakoff D, Enserink M. In Dramatic Move, Flu Researches Announce Moratorium on Some H5N1 Flu Research, Call for Global Summit. *ScienceInsider*. January 2012.
3. Nigmatulina K. Modeling and Responding to Pandemic Influenza: Importance of Population Distributional Attributes and Non-Pharmaceutical Interventions (Doctoral dissertation).
4. U.S. Department of Health & Human Services. About Pandemics. *FLU.gov*. Available at: http://www.flu.gov/pandemic/about/index.html#comparison. Accessed January 10, 2012.
5. World Health Organization. Pandemic influenza vaccine manufacturing process and timeline. *World Health Organization*. August 6, 2009. Available at: http://www.who.int/csr/disease/swineflu/notes/h1n1_vaccine_20090806/en/index.html. Accessed January 12, 2012.
6. Centers for Disease Control and Prevention. Seasonal Flu Shot Questions & Answers. Available at: http://www.cdc.gov/FLU/about/qa/flushot.htm. Accessed January 6, 2011.
7. Maugh T. Cell-culture influenza vaccine proves effective, could speed production. *Los Angeles Times*. February 2011.
8. Rappuoli R. Cell-Culture-Based vaccine production: technological options. *The Bridge*. 2006;36(3):25-30.
9. U.S. Department of Health and Human Services. First U.S. cell-based flu vaccine plant set for dedication
10. Taubenberger J, Morens D. 1918 Influenza: the Mother of All Pandemics. *Emerg Infect Dis [serial on the Internet]*. January 2006;12(1).
11. GlobalSecurity.org. 1957 Asian Flu Pandemic. Available at: http://www.globalsecurity.org/security/ops/hsc-scen-3_pandemic-1957.htm. Accessed January 22, 2012.
12. U.S. Department of Health & Human Services. History of flu Pandemics. Available at: http://www.pandemicflu.gov/pandemic/history/index.html. Accessed January 6, 2012.
13. McKibbin W, Sidorenko A. Global Macroeconomic Consequences of Pandemic Influenza. *Lowy Institute for International Policy, Sidney*. February 2006.
14. Lo J, Tsang T, Leung Y, Yeung E, Wu T, Lim W. Respiratory Infections during SARS Outbreak, Hong Kong 2003. *Emerging Infectious Diseases*. November 2005;11(111):1738-1741.
15. Centers for Disease Control and Prevention. The 2009 H1N1 Pandemic: Summary Highlights, April 2009-April 2010. June 16, 2010. Available at: http://www.cdc.gov/h1n1flu/cdcresponse.htm. Accessed January 12, 2011.
16. World Health Organization. World now at the start of 2009 influenza pandemic. June 11, 2009. Available at: http://www.who.int/mediacentre/news/statements/2009/h1n1_pandemic_phase6_20090611/en/index.html. Accessed January 12, 2011.
17. World Health Organization. Pandemic (H1N1) 2009 - update 112. May 23, 2010. Available at: http://www.who.int/csr/don/2010_08_06/en/index.html. Accessed January 20, 2011.
18. Centers for Disease Control and Prevention. CDC Estimates of 2009 H1N1 Influenza Cases, Hospitalizations and Deaths in the United States, April 2009 – March 13, 2010. April 19, 2010. Available at: http://www.cdc.gov/h1n1flu/estimates/April_March_13.htm. Accessed January 20, 2012.

19. Centers for Disease Control and Prevention. Press Briefing Transcripts: Media Briefing: Update on 2009 H1N1 Flu. September 3, 2009. Available at: http://www.cdc.gov/media/transcripts/2009/t090903.htm. Accessed January 20, 2011.
20. Centers for Disease Control and Prevention. 2009 H1N1 Vaccine Doses Allocated, Ordered, and Shipped by Project Area. February 2, 2010. Available at: http://www.cdc.gov/h1n1flu/vaccination/vaccinesupply.htm. Accessed January 20, 2011.
21. Hopkins J. H1N1 After-action reports: lessons on vaccine distribution. *MIT ESD Working Paper.* 2011.
22. Jefferson T, Jones M, Doshi P, Del Mar C. Neuraminidase inhibitors for preventing and treating influenza in healthy adults. *Cohrane Database Syst Rev.* 2006.
23. Greenberg M, Lai M, Hartel G, Wichems C. Response after one dose of a monovalent influenza A (H1N1) 2009 vaccine - preliminary report. *N Engl J med.* 2009;361:2405-2413.
24. Diekmann O, Heesterbeek J, Metz J. On the definition and the computation of the basic reproduction ration R0 in models for infectious diseases. *J. Math. Biol.* 1990;35:503-522.
25. Lipsitch M, Mills C, Robbins J. Transmissibility of 1918 pandemic influenza. *Nature.* December 2004;432:904-906.
26. Vynnyvky E, Edmunds W. Analyses of the 1957 (Asian) influenza pandemic in the United Kingdom and the impact of school closures. *Epidemiol Infect.* February 2008;136(2):166-179.
27. Anderson R, Fraser C, Ghani A, Donnelly C. Epidemiology, transmission dynamics and control of SARS the 2002-2003 epidemic. *Phil. Trans. R. Soc.* July 2004;359(1447):1091-1105.
28. White L, Wallinga J, Finelli L, Reed C. USA, Estimation of the reproductive number and the serial interval in early phase of the 2009 influenza A/H1N1 pandemic in the. *Influenza and Other Respiratory Viruses.* September 2009;3(6):267-276.
29. Anderson R. The concept of herd immunity and the design of community-based immunization programmes. *Vaccine.* 1992;10:928-35.
30. Anderson R, May R. *Infectious Diseases of Humans: Dynamics and Control.* Oxford: Oxford University Press; 1991.
31. Kermack W, McKendrick A. A contribution to the mathematical theory of epidemics. *Proc. Roy. Soc. Lond.* 1927;115:700-721.
32. Typhoid Mary Dies of a Stroke At 68. Carries of Disease, blamed for 51 Cases and 3 Deaths, but She Was Held Immune Services this Morning Epidemic is Traced. *New York Times.* November 1938.
33. Keeling N, Rohani P. *Modeling infectious diseases in humans and animals.* Princeton, NJ: Princeton Univ Press; 2007.
34. Lloyd-Smith J, Schreiber S, Kopp P, Getz W. Superspreading and the effect of individual variation on disease emergence. *Nature.* 2005;438:355-359.
35. Nigmatulina K, Larson R. Living with influenza: Impacts of government imposed and voluntarily selected interventions. *EJOR.* 2009;195(2):613-627.
36. Larson R. Simple models of influenza progression within a heterogeneous population. *Operations Research.* 2007;55(3):399-412.
37. Mikolajczyk R, Akmatov M, Kretzschmar M. Social contacts of school children and the transmission of respiratory-spread pathogens. *Epidemiol Infect.* Jun 2008;6(813-22):136.
38. Fu Y. Measuring personal networks with daily contacts: A single-item survey question and the contact diary. *Soc. Networks.* 2005;27:169-186.
39. Mossong J, Hens N, Jit M, Beutels P, Auranen K. Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS Med.* 2008.

40. Hethcote H, Van Ark J. Epidemiological models for heterogeneous populations: proportionate mixing, parametere estimation, and immunization programs. *Mathematical Bioscience.* 1987;84(1):85-118.
41. Nold A. Heterogeneity in Disease Transmission Modeling. *Mathematical Biosciences.* 1980;52:227-240.
42. van den Driessche P, Watmough J. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Mathematical Biosciences.* 2002;180:29-48.
43. Diekmann O, Heesterbeek J. *Mathematical Epidemiology of Infectious Diseases: Model Building Analysis and Interpretation.* New York; 2000.
44. Ames G, George D, Hampson C, Kanarek A, McBee C. Using network properties to predict disease dynamics on human contact networks. *Proc R Soc B.* December 2011;278(1724):3544-3550.
45. Christley R, Pinchbeck G, Bowers R, et al. Infection in social networks: using network analysis to identify high-risk individuals. *Am J Epidemiol.* November 2005;162(10):1024-1031.
46. Salathe M, Kazandjieva M, Lee J, Levis P, Feldman M, Jones J. A high-resolution human contact network for disease transmission. *PNAS.* December 2010;107(51):22020-22025.

Chapter II: Continuous-attribute heterogeneous population model

1. Introduction

Despite a large body of research dealing with mathematical modeling of influenza spread, there is still no widely accepted way to inform decision making in the case of a global event such as the 2009-2010 H1N1 "Swine Flu" pandemic. The H1N1 influenza outbreak demonstrated that limited information and short time allowances during a pandemic make it difficult to model accurately the spread of influenza and to make appropriate and effective decisions with regard to public policy and vaccine distribution.

As a pandemic evolves, decision makers receive aggregate statistics in the form of the number of people reporting to physicians with flu-like symptoms, number of related hospital admissions, number of flu-related deaths, and number of vaccinations administered. Yet, aggregate statistics hide the fact that early transmission and propagation of the disease are driven largely by particular segments of the population: (1) those who are highly active in daily face-to-face encounters; (2) those who are overly prone to become infected given exposure; and/or (3) those who shed virus and spread the disease much more than average. Any person can be characterized along a spectrum of these three attributes: social activity, proneness to infection, and proneness to shed virus and spread infection. Those who are at the "right-hand-tails" of one or more of these distributional attributes play a significant role in the early spread of the disease. Such individuals, due to early infection, drop out of the susceptible population near the middle and almost certainly by the end of the outbreak.

The correlation between these types of heterogeneity also plays a role in the characterizing the speed and duration of infection. It is reasonable to assume that people who are most susceptible to infection are also those that are most likely to spread it to others. Those with high levels of both sources of heterogeneity are likely to get infected first and quickly leave the susceptible population. However, a negative correlation between susceptibility and infectiousness will cause the infection to spread much slower and be sustained for a longer period of time within a community. To understand the dynamics of flu spread, or the spread of any human-to-human infectious disease, one cannot ignore such population heterogeneities.

In this chapter, we generalize the heterogeneous-population, influenza-spread model developed by Larson [**1**]. We improve analytical tractability by incorporating a more general, continuous form of population heterogeneity. In this chapter we generalize Larson's approach to modeling influenza spread in heterogeneous population by eliminating the need for a finite number of discrete classes of statistically identical individuals. Instead, we introduce a continuous distribution for all the parameters in question, in essence employing an infinite number of classes. Eventually we deal with all three attributes introduced above: social activity, proneness to infection, and proneness to spread infection. Initially we focus only on contact rates, the measure of social activity. The model relies on just a few equations that define the state of infection at a given time. We present numerical examples showing how various segments of the population are affected throughout the outbreak and the importance of determining a good starting distribution for contact rates. We also discuss the calculation of R_0, and particularly the effects of heterogeneity on $R(t)$. We conclude by using contact data from four European countries to demonstrate the empirical uses of our model. In addition to generalizing the traditional compartmental model by using a spectrum of heterogeneity as opposed to arbitrary classes, our model introduces closed-form equations to describe the state of the epidemic in a community at different points in time. These

equations can be used to mathematically prove results regarding the changing characteristics of the population and the progression of the virus within the community. We provide some such proofs of in Section 2.3.

2. Generalized model

Larson employs a discrete-time model where the unit of time is a generation of influenza, or a "day". A generation for the purposes of this model is the period of infection during which a person is infectious and actively interacting in society. Larson's model assumes a finite number of classes of people, differing by their activity level. Specifically,

λ_C = Poisson rate of human contacts per day of an individual in class C

p = probability that a susceptible person becomes infected after a contact with an infectious individual

During each day i, a person may be classified as susceptible, infected, or immune. A person's state may change on day $i + 1$, where each susceptible person independently may become infected or remain susceptible, an infected person recovers and becomes immune, and each immune person remains immune. For simplicity here, we assume that there are no influenza-related deaths during this outbreak.

We generalize on this model by making the Poisson contact rate λ a random variable, and assuming a distribution of λ over the entire population. For a given person in the population, we sample from the λ-distribution and find $\lambda = \lambda_0$. Then, that person has a number of potentially infectious daily contacts n, selected from a Poisson distribution with mean λ_0. Here λ_0 corresponds to the individual's "class". We introduce the following notation:

$f(\lambda)$ = population distribution of λ, the individual Poisson rate of contacts per day[1]

$f_i^I(\lambda)$= distribution of λ, the individual Poisson rate of contacts per day, in the infectious population on day i. That is, if we were to select a single person, at random, from the infectious population, $f_i^I(\lambda)$ would represent the distribution corresponding to that person's daily contact rate, λ.

$f_i^S(\lambda)$= distribution of λ, the individual Poisson rate of contacts per day, in the susceptible population on day i. That is, if we were to select a single person at random from the susceptible population, $f_i^I(\lambda)$ would represent the density function corresponding to that person's daily contact rate, λ.[2]

N_i^I= number of infectious individuals on day i.

N_i^S= number of susceptible individuals on day i.

N = total number of people in a population.

---

[1] We assume here that $f(\lambda)$ is a continuous probability density function. The arguments in this section may also be applied to discrete distributions, for which $f(\lambda)$ represents the probability that the Poisson contact rate equals λ. We will use $f(\lambda)$ and the word "distribution" in both contexts to refer to either the probability density function or the probability mass function and use summations instead of integrals where appropriate.

[2] Some individuals may change their contact rates during their infectious generation, e.g. by staying at home while

Once we know the initial conditions, we can model a typical epidemic curve and approximate the social-contact characteristics of the population by sequentially applying steps similar to those used by Larson:

Day i:

We start with the known values of all our quantities of interest during day i: $N_i^S, N_i^I, f_i^S(\lambda)$ and $f_i^I(\lambda)$. We now show how to find all the necessary values of $N_{i+1}^S, N_{i+1}^I, f_{i+1}^S(\lambda)$ and , $f_{i+1}^I(\lambda)$. As per Larson:

β_i = the day i probability that a random person's next interaction is with an infected person.

We now consider the total expected number of contacts that all infectious people make on day i:

$$\int_0^\infty N_i^I \lambda\, f_i^I(\lambda)\, d\lambda = N_i^I \int_0^\infty \lambda\, f_i^I(\lambda)\, d\lambda = N_i^I E_i^I(\lambda)$$

The mean total number of day i contacts by infectious people is the product of the average number of contacts by an infectious person and the number of infectious people in the population. Likewise, the mean total number of human contacts each day is

$$\int_0^\infty N\lambda f(\lambda)\, d\lambda = N \int_0^\infty \lambda f(\lambda)\, d\lambda = NE(\lambda)$$

and, with no behavioral changes or deaths, is assumed to be constant throughout the outbreak. Therefore, the probability that any interaction is with an infectious person, is simply the ratio,

$$\beta_i = \frac{N_i^I E_i^I(\lambda)}{NE(\lambda)}. \tag{1}$$

We next define the day-specific infection probability

$p_i(I|\lambda)$ = probability that a susceptible individual with contact rate λ becomes infected on day i

which Larson showed to be

$$p_i(I|\lambda) = 1 - e^{-\lambda p \beta_i} \tag{2}$$

We now have all the necessary tools to find $N_{i+1}^S, N_{i+1}^I, f_{i+1}^S(\lambda)$ and $f_{i+1}^I(\lambda)$.

$$N_{i+1}^I = \int_0^\infty p(I|\lambda) N_i^S f_i^S(\lambda) d\lambda = N_i^S \int_0^\infty p(I|\lambda) f_i^S(\lambda) d\lambda \tag{3}$$

$$N_{i+1}^S = N_i^S - N_{i+1}^I \tag{4}$$

The λ-distributions over different segments of the population are slightly more complex. The distribution of contact rates of infectious people changes on day $i + 1$ from what it was on day i. On day $i + 1$, it is the distribution of contact rates of those people who were susceptible and became infected on day i.

$$f_{i+1}^I(\lambda) = f(\lambda | \textit{individual was infected on day } i) = \frac{p_i(I|\lambda) * f_i^S(\lambda)}{\int_0^\infty p_i(I|\lambda) * f_i^S(\lambda)\, d\lambda} \tag{5}$$

Now for the susceptibles: their day $i+1$ contact-rate distribution is that of those who were susceptible on day i and who were not infected on day $i+1$.

$$f_{i+1}^{S}(\lambda) = f(\lambda | individual\ was\ not\ infected\ on\ day\ i) = \frac{\left(1 - p_i(I|\lambda)\right) * f_i^S(\lambda)}{\int_0^\infty \left(1 - p_i(I|\lambda)\right) * f_i^S(\lambda)\, d\lambda} \quad (6)$$

These formulae describe entirely the state of the epidemic in the community on day i. By starting with some known boundary values of parameters and iteratively applying these rules over time, we obtain a model of complete epidemic dynamics within a continuous heterogeneous context.

Model Summary:

Starting with the following quantities on day i: $N_i^S, N_i^I, f_i^S(\lambda)$ and $, f_i^I(\lambda)$

Compute the intermediate quantities:

$$\beta_i = \frac{N_i^I E_i^I(\lambda)}{NE(\lambda)}$$

$$p_i(I|\lambda) = 1 - e^{-\lambda p \beta_i}$$

Compute the relevant quantities on day $i+1$:

$$N_{i+1}^I = N_i^S * \int_0^\infty p(I|\lambda) f_i^S(\lambda) d\lambda$$

$$N_{i+1}^S = N_i^S - N_{i+1}^I$$

$$f_{i+1}^I(\lambda) = \frac{p_i(I|\lambda) * f_i^S(\lambda)}{\int_0^\infty p_i(I|\lambda) * f_i^S(\lambda)\, d\lambda}$$

$$f_{i+1}^S(\lambda) = \frac{\left(1 - p_i(I|\lambda)\right) * f_i^S(\lambda)}{\int_0^\infty \left(1 - p_i(I|\lambda)\right) * f_i^S(\lambda)\, d\lambda}$$

3. Model details

3.1 Numerical examples:

3.1.1 Uniform distribution

We first model a population of a small hypothetical community of 10,000 people. As a boundary condition, consider $f_0(\lambda)$ to be uniform between 0 and 40. Iteratively running our model, results are summarized in Figures 1 and 2 and Table 1. Additional details are provided in Appendix C.

Outbreak Statistics	
Duration of Outbreak	17 days
Total Number of Infections	7,168
$E_1^I(\lambda)$	26.7
$E_{17}^I(\lambda)$	18.2
$E_{17}^S(\lambda)$	10.4

Table 1: Outbreak statistics. The initial distribution of λ is uniform between 0 and 40.

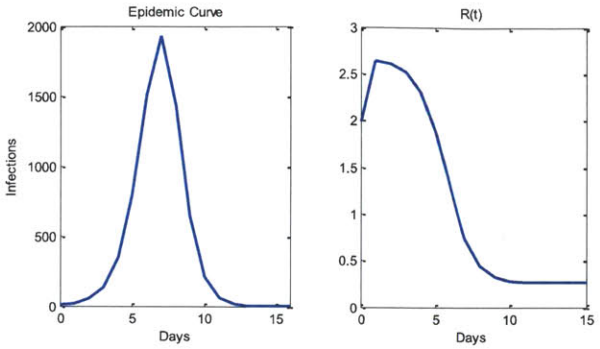


Figure 1: Under an initial uniform distribution of daily contact rates, the epidemic curve and R(t) of the outbreak. R(t) is computed by dividing the number of infectious people at time t+1 by the number of infectious people at time t.

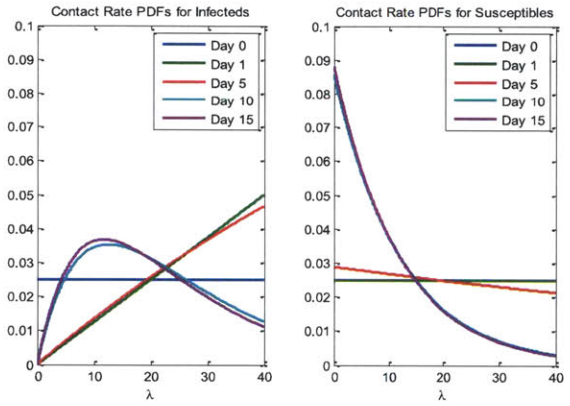


Figure 2: Contact rate characteristics for the infectious and susceptible subsections of the population at different times during the outbreak.

Talking through the example, on day 0, we introduced 10 infectious individuals into the population, each having a λ value drawn from the uniform distribution over 0 to 40. However, on day 1, the distribution of contact rates of infectious persons resembles a straight line with slope $\frac{1}{8}$. That is, if on day 1 we were to select a person at random from the infectious population, that person would be much more likely to have a high daily contact rate, than a low one. In fact, the expected contact rate by an infectious person on day 1 is 26.7 (Table 1). Note, also, that it is impossible for a person from the infectious population to have a daily contact rate of 0 because a person with no contacts has no chance of being infected. The infected population becomes "high-activity" heavy. This is consistent with Larson's model and supports the conclusion that high-activity people are the drivers of the infection at the beginning of the outbreak. As the outbreak continues, however, the high-activity tail dips down, as high-activity individuals become infected early, become immune and can no longer be infected. Note that, on the contrary, the distribution of contact rates in the susceptible population becomes more and more "low-activity" heavy. On the last day of the outbreak, the average contact rate in the susceptible population is 10.4, almost half of the original average contact rate of 20.

3.1.2 Poisson distribution

We now consider the same community with different initial conditions. We propose a more realistic, Poisson distribution for $f(\lambda)$ with mean 20. Firstly, we are now dealing with a discrete distribution for λ as opposed to a continuous density function. There are some notable differences from the first example.

Outbreak Statistics	
Duration of Outbreak	21 days
Total Number of Infections	7,865
$E_1^I(\lambda)$	21.0
$E_{20}^I(\lambda)$	19.5
$E_{\infty}^S(\lambda)$	18.4

Table 2: Outbreak statistics. The initial distribution of λ is Poisson with mean 20.

Here too, we see that the characteristics of the subsections of the population change as the outbreak continues, although the effect is much smaller. On the first day, the expected contact rate in the infectious population is 21, a slight increase from 20 (Table 2).

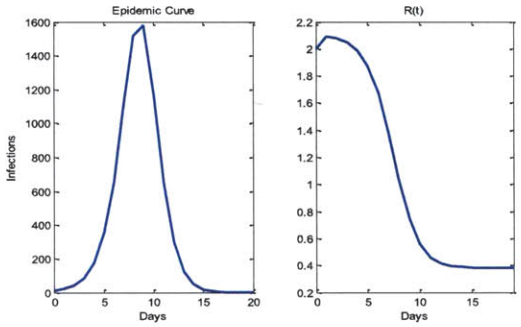


Figure 3: The epidemic curve and R(t) of the outbreak. The initial distribution is Poisson with mean 20.

Toward the end, the expected contact rates in different subsections of the population change slightly, but generally the distributions remain similar to the original contact rate distribution. This behavior is noticeably different from the changing distributions in the uniform distribution case.

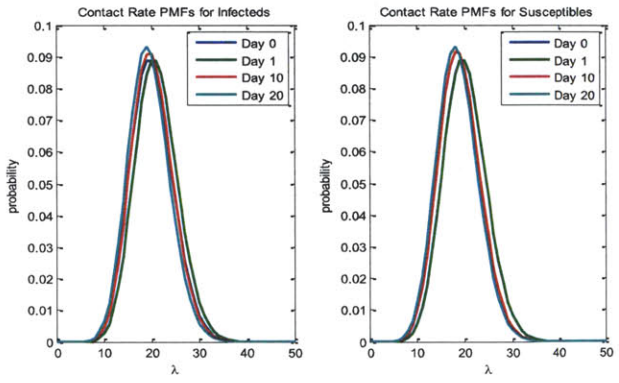


Figure 4: Contact rate characteristics for the infectious and susceptible subsections of the population at different times during the outbreak. The initial distribution is Poisson with mean 20.

3.2 Sensitivity to Original Distribution

The two numerical examples show that the original distribution of contact rates plays a major role in the evolution of all four of our relevant quantities. The total expected number of infections in the outbreak is 10% larger in the case of a Poisson original distribution than in the case of a uniform distribution. In the case of the uniform distribution, the susceptible and the infectious subsections of the populations have extremely different contact rate characteristics. The susceptible population is low-activity heavy, while the contact rate distribution for the infectious population is a concave function with low probability at the left tail and decreasing probability at the right.

Calculation of R_0

The differences between the two cases are particularly evident in the calculation of the basic reproductive number, R_0. Recall that R_0 is defined as the expected number of secondary infections caused by a typical infected individual in a fully susceptible population **2**. On day 0, since almost everybody is susceptible, we can assume that we are in a fully susceptible population. There is some uncertainty as to the exact manner of calculating R_0. The cause of ambiguity is the term "typical". Suppose we pick a person at random from our infectious population on day 0. In both the case of the uniform and the Poisson original distributions, the average contact rate for a random person, call her patient zero, would be 20. If the Poisson contact rate is 20, the average number of patient zero's contacts is also 20. Since each contact has a .1 probability of spreading infection, this interpretation of the definition would give us an R_0 of 2. If R_0 is the same for both original distributions of contact rates, we would expect the total number of infections to be approximately the same, but they are not.

Instead, we recommend the logic followed by network models in the established epidemiological literature [**2, 3, 4**], where R_0 is calculated as an expected value of a distribution that is weighted by the magnitude of the contact rate. That is, a person in the infectious population with a high contact rate should be weighed proportionally higher, as that person is more likely to acquire and then spread the infection than a person with a low contact rate. In other words, our "typical" person would no longer be selected at random, but according to a distribution $f_R(\lambda)$. *Such a distribution will no longer be sampling from the population of people, but rather population of contacts*. Consider each person dropping a piece of paper that is marked with that person's λ value on the floor during each and every contact. We then pick R_0 by drawing a random piece of paper. This new distribution would favor high activity people proportionally to their activity level, then $f_R(\lambda) = \frac{f_0^I(\lambda)*\lambda}{E_0^I(\lambda)}$.

Note that this is exactly the formula used to describe the concept of time-average duration of an interval in a renewal process [**5**]. Using this scenario, we can calculate R_0:

For the uniform distribution case,

$$R_0 = \int_0^{40} p\frac{\lambda^2}{40E(\lambda)}d\lambda = \frac{40^2}{30*20} = 2.7$$

For the Poisson case,

$$R_0 = \sum_0^{\infty} p\frac{\lambda^2}{E(\lambda)}\frac{20^\lambda}{\lambda!}e^{-20} = \frac{420}{200} = 2.1$$

Note that even though the uniform distribution has a larger value of R_0, it actually results in fewer total infections! This phenomenon is caused by the fact that the uniform distribution has a higher variance than a Poisson distribution with the same mean. In the uniform case, there are fewer very high-activity people who are depleted much more quickly than in the Poisson case. Once the high-activity people are gone, the outbreak quickly subsides resulting in a shorter total outbreak time and fewer total infections. It is important to note, also, that the effect of depletion of high-activity individuals from the population comes from the fact that high-activity people are more susceptible to infection; we show in Section 3.4.2, that the fact that these people are also more infectious does not contribute to this effect. Andersson and Britton confirm, using compartmental model methods, that epidemics are particularly devastating when susceptibility does not vary, i.e. there is no early depletion [**6**].

This is another reminder that an influenza model is a simplification of a real life problem that is too complex to describe fully. To avoid critical mistakes, healthcare decision-makers must be extremely cautious in choosing initial values of the model's parameters such as the initial distribution of contact rates to avoid critical mistakes. Many published heterogeneous modeling published papers describe deriving R_0 in a manner consistent with the second approach above, arguing that R_0 is the mean of a distribution that is skewed toward high-activity individuals [**2**, **3**, **4**]. However, one needs to be careful when transferring this idea into simulation modeling. Often, we find that flu simulations are seeded by selecting a few typical "patient zeros" from the population to enter the susceptible population in an infectious state. These typical people are selected to be representative of the overall general population, and their average contact rate is used to calibrate the simulation to known R_0 parameters. "Patient zeros" that are selected in this way, result in an underestimate of the true value of R_0. For the calculation of R_0 that includes heterogeneity in contact, infectivity, and susceptibility levels as well as the possibility of mortality, see Appendix D.

3.3 Common properties

Despite the fact that the starting distribution plays a significant role in the way an epidemic curve for the outbreak is calculated using this model, certain properties remain true regardless of original distribution.

Theorem 1: Tilting of contact rate distributions in susceptible and infected populations.

Starting with any initial conditions, $f_0^I(\lambda)$, and $f_0^S(\lambda)$,

for any $i > 0$

$$f_i^I(\lambda) = C_i^I \left(e^{-\lambda p \sum_{j=0}^{i-2} \beta_j} - e^{-\lambda p \sum_{j=0}^{i-1} \beta_j} \right) f_0^S(\lambda) \tag{7}$$

and

$$f_i^S(\lambda) = C_i^S \left(e^{-\lambda p \sum_{j=0}^{i-1} \beta_j} \right) f_0^S(\lambda) \tag{8}$$

where C_i^I and C_i^S are normalizing constants. Equations (7) and (8) imply that as the outbreak progresses, contact rate distributions become tilted distributions with tails decaying exponentially faster than in the original contact rate distributions. Note also, that the initial distribution of infectious individuals, $f_0^I(\lambda)$, plays no role in the distributions for any generation i>0.

Proof:[3]

We next introduce a result that will be useful in describing the behavior of distributions as the outbreak progresses in more detail.

Lemma 1: For any constants $0 < a \leq 1$, $0 < b \leq 1$, and $\lambda > 0$, the function

$$g(\lambda) = \left(\frac{\left(1 - e^{-\lambda a}\right)}{(e^{\lambda b} - 1)}\right)$$

is monotonically decreasing and convex in λ.

Theorem 2: Stochastic ordering of susceptible and infectious populations.

Let Λ_i^Sbe the random variable distributed as $f_i^S(\lambda)$, and similarly Λ_i^S as $f_i^S(\lambda)$.

Then, as long as $f_0^S(\lambda), f_0^I(\lambda)$ are non-deterministic, for any $i > 0$,

$$\Lambda_{i+1}^S <_{ST} \Lambda_i^S \quad (9)$$
$$\Lambda_i^S <_{ST} \Lambda_i^I \quad (10)$$
$$\Lambda_{i+1}^I <_{ST} \Lambda_i^I \quad (11)$$

Where $X <_{ST} Y$ implies that X is strictly smaller than Y in the usual stochastic order [**7**].

Stochastic ordering of the various distributions as time goes on provides us with an analytical expression of our intuition. That is, according to equation (9), the contact rates in the susceptible population become "smaller" with time in the stochastic ordering sense. In other words, the susceptible population becomes more and more densely packed with inactive individuals as the high activity people become infected. Similarly with the infectious population.

Theorem 3: Depletion of high-activity individuals in the susceptible and infectious populations.

As long as $f_0^S(\lambda), f_0^I(\lambda)$ are non-deterministic, for any $i > 0$,

$$f_i^S(\lambda) > f_{i+1}^S(\lambda) \; for \; \lambda \geq E_i^S(\lambda) \quad (12)$$
$$f_i^S(\lambda) < f_{i+1}^I(\lambda) \; for \; \lambda \geq E_i^S(\lambda) \quad (13)$$
$$f_i^I(\lambda) > f_{i+1}^I(\lambda) \; for \; \lambda \geq E_i^I(\lambda) \quad (14)$$

These are true regardless of initial contact rate distribution. Equation (12) implies that that after the first generation, the susceptible population becomes increasingly low-activity heavy. In equation (14) we show that that the distribution of infected people also becomes more low-activity heavy with time since high-activity people become depleted from the population. For illustration, see Figure 5. We note that since we model infection spread in discrete time increments, the high-activity people get infected in large numbers within each generation. However, in a real-life scenario, these generations overlap and we do not need to "wait for a generation to end" for new people to get infected. As a result, the depletion of high activity individuals should occur even more rapidly in a continuous-time, real-world outbreak. Finally, we note that expectation of contact rate distributions in both the susceptible and infected subsections of the population is decreasing with time.

[3] All proofs in this section are presented in Appendix A

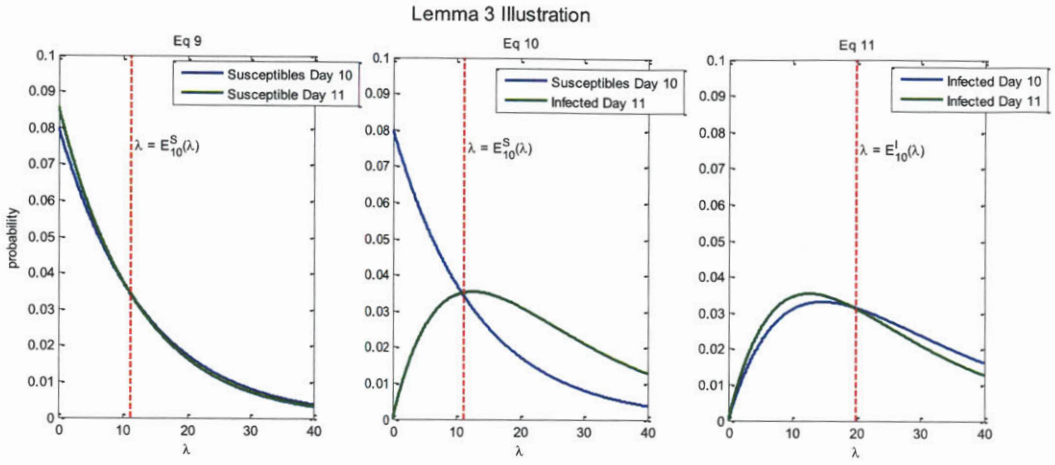


Figure 5: Illustration to theorem 3: the probability density functions of contact rates for the susceptible and infected populations on days 10 and 11. The initial distribution of contact rates was uniform, ranging from 0 to 40.

Corollary 1: Overall decrease in expected contact rates. For $i > 0$, and non-deterministic $f_0^S(\lambda), f_0^I(\lambda)$,

$$E_i^S(\lambda) < E_{i+1}^S(\lambda) \tag{15}$$
$$E_i^I(\lambda) < E_{i+1}^I(\lambda) \tag{16}$$

3.4 Further generalizations

The methodology developed above can be extended to include heterogeneity due to vulnerabilities of different groups as well as the effects of changing social behaviors during an ongoing pandemic.

3.4.1 Mortality:

So far in our model, we have assumed that all infected people recover and rejoin the population within one generation. However, any devastating pandemic influenza that we fear is likely to be deadly; some infected individuals are likely to die of the infection. Moreover, the likelihood of death from an infection will almost certainly not be uniform throughout the population. The Centers for Disease Control and Prevention (CDC) suggests that pregnant women and the elderly are at most risk from death during a seasonal influenza outbreak [**8**, **9**]. Meanwhile, one of the characteristics of a particularly dangerous novel influenza strain pandemic, is the heightened mortality rates of the people in their 20's [**10**].

To add this layer of complexity, we add a probability of death after an infection, q. That is, an infectious person dies at the end of her infectious period with probability q, and rejoins the population with probability $(1 - q)$. Moreover, this probability cannot be constant, as certain people are at more risk than others. We must impose a probability distribution on q that is not independent from λ.

The quantities of interest become:

$f_i(\lambda, q)$ = the joint distribution of λ, the Poisson rate of contacts per day, and q, the probability of death given infection, in the entire population alive at the start of day i.

$f_i^I(\lambda, q)$= the joint distribution of λ, the Poisson rate of contacts per day, and q, the probability of death given infection, in the infectious population on day i.

$f_i^S(\lambda, q)$= the joint distribution of λ, the Poisson rate of contacts per day, and q, the probability of death given infection, in the susceptible population on day i.

N_i = the number of people alive in the population at the start of day i.

N_i^I, N_i^S as before.

Replicating equations (1)-(6):

The total number of contacts made by infectious people on day i remains as before,

$$\int_0^1 \int_0^\infty N_i^I \lambda f_i^I(\lambda, q)\, d\lambda\, dq = N_i^I * \int_0^1 \int_0^\infty \lambda f_i^I(\lambda, q)\, d\lambda\, dq = N_i^I E_i^I(\lambda)$$

However, the total number of interactions made by the entire population on day i, now needs to account for the mortality rate of the population.

$$\int_0^1 \int_0^\infty N_i \lambda f_i(\lambda, q)\, d\lambda\, d\, q = N_i * \int_0^1 \int_0^\infty \lambda f_i(\lambda, q)\, d\lambda\, dq = N_i E_i(\lambda)$$

We can now solve for β_i:

$$\beta_i = \frac{N_i^I E_i^I(\lambda)}{N_i E_i(\lambda)}. \tag{1a}$$

The dynamics of the spread of infection on day i remain unchanged as the deaths occur at the end of each generation; equations (3) -(6) remain unchanged except for the new distribution functions that account for death probabilities.

$$N_{i+1}^I = \int_0^1 \int_0^\infty p(I|\lambda) N_i^S f_i^S(\lambda, q) d\lambda\, dq = N_i^S \int_0^1 \int_0^\infty p(I|\lambda) f_i^S(\lambda, q) d\lambda\, dq \tag{3a}$$

$$N_{i+1}^S = N_i^S - N_{i+1}^I \tag{4a}$$

$$f_{i+1}^I(\lambda, q) = \frac{p_i(I|\lambda) * f_i^S(\lambda, q)}{\int_0^\infty p_i(I|\lambda) * f_i^S(\lambda, q)\, d\lambda} \tag{5a}$$

$$f_{i+1}^S(\lambda, q) = \frac{\big(1 - p_i(I|\lambda)\big) * f_i^S(\lambda, q)}{\int_0^\infty \big(1 - p_i(I|\lambda)\big) * f_i^S(\lambda, q)\, d\lambda} \tag{6a}$$

We are left to find the effects of the deaths on day i. Consider the probability that a person with contact rate λ and mortality probability q died in period i:

$$P_i(death\,|\lambda, q) = P_i(infection\,|\lambda, q) * q$$

$$P_i(infection|\ \lambda, q) = \frac{f_i^I(\lambda, q) * \frac{N_i^I}{N_i}}{f_i(\lambda, q)} = \frac{f_i^I(\lambda, q) * N_i^I}{f_i(\lambda, q) * N_i}$$

for all values of λ, q for which $f_i(\lambda, q) > 0$.

So the probability that a person with these parameters survived day i, is then

$$1 - \frac{f_i^I(\lambda, q) * N_i^I}{f_i(\lambda, q) * N_i} * q$$

We can then derive $f_{i+1}(\lambda, q)$

$$f_{i+1}(\lambda, q) = \frac{f_i(\lambda, q) * P_i(no\ death|\lambda, q)}{P(no\ death)} = \frac{f_i(\lambda, q) * \left(1 - \frac{f_i^I(\lambda, q) * N_i^I}{f_i(\lambda, q) * N_i} * q\right)}{1 - \frac{N_i^I E_i(q)}{N_i}} \tag{17}$$

And finally, assuming that deaths are independent of each other,

$$N_{i+1} = N_i - N_i^I \int_0^1 \int_0^\infty q f_i^I(\lambda, q)\, d\lambda\, dq = N_i - N_i^I E_i(q) \tag{18}$$

3.4.2 Variability in susceptibility and infectivity:

The second additional layer of complexity is also acknowledged by Larson. We have so far assumed that if a susceptible and infectious individual come into contact, regardless of their respective contact rates, the probability that a susceptible person becomes infected from the contact is p. However, particularly in recent times, there has been added concern about people who are inherently more susceptible to disease than others. For example, the recent H1N1 pandemic of 2009 was shown to be less likely to affect people over 65 [**11**]. The "Asian Flu" of 1957 reported similar tendencies, likely due to acquired immunity by the older generation. Similarly, some portion of the population includes "super-shedders;" these people are much more likely to spread the disease than the average person. These differences in people's ability to shed and be infected by the flu virus, cause variability in the parameter p.

Larson brings up this possibility by splitting p into two different parameters. We will change notation here to call these parameters r and s.

r = the infectiousness index of an individual. r can take on any value between 0 and 1; it describes the level of infectiousness of an individual, with 0 implying a person does not shed any virus, and 1 implying a person is maximally infectious.

s = the level of susceptibility of an individual. That is, the probability that a susceptible individual gets infected after an interaction with an infectious individual with infectiousness index $r = 1$.

Using this notation, if an infectious person with infectiousness index r and a susceptible person with susceptibility parameter s make contact on a given day, the probability that the susceptible person becomes infected is $p = r * s$.

We alter our epidemic parameters to include r and s:

$f_i(\lambda, q, r, s)$ = joint distribution of λ, q, r, and s in the entire population alive at the start of day i.

$f_i^I(\lambda, q, r, s)$ = joint distribution of λ, q, r, and s in the infectious population on day i.

$f_i^S(\lambda, q, r, s)$ = joint distribution of λ, q, r, and s in the susceptible population on day i.

N_i = number of people alive in the population at the start of day i.

N_i^I, N_i^S as before.

Once again we calculate equations analogous to equations (1)-(6), taking extra care with notation:

$\beta_i(r)$ = the probability that a random person's next interaction is with an infected person having infectivity parameter r, on day i.

$$\beta_i(r) = \frac{\int_0^1 \int_0^1 \int_0^\infty N_i^I \lambda f_i^I(\lambda, q, r, s)\, d\lambda\, dq d\, s}{\int_0^1 \int_0^1 \int_0^1 \int_0^\infty N_i \lambda f_i(\lambda, q, r, s)\, d\lambda\, d\, q dr ds} = \frac{N_i^I E_i^I(\lambda|r) f_i^I(r)}{N E(\lambda)} \quad (1b)$$

Let

$p_i(I|\lambda, q, s)$= Probability that a susceptible person with susceptibility parameter s and contact-rate parameter λ, becomes infected on day i (note that the person's r value is irrelevant) (2b)

$$p(I|\lambda, s) = 1 - e^{-\lambda s \int_0^1 \tilde{r} \beta_i(\tilde{r}) d\tilde{r}} = 1 - e^{-\lambda s \frac{N_i^I E_i^I(\lambda r)}{N E(\lambda)}}$$

See Larson, [**1**] supplementary materials for derivation.

Then we calculate the usual quantities:

$$N_{i+1}^I = N_i^S \int_0^1 \int_0^1 \int_0^1 \int_0^\infty p(I|\lambda, s) * f_i^S(\lambda, q, r, s) d\lambda\, dq dr ds \quad (3b)$$

$$N_{i+1}^S = N_i^S - N_{i+1}^I \quad (4b)$$

$$f_{i+1}^I(\lambda, q, r, s) = \frac{p_i(I|\lambda, q, s) * f_i^S(\lambda, q, r, s)}{\int_0^1 \int_0^1 \int_0^1 \int_0^\infty p(I|\lambda, q, s) * f_i^S(\lambda, q, r, s) d\lambda\, dq dr ds} \quad (5b)$$

$$f_{i+1}^S(\lambda, q, r, s) = \frac{\left(1 - p_i(I|\lambda, q, s)\right) * f_i^S(\lambda, q, r, s)}{\int_0^1 \int_0^1 \int_0^1 \int_0^\infty \left(1 - p(I|\lambda, q, s)\right) * f_i^S(\lambda, q, r, s) d\lambda\, dq dr ds} \quad (6b)$$

$$f_{i+1}(\lambda, q, r, s) = \frac{f_i(\lambda, q, s) * \left(1 - \frac{f_i^I(\lambda, q, r, s) * N_i^I}{f_i(\lambda, q, r, s) * N_i} * q\right)}{1 - \frac{N_i^I E_i(q)}{N_i}} \quad (14b)$$

and

$$N_{i+1} = N_i - N_i^I E_i(q) \quad (15b)$$

For the equivalents of Theorems 1 to 3 for the generalized model see Appendix E.

3.4.3 Behavioral changes:

We have so far been working with a static model, where people behave in the same way every day and do not change their routine as influenza enters their community. However, previous outbreaks such as

the 2002 SARS epidemic show evidence that people change their behavior throughout their outbreak even if their region has not yet been affected by the epidemic [**12, 13**]. These behavioral changes have been shown to make an impact on the spread of infection. A good model, therefore, should take into account the fact that people will try to limit their daily activity in the event of a pandemic scare [**12**]. This can be easily incorporated our model by introducing a transformation function $G_i(\lambda)$ that represents the effect of behavioral changes in λ on day i. If G is invertible, the distribution functions can be easily determined at the end of each day.

On day i consider the quantities $f_i^*(\lambda, q, r, s)$, $f_i^{I^*}(\lambda, q, r, s)$ and $f_i^{S^*}(\lambda, q, r, s)$ to be the relevant distributions computed without taking into account behavioral interventions. Then,

$$f_i(\lambda, q, r, s) = f_i^*(G_i(\lambda), q, r, s)$$

$$f_i^I(\lambda, q, r, s) = f_i^{I^*}(G_i(\lambda), q, r, s)$$

$$f_i^S(\lambda, q, r, s) = f_i^{S^*}(G_i(\lambda), q, r, s)$$

For example, we can assume that each day, all individuals reduce their number of contacts by $\frac{1}{10}$. Then $G_i(\lambda) = \frac{9}{10}\lambda$, and the distributions are changed accordingly. This is a simple example of how the model can be used to reflect behavioral changes during the outbreak. The model can be extended further by using non-deterministic or state-based functions for partial compliance with government programs.

4. European data

We conclude with an example based on research by Mossong et al [**14**]. In this section, we will use the word 'day' to refer to an actual real-time day, and the word 'generation' to refer to a generation of influenza. The authors conducted a thorough study of contacts by participants in 8 European countries. Participants were asked to record their daily contacts, "defined as either skin-to-skin contact such as a kiss or handshake… or a two-way conversation with three or more words in the physical presence of another person" [**14**]. The information from the participants' diaries was weighed to match the demographics of participating countries. The group published distributions of daily contacts by individuals from each country. We ran our model on the populations of four of those countries, namely, Belgium, Great Britain, Germany, and Poland, to demonstrate empirical use of our model.

To incorporate this information into our model, we note that these data were selected by sampling distinct persons and noting the number of contacts that person had in one day. This is not exactly $f(\lambda)$ as we have defined it for two reasons. Firstly, $f(\lambda)$ is the distribution of contact rates within a generation, not one day. To correct for this, we assumed that a generation of the flu is approximately 2.5 days [**15, 16**], and extrapolated the available data accordingly. Secondly, $f(\lambda)$ is the distribution of contact rates within a population, and not the distribution of the number of daily contacts by an individual, which is the quantity being sampled. Recall that the number of daily contacts by a person is Poisson distributed with mean λ. Each sample point in the empirical data is a value picked from this Poisson distribution. However, we note that this empirical distribution is the maximum-likelihood estimator for $f(\lambda)$ in a specific country and we used it to estimate $f(\lambda)$ for our model. If the data from Mossong constitute a representative sample of the population, then the daily contact distribution should be equivalent to $f(\lambda)$. We used the same basic parameters for each country, so that the only differences between countries

affecting the epidemic curve were their respective sizes and their contact rate distributions. Finally, note that in this example we model heterogeneity based solely on contact rates, and none of the other possible heterogeneous parameters.

Basic Parameters	
$q(\lambda)$	.002 (estimated H1N1 mortality rate) [17]
P	.04
N_i	10

Table 3: Initial global parameters used in model instance.

Country-Dependent Parameters				
	Belgium	**Great Britain**	**Germany**	**Poland**
Population	10,423,493*	62,348,447*	82,282,988*	38,463,689*
Average Daily Contact Rate	11.84	11.74	7.95	16.31

Table 4: Initial country dependent parameters used in model. *Data accessed from CIA Factbook, [18]

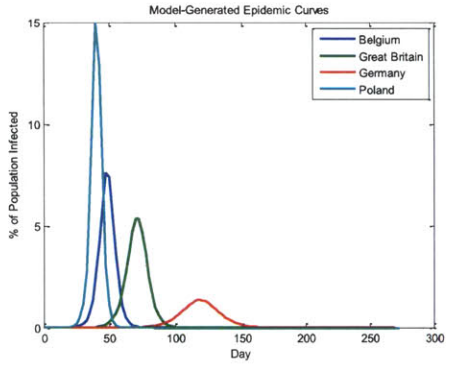


Figure 6: Epidemic curves for the 4 countries as percentage of the population infected on a given day.

Table 5 summarizes the results of our model for each of the four countries. Figures 6 and 7 further illustrate the epidemic curves and the characteristics of the population before and after the outbreaks.

Results				
	Belgium	**Great Britain**	**Germany**	**Poland**
$\mathbf{R_0}$	1.98	1.66	1.33	2.54
Total Infections	4,620,000 (44%)	25,780,000 (41%)	15,788,000 (19%)	23,787,000 (62%)
Duration	103 days	118 days	258 days	83 days

Table 5: The results of running the model using initial parameters described in Table 3 and Table 4.

It is not surprising with all transmissibility parameters being equal, that the R_0 as well as the total number of infections caused by the outbreak is correlated with the mean contact rate in the population. It is also worth noting, that as suggested in our model analysis, the variance in the contact rates in the country is correlated with change in the susceptible population characteristics that occurred throughout the outbreak, see Table 6. Recall, also, that variance plays a role in the R_0 calculations, and as a result while Belgium and Great Britain have similar average contact rates in their respective populations, the values of R_0 for the two countries differ.

Higher variance implies that a country has a combination of highly active people and relatively non-active people in its population. During the first few generations of the outbreak, the high activity super-spreaders become infected and immune, while the susceptible population becomes dominated by low-activity individuals.

Contact Rate Variance and Population Characteristic Change				
	Belgium	**Great Britain**	**Germany**	**Poland**
Contact Rate Variance	9.85	7.67	6.26	11.45
Mean Square Population Change	6.6e-005	4.8e-005	1.0e-005	6.6e-004

Table 6: The variance of a region's contact rate distribution compared to the total change in its population characteristics.

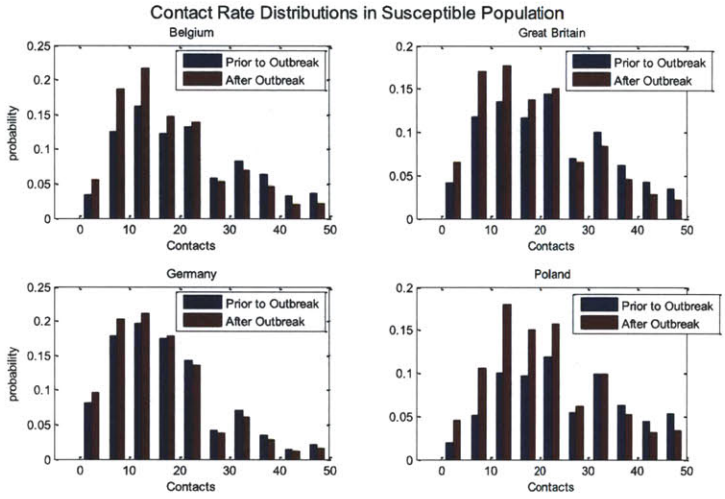


Figure 7: The contact rate distributions in the susceptible population of each country before the beginning and after the end of the outbreak.

5. Discussion

Current policy literature gives little consideration to changing characteristics of the population seen throughout the outbreak. However, an argument can be made, that heterogeneity should play a significant role in certain policy decisions. Consider for example, vaccine administration. In a population with little heterogeneity and low contact rate variance, the characteristics do not change significantly throughout the outbreak, and vaccines administered early in an outbreak reach similar classes of people as vaccines administered later on. In a high-variance population, vaccines administered early will be more likely given to high-activity people, whereas vaccines administered later on will be particularly ineffective, as most of the recipients will already be low-activity individuals. We must not forget, however, that other considerations must play a role in policy decisions. Researchers have suggested that it is often preferable to vaccinate high-risk individuals like pregnant women and people with comorbidities [**8**]. However, mathematical models such as this one may provide insights into when and why activity-focused vaccination policies may be preferable.

The model we provide in this chapter is generalizes existing heterogeneous compartmental models. The continuous spectrum of heterogeneous properties allows modelers a natural alternative to discrete compartments with identical individuals and proportional inter-compartmental mixing. Inclusion of heterogeneity in population modeling is especially important in policy decisions such as occur in designing vaccine allocation plans and message-targeting in public awareness campaigns. The framework we provide not only simplifies modeling heterogeneous populations, but also provides insight into changes of population attributes over time. The closed form equations describing evolving population characteristics should invite further mathematical analysis of epidemics.

Please note that the content in this chapter has been published in the European Journal of Operations Research [**19**].

References

1. Larson R. Simple models of influenza progression within a heterogeneous population. *Operations Research*. 2007;55(3):399-412.
2. Diekmann O, Heesterbeek J, Metz J. On the definition and the computation of the basic reproduction ration R0 in models for infectious diseases. *J. Math. Biol.* 1990;35:503-522.
3. Diekmann O, Heesterbeek J. *Mathematical Epidemiology of Infectious Diseases: Model Building Analysis and Interpretation*. New York; 2000.
4. May R. Network structure and the biology of populations. *Trends Ecol Evol*. 2006;21:394-399.
5. Gallager R. *Discrete Stochastic Processes*. Boston: Kluwer Academic Publishers; 1996.
6. Andersson H, Britton T. *Stochastic Epidemic Models and Their Statistical Analysis*. NY: Springer-Verlag; 2000.
7. Shaked M, Shanthikumar G. *Stochastic Orders*. NY: Springer Science; 2007.
8. Longini I, Halloran E. Strategy for Distribution of Influenza Vaccine to High-Risk Groups and Children. *Am J Epidemiol*. 2005;161(4):303-306.
9. Centers for Disease Control and Prevention. Seasonal Flu Shot Questions & Answers. Available at: http://www.cdc.gov/FLU/about/qa/flushot.htm. Accessed January 6, 2011.
10. Taubenberger J, Morens D. 1918 Influenza: the Mother of All Pandemics. *Emerg Infect Dis [serial on the Internet]*. January 2006;12(1).
11. Centers for Disease Control and Prevention. 2009 H1N1 Early Outbreak and Disease Characteristics. Available at: http://www.cdc.gov/h1n1flu/surveillanceqa.htm. Accessed Feburary 25, 2011.
12. Nigmatulina K. Modeling and Responding to Pandemic Influenza: Importance of Population Distributional Attributes and Non-Pharmaceutical Interventions (Doctoral dissertation).
13. Shi K, Lu J, Fan H, et al. Rationality of 17 cities' public perception of SARS and predictive model of psychological behavior. *Chinese Science Bulletin*. 2003;48(13):1297-1303.
14. Mossong J, Hens N, Jit M, Beutels P, Auranen K. Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS Med*. 2008.
15. White L, Wallinga J, Finelli L, Reed C. Estimation of the reproductive number and serial inteval in early phase of the 2009 influenza A/H1N1 pandemic in the USA. *Influenza and Other Respiratory Viruses*. 2009;3:267-276.
16. Tuite A, Greer A, Whelan M, et al. Estimated epidemiologic parameters and morbidity associated with pandemic H1N1 influenza. *CMAJ*. 2010;182(2):131-136.
17. Centers for Disease Control and Prevention. CDC Estimates of 2009 H1N1 Influenza Cases, Hospitalizations and Deaths in the United States, April 2009 – March 13, 2010. April 19, 2010. Available at: http://www.cdc.gov/h1n1flu/estimates/April_March_13.htm. Accessed January 20, 2012.
18. Central Intelligence Agency. The World Factbook. Available at: https://www.cia.gov/library/publications/the-world-factbook/index.html. Accessed February 25, 2011.
19. Teytelman A, Larson R. Modeling influenza progression within a continuous-attribute heterogeneous population. *European Journal of Operations Research*. July 2012;220(1):238-250.

20. Erdos P, Renyi A. On random graphs. *Publ Math.* 1959;6:290-297.
21. Bender E, Canfield E. The asymptotic number of labelled graphs with given degree sequences. *J. Comb. Th.* 2006:1377-1397.
22. Britton T, Deijfen M, Anders M. Generating simple random graphs with prescribed distribution. *J Stat Phys.* 2006;124(6):1377-1397.

Appendix A: Proofs

Theorem 1: Starting with any initial conditions, $f_0^I(\lambda)$, and $f_0^S(\lambda)$; for any $i > 0$,

$$f_i^I(\lambda) = C_i^I \left(e^{-\lambda p \sum_{j=0}^{i-2} \beta_j} - e^{-\lambda p \sum_{j=0}^{i-1} \beta_j}\right) f_0^S(\lambda) \tag{A.7}$$

$$f_i^S(\lambda) = C_i^S \left(e^{-\lambda p \sum_{j=0}^{i-1} \beta_j}\right) f_0^S(\lambda) \tag{A.8}$$

Where C_i^I and C_i^S are normalizing constants.

Proof:

We proceed by induction. Suppose that for some $i > 0$,

$$f_i^I(\lambda) = C_i^I \left(e^{-\lambda p \sum_{j=0}^{i-2} \beta_j} - e^{-\lambda p \sum_{j=0}^{i-1} \beta_j}\right) f_0^S(\lambda)$$

$$f_i^S(\lambda) = C_i^S \left(e^{-\lambda p \sum_{j=0}^{i-1} \beta_j}\right) f_0^S(\lambda)$$

Then,

$$f_{i+1}^I(\lambda) = \frac{p_i(I|\lambda) * f_i^S(\lambda)}{\int_0^\infty p_i(I|\lambda) * f_i^S(\lambda) d\lambda} = C_{i+1}^I \left(1 - e^{-\lambda p \beta_i}\right) \left(e^{-\lambda p \sum_{j=0}^{i-1} \beta_j}\right) f_0^S(\lambda)$$
$$= C_{i+1}^I \left(e^{-\lambda p \sum_{j=0}^{i-1} \beta_j} - e^{-\lambda p \sum_{j=0}^{i} \beta_j}\right) f_0^S(\lambda)$$

Since $\int_0^\infty p_i(I|\lambda) f_i^S(\lambda)$ is a constant. Similarly,

$$f_{i+1}^S(\lambda) = \frac{\left(1 - p_i(I|\lambda)\right) * f_i^S(\lambda)}{\int_0^\infty \left(1 - p_i(I|\lambda)\right) * f_i^S(\lambda) d\lambda} = C_{i+1}^S \left(e^{-\lambda p \beta_i}\right) \left(e^{-\lambda p \sum_{j=0}^{i-1} \beta_j}\right) f_0^S(\lambda)$$
$$= C_{i+1}^S \left(e^{-\lambda p \sum_{j=0}^{i} \beta_j}\right) f_0^S(\lambda).$$

Now for the base case; note that since there were no infected people prior to day 0, $\beta_{-1} = 0$.

$$f_1^I(\lambda) = \frac{p_0(I|\lambda) * f_0^S(\lambda)}{\int_0^\infty p_0(I|\lambda) * f_0^S(\lambda) d\lambda} = C_1^I \left(1 - e^{-\lambda p \beta_0}\right) f_0^S(\lambda) = C_1^I \left(1 - e^{-\lambda p \beta_0}\right) \left(e^{-\lambda p \beta_{-1}}\right) f_0^S(\lambda)$$
$$= C_1^I \left(e^{-\lambda p \sum_{j=0}^{-1} \beta_j} - e^{-\lambda p \sum_{j=0}^{0} \beta_j}\right) f_0^S(\lambda)$$

and

$$f_1^S(\lambda) = \frac{\left(1 - p_0(I|\lambda)\right) * f_0^S(\lambda)}{\int_0^\infty \left(1 - p_0(I|\lambda)\right) * f_0^S(\lambda)\, d\lambda} = C_1^S e^{-\lambda p \beta_0} f_0^S(\lambda) = C_1^S \left(e^{-\lambda p \beta_0}\right) = C_i^S \left(e^{-\lambda p \sum_{j=0}^{0} \beta_j}\right) f_0^S(\lambda)$$

Lemma 1: For any constants $0 < a \leq 1$, $0 < b \leq 1$, the function

$$g(\lambda) = \left(\frac{\left(1 - e^{-\lambda a}\right)}{\left(e^{\lambda b} - 1\right)}\right)$$

is monotonically decreasing and convex in λ for $\lambda > 0$.

Proof: We first prove convexity.

Let $\lambda b = \alpha$, and $\frac{a}{b} = \gamma$, and $h(\alpha) = \left(\frac{(1-e^{-\alpha\gamma})}{(e^{\alpha}-1)}\right)$. Then $g(\lambda)$ is convex in λ if and only if $h(\alpha)$ is convex in α. Differentiating:

$$\frac{d^2h}{d\alpha^2} = \frac{e^{-\alpha\gamma}\left((e^{2\alpha}+e^{\alpha})e^{\alpha\gamma} + (-e^{2\alpha}+2e^{\alpha}-1)\gamma^2 + (2e^{\alpha}-2e^{2\alpha})\gamma - e^{2\alpha} - e^{\alpha}\right)}{e^{3\alpha}-3e^{2\alpha}+3e^{\alpha}-1}$$

The denominator, $e^{3\alpha} - 3e^{2\alpha} + 3e^{\alpha} - 1 = (e^{\alpha}-1)^3 > 0$ since $\alpha > 0$.

Consider now the numerator. The numerator evaluated at $\alpha = 0$ is 0. We want to show that it is positive for all other α, and therefore the numerator is positive for all $\alpha > 0$. Let $e^{\alpha} = \mu$, and let's factor out the positive term $e^{-\alpha\gamma}$. We are left to examine

$$\kappa(\mu) = \mu^{2+\gamma} + \mu^{1+\gamma} - (\gamma+1)^2\mu^2 + (2\gamma^2 + 2\gamma - 1)\mu - \gamma^2$$

Differentiating with respect to μ,

$$\frac{d\kappa}{d\mu} = (2+\gamma)\mu^{1+\gamma} + (1+\gamma)\mu^{\gamma} - 2(\gamma+1)^2\mu + (2\gamma^2 + 2\gamma - 1)$$

Evaluating at $\mu = 1$,

$$\left.\frac{d\kappa}{d\mu}\right|_{\mu=1} = (2+\gamma) + 1 + \gamma - 2(\gamma+1)^2 + (2\gamma^2 + 2\gamma - 1) = 0$$

Differentiating again,

$$\frac{d^2\kappa}{d\mu^2} = (1+\gamma)(2+\gamma)\mu^{\gamma} + (1+\gamma)\gamma\mu^{\gamma-1} - 2(\gamma+1)^2$$

$$\left.\frac{d^2\kappa}{d\mu^2}\right|_{\mu=1} = (1+\gamma)(2+\gamma) + (1+\gamma)\gamma - 2(\gamma+1)^2 = 0$$

$$\frac{d^3\kappa}{d\mu^3} = (1+\gamma)(2+\gamma)\gamma\mu^{\gamma-1} + (1+\gamma)\gamma\mu^{\gamma-1}$$

Since all quantities in the above equation are positive, the third derivative of κ is positive for all $\mu \geq 1$. Note also, that κ and its first and second derivatives are 0 at $\mu = 1$, which implies that the first and second derivatives of κ as well as κ itself must be positive for all $\mu > 1$. This in turn implies that $\frac{d^2h}{d\alpha^2}$ is positive for all $\alpha > 0$ and the function g is convex for all $\lambda > 0$.

To prove the monotonic descent property, we note that $h(\alpha)$is monotonically decreasing if and only if $g(\lambda)$is monotonically decreasing.

$$\frac{dh}{d\alpha} = -\frac{e^{-\alpha\gamma}(e^{\alpha\gamma+\alpha} + (1-e^{\alpha})\gamma - e^{\alpha})}{(e^{\alpha}-1)^2} = -\frac{e^{-\alpha\gamma}(e^{\alpha}(e^{\alpha\gamma}-1) + (1-e^{\alpha})\gamma)}{(e^{\alpha}-1)^2} < 0$$

For all $\alpha > 0$. Consequently, $g(\lambda)$is monotonically decreasing.

Theorem 2: Stochastic ordering of susceptible and infectious populations.

Let Λ_i^S be the random variable distributed as $f_i^S(\lambda)$, and similarly Λ_i^S as $f_i^S(\lambda)$.

Then, as long as $f_0^S(\lambda), f_0^I(\lambda)$ are non-deterministic, for any $i > 0$,

$$\Lambda_{i+1}^S <_{ST} \Lambda_i^S \quad \text{(A.9)}$$
$$\Lambda_i^S <_{ST} \Lambda_{i+1}^I \quad \text{(A.10)}$$
$$\Lambda_{i+1}^I <_{ST} \Lambda_i^I \quad \text{(A.11)}$$

Where $X <_{ST} Y$ implies that X is strictly smaller than Y in the usual stochastic order.

Proof: Take any $i > 0$.

$$f_{i+1}^S(\lambda) = \frac{\left(1 - p_i(I|\lambda)\right) * f_i^S(\lambda)}{\sum_0^\infty \left(1 - p_i(I|\lambda)\right) * f_i^S(\lambda)} = f_i^S(\lambda)\frac{\left(e^{-\lambda p \beta_i}\right)}{\sum_0^\infty (e^{-\lambda p \beta_i}) * f_i^S(\lambda)} = f_i^S(\lambda)\frac{\left(e^{-\lambda p \beta_i}\right)}{E_i^S(e^{-\lambda p \beta_i})}$$

Since $\beta_i > 0$, $e^{-\lambda p \beta_i}$ is strictly monotonically decreasing in λ. Therefore,

$$\frac{f_{i+1}^S(\lambda)}{f_i^S(\lambda)} = \frac{\left(e^{-\lambda p \beta_i}\right)}{E_i^S(e^{-\lambda p \beta_i})}$$

is strictly monotonically decreasing in λ implying that Λ_{i+1}^S is smaller than Λ_i^S in likelihood ratio order, which in turn implies Equation (9). See Shaked and Shanthikumar for further detail **7**.

Similarly,

$$f_{i+1}^I(\lambda) = \frac{p_i(I|\lambda) * f_i^S(\lambda)}{\sum_0^\infty p_i(I|\lambda) * f_i^S(\lambda)} = f_i^S(\lambda)\frac{\left(1 - e^{-\lambda p \beta_i}\right)}{E_i^S(1 - e^{-\lambda p \beta_i})}.$$

Here $\left(1 - e^{-\lambda p \beta_i}\right)$ is strictly monotonically increasing, so

$$\frac{f_i^S(\lambda)}{f_{i+1}^I(\lambda)} = \frac{E_i^S(1 - e^{-\lambda p \beta_i})}{\left(1 - e^{-\lambda p \beta_i}\right)}$$

Is strictly monotonically decreasing in λ, implying Equation (10).

Finally,

$$f_i^I(\lambda) = C_i^I\left(e^{-\lambda p \sum_{j=0}^{i-2} \beta_j} - e^{-\lambda p \sum_{j=0}^{i-1} \beta_j}\right) f_0^S(\lambda) = \frac{C_i^I}{C_i^S}\left(e^{\lambda p \beta_{i-1}} - 1\right) f_i^S(\lambda)$$

$$f_i^S(\lambda) = \frac{\frac{C_i^S}{C_i^I} f_i^I(\lambda)}{(e^{\lambda p \beta_{i-1}} - 1)}$$

Furthermore,

$$f_{i+1}^{I}(\lambda) = f_i^S(\lambda)\frac{\left(1-e^{-\lambda p\beta_i}\right)}{E_i^S(1-e^{-\lambda p\beta_i})} = \frac{\frac{C_i^S}{C_i^I}f_i^I(\lambda)}{(e^{\lambda p\beta_{i-1}}-1)} * \frac{\left(1-e^{-\lambda p\beta_i}\right)}{\frac{C_i^S}{C_i^I}E_i^I(\frac{(1-e^{-\lambda p\beta_i})}{(e^{\lambda p\beta_{i-1}}-1)})}$$

$$= f_i^I(\lambda) * \frac{\left(\frac{\left(1-e^{-\lambda p\beta_i}\right)}{(e^{\lambda p\beta_{i-1}}-1)}\right)}{E_i^I\left(\frac{(1-e^{-\lambda p\beta_i})}{(e^{\lambda p\beta_{i-1}}-1)}\right)}$$

From the definition of β_i, and the fact that $i > 0$; $\beta_{i-1} > 0$ and $\beta_i > 0$.

From lemma 1, we know that $\left(\frac{(1-e^{-\lambda p\beta_i})}{(e^{\lambda p\beta_{i-1}-1})}\right)$ is monotonically decreasing $\lambda > 0$, so

$$\frac{f_{i+1}^I(\lambda)}{f_i^I(\lambda)} = \frac{\left(\frac{\left(1-e^{-\lambda p\beta_i}\right)}{(e^{\lambda p\beta_{i-1}}-1)}\right)}{E_i^I\left(\frac{(1-e^{-\lambda p\beta_i})}{(e^{\lambda p\beta_{i-1}}-1)}\right)}$$

is monotonically decreasing, and Equation (11) is proved.

Theorem 3: As long as $f_0^S(\lambda), f_0^I(\lambda)$ are non-deterministic, for any $i > 0$,

$$f_i^S(\lambda) > f_{i+1}^S(\lambda)\ for\ \lambda \geq E_i^S(\lambda) \quad (A.12)$$
$$f_i^S(\lambda) < f_{i+1}^I(\lambda)\ for\ \lambda \geq E_i^S(\lambda) \quad (A.13)$$
$$f_i^I(\lambda) > f_{i+1}^I(\lambda)\ for\ \lambda \geq E_i^I(\lambda) \quad (A.14)$$

regardless of initial contact rate distribution.

Proof : Take any $i > 0$.

By Jensen's inequality, since $e^{-\lambda p\beta_i}$ is a strictly convex function, $e^{-E_i^S(\lambda)p\beta_i} < E_i^S(e^{-\lambda p\beta_i})$

Furthermore, since $e^{-\lambda p\beta_i}$ is strictly decreasing, whenever $\lambda \geq E_i^S(\lambda)$;

$e^{-\lambda p\beta_i} \leq e^{-E_i^S(\lambda)p\beta_i} < E_i^S(e^{-\lambda p\beta_i})$, and using the derivation from theorem 2,

$f_{i+1}^S(\lambda) = f_i^S(\lambda)\frac{(e^{-\lambda p\beta_i})}{E_i^S(e^{-\lambda p\beta_i})} < f_i^S(\lambda)$, proving Equation (12).

Now for Equation (13), note that since $f_0^S(\lambda)$ is non-deterministic and $\lambda \geq 0$, $E_i^S(\lambda) > 0$. We are interested in $\lambda \geq E_i^S(\lambda) > 0$.

Since $\left(1-e^{-\lambda p\beta_i}\right)$ is strictly concave for $\lambda > 0$, we again use Jensen's inequality, to show that

$1-e^{-E_i^S(\lambda)p\beta_i} > E_i^S(1-e^{-\lambda p\beta_i})$.

Whenever $\lambda \geq E_i^S(\lambda)$; $1-e^{-\lambda p\beta_i} \geq 1-e^{-E_i^S(\lambda)p\beta_i} > E_i^S(1-e^{-\lambda p\beta_i})$ and

$$f_{i+1}^{I}(\lambda) = f_i^S(\lambda)\frac{\left(1-e^{-\lambda p\beta_i}\right)}{E_i^S(1-e^{-\lambda p\beta_i})} > f_i^S(\lambda)$$

Finally, we use Jensen's inequality for the third time for the last equation. Recall from lemma 1, that $\left(\frac{(1-e^{-\lambda p\beta_i})}{(e^{\lambda p\beta_{i-1}}-1)}\right)$ is monotonically decreasing and convex for $\lambda > 0$.

Whenever $\lambda \geq E_i^I(\lambda)$,

$$\left(\frac{\left(1-e^{-\lambda p\beta_i}\right)}{(e^{\lambda p\beta_{i-1}}-1)}\right) \leq \left(\frac{\left(1-e^{-E_i^I(\lambda)p\beta_i}\right)}{\left(e^{E_i^I(\lambda)p\beta_{i-1}}-1\right)}\right) < E_i^I\left(\frac{\left(1-e^{-\lambda p\beta_i}\right)}{(e^{\lambda p\beta_{i-1}}-1)}\right)$$

and

$$f_{i+1}^{I}(\lambda) < f_i^I(\lambda).$$

Corollary 1: For $i > 0$, and $f_0^S(\lambda), f_0^I(\lambda)$ non deterministic

$$E_i^S(\lambda) < E_{i+1}^S(\lambda) \quad \text{(A.15)}$$
$$E_i^I(\lambda) < E_{i+1}^I(\lambda) \quad \text{(A.16)}$$

Proof: The proof follows immediately from the stochastic ordering of Λ_i^S and Λ_i^I.

Appendix B: Feasibility of contact rate distributions.

The model assumes that each individual person's number of daily contacts is randomly distributed as a Poisson distribution with rate λ, independently of all other people in the community. Since all contacts must involve two parties, we must examine the feasibility of such an assumption. In order to visualize a scenario of daily contact activity, we view the community as a graph in which each individual is represented by a node and each contact, by a segment connecting two nodes. Since any pair of individuals may have multiple contacts on any given day, we allow for multiple segments connecting two nodes but we do not allow loops that represent self-contacts.

We address here, the feasibility of such a graph given that at any time. In small community of, say, 3 people, we might not be able to create a scenario that corresponds to given contact rate constraints. For example, there is no reasonable way to model a situation in which one super-spreader has 5 contacts, while the other two people have only 1 contact each on a given day. Random graph literature suggests, however, that for large enough graphs, this problem is no longer significant.

Random graphs have been a focus of research for at least a half a century. Erdos and Renyi first proposed a way to create graphs with Poisson-distributed degrees at random [**20**]. The problem of generating a random graph when each node has a known degree distribution was soon addressed by the configuration model [**21**]. Later research has shown that with large N, the configuration model can be adapted to simple graphs, i.e. undirected graphs with no self-loops or multiple edges, to create a random graph with desired degree distributions [**22**]. The same methods may be applied to create undirected graphs like ours, which allow multiple edges but not self-loops. We do not describe the methodology here, but refer the reader to these works [**20**, **21**, **22**] to confirm that in graphs with relatively large numbers of nodes and, therefore, in any reasonable susceptible population, the contact constraints defined by Poisson rates can be easily satisfied.

Appendix C: Numerical examples

Example 1: Uniform Distribution

Since λ is inherently greater than 0, one simple, reasonable selection for $f(\lambda)$ is the uniform distribution. On day 0, we assume that $f_0(\lambda)$ is a uniform distribution of Poisson contact rates between 0 and 40, and we introduce 10 infectious individuals into the population.

We start with the following initial parameters:

$$f(\lambda) = \frac{1}{40} \; for \; \lambda \in [0,40]$$
$$f_0^I(\lambda) = \frac{1}{40} \; for \; \lambda \in [0,40]$$
$$f_0^S(\lambda) = \frac{1}{40} \; for \; \lambda \in [0,40]$$
$$N_0^I = 10$$
$$N_0^S = 9{,}990$$
$$N = \; 10{,}000$$
$$p = .1$$

Day 1:

$$\beta_0 = \frac{N_0^I E_0^I(\lambda)}{NE(\lambda)} = \frac{500}{500{,}000} = \frac{1}{1{,}000}$$
$$p_0(I|\lambda) = 1 - e^{-\lambda p \beta_0} = 1 - e^{-\frac{\lambda}{10{,}000}}$$
$$N_1^I = N_0^S * \int_0^\infty p_0(I|\lambda) f_0^S(\lambda) d\lambda = 9{,}990 * \int_0^{100} \frac{\left(1 - e^{-\frac{\lambda}{10{,}000}}\right)}{100} d\lambda \approx 20$$
$$N_1^S = N_0^S - N_1^I \approx 9{,}990 - 20 = 9{,}970$$
$$f_1^I(\lambda) = \frac{p_0(I|\lambda) * f_0^S(\lambda)}{\int_0^\infty p_0(I|\lambda) * f_0^S(\lambda) d\lambda} = \frac{\left(1 - e^{-\frac{\lambda}{10{,}000}}\right)}{\int_0^{100} \left(1 - e^{-\frac{\lambda}{10{,}000}}\right) d\lambda} \approx \frac{\left(1 - e^{-\frac{\lambda}{10{,}000}}\right)}{20} \; for \; \lambda \in [0,40]$$
$$f_1^S(\lambda) = \frac{\left(1 - p_0(I|\lambda)\right) * f_0^S(\lambda)}{\int_0^\infty \left(1 - p_0(I|\lambda)\right) * f_0^S(\lambda) d\lambda} = \frac{\left(e^{-\frac{\lambda}{10{,}000}}\right)}{\int_0^{40} \left(e^{-\frac{\lambda}{10{,}000}}\right) d\lambda} \; for \; \lambda \in [0,40]$$

The latter is just a version of the exponential distribution with rate $p\beta_0$ truncated at 40. We used MATLAB to implement the remainder of the outbreak according to this model.

Example 2: Poisson Distribution

A more realistic distribution for λ is the Poisson distribution. On day 0, we assume that λ is a discrete random variable and that $f_0(\lambda)$ is a Poisson distribution with mean 20, and we introduce 10 infectious individuals into the population.

$$f(\lambda) = \frac{20^\lambda}{\lambda!} e^{-20} \; for \; \lambda = 0, 1, 2 \ldots$$

$$f_0^I(\lambda) = \frac{20^\lambda}{\lambda!} e^{-20} \; for \; \lambda = 0, 1, 2 \ldots$$

$$f_0^S(\lambda) = \frac{20^\lambda}{\lambda!} e^{-20} \; for \; \lambda = 0, 1\, 2 \ldots$$

$$N_0^I = 10$$

$$N_0^S = 9{,}990$$

$$N = 10{,}000$$

$$p = .1$$

Day 1:

$$\beta_0 = \frac{N_0^I E_0^I(\lambda)}{NE(\lambda)} = \frac{200}{200{,}000} = \frac{1}{1{,}000}$$

$$p_0(I|\lambda) = 1 - e^{-\lambda p \beta_0} = 1 - e^{-\frac{\lambda}{10{,}000}}$$

Replacing the integral in the equation for the number of susceptible and infected people in the population with a summation,

$$N_1^I = N_0^S * \sum_0^\infty p_0(I|\lambda) f_0^S(\lambda) = 9{,}990 * \sum_0^\infty \left(1 - e^{-\frac{\lambda}{10{,}000}}\right) \frac{20^\lambda}{\lambda!} e^{-20} \approx 20$$

$$N_1^S = N_0^S - N_1^I \approx 9{,}990 - 20 = 9{,}970$$

$$f_1^I(\lambda) = \frac{p_0(I|\lambda) * f_0^S(\lambda)}{\sum_0^\infty p_0(I|\lambda) * f_0^S(\lambda)} = \frac{\left(1 - e^{-\frac{\lambda}{10{,}000}}\right) * \frac{20^\lambda}{\lambda!} e^{-20}}{\sum_0^\infty \left(1 - e^{-\frac{\lambda}{10{,}000}}\right) \frac{20^\lambda}{\lambda!} e^{-20}} = \frac{\left(1 - e^{-\frac{\lambda}{10{,}000}}\right) * \frac{20^\lambda}{\lambda!} e^{-20}}{20}$$

$$for \; \lambda = 0, 1, 2 \ldots$$

$$f_1^S(\lambda) = \frac{\left(1 - p_0(I|\lambda)\right) * f_0^S(\lambda)}{\sum_0^\infty \left(1 - p_0(I|\lambda)\right) * f_0^S(\lambda)} = \frac{\left(e^{-\frac{\lambda}{10{,}000}}\right) * \frac{20^\lambda}{\lambda!} e^{-20}}{\sum_0^\infty \left(e^{-\frac{\lambda}{10{,}000}}\right) \frac{20^\lambda}{\lambda!} e^{-20}} \; for \; \lambda = 0, 1, 2 \ldots$$

Once again we used MATLAB to generate the remainder of the outbreak.

Appendix D: R_0 calculation for three-dimensional heterogeneity

In Section 3.2 we derived the value of R_0 for a population with activity-level heterogeneity. We perform a similar calculation here for a population with all three types of heterogeneity. The main consideration in Section 3.2, was the selection of a "random" person from the infectious population. Since the initial distribution in the infectious population is arbitrary and we assume that the number of infectious people in the first generation is very small, we need a better approximation of the joint distribution of λ, r and s in the susceptible population. We can do this by selecting a person following the distribution $f_1^I(\lambda, r, s)$. This distribution is independent from the initial infectious distribution (see equation (5b)) and it represents the "true" infectious distribution at the very beginning of the outbreak.

Suppose now, that we select a "patient zero" from the distribution $f_1^I(\lambda, r, s)$. We know this person's value of λ and r. However, we still need to find out how many of this person's contacts result in infection. Since we are in a fully susceptible population, the last piece of the puzzle is the expected susceptibility level of a susceptible person who comes in contact with our patient zero. Using the same logic as in section 3.2, we know that most active people are the most likely ones to come in contact with patient zero. Thus, the joint distribution of s and λ in the susceptible population given that a person has come in contact with an infectious individual is $\frac{\lambda f_0^S(\lambda,s)}{E_0^S(\lambda)}$. The expected value of s for this scenario is,

$$E_0^S(s|contact\ with\ patient\ zero) = \int_0^\infty \int_0^1 \left(\frac{s\lambda f_0^S(\lambda,s)}{E_0^S(\lambda)}\right) ds d\lambda = \frac{E_0^S(s\lambda)}{E_0^S(\lambda)}$$

Since we are in a fully susceptible population, we can assume that every person "patient zero" comes in contact with is susceptible, and all of his contacts are with distinct individuals. We can now calculate R_0:

$$R_0 = E_1^I(\lambda r) * \frac{E_0^S(s\lambda)}{E_0^S(\lambda)}$$

Recall from equation (5b) that

$$f_1^I(\lambda, q, r, s) = \frac{p_0(I|\lambda, s) * f_0^S(\lambda, q, r, s)}{\int_0^1 \int_0^\infty p(I|\lambda, s) * f_0^S(\lambda, s) d\lambda\, ds}$$

where

$$p_0(I|\lambda, s) = 1 - e^{-\lambda s \int_0^1 \tilde{r}\beta_i(\tilde{r})d\tilde{r}}$$

We can further simplify this equation by noticing that $\lambda s \int_0^1 \tilde{r}\beta_i(\tilde{r})d\tilde{r} \ll 1$ at the very beginning of the outbreak. We can approximate

$$p_0(I|\lambda, s) \approx \lambda s \int_0^1 \tilde{r}\beta_i(\tilde{r})d\tilde{r}$$

Using this, we can express R_0 in terms of the initial susceptible distribution as follows:

$$R_0 = \frac{E_0^S(\lambda r * p_0(I|\lambda, s))}{E_0^S(p_0(I|\lambda, s))} * \frac{E_0^S(s\lambda)}{E_0^S(\lambda)} \approx \frac{E_0^S(\lambda^2 rs)\int_0^1 \tilde{r}\beta_i(\tilde{r})d\tilde{r}}{E_0^S(\lambda s)\int_0^1 \tilde{r}\beta_i(\tilde{r})d\tilde{r}} * \frac{E_0^S(s\lambda)}{E_0^S(\lambda)} = \frac{E_0^S(\lambda^2 rs)}{E_0^S(\lambda)}.$$

As a final note, if r, and s are constant and the only source of heterogeneity is λ, then $p = r * s$, and

$$R_0 = p\frac{E_0^S(\lambda^2)}{E_0^S(\lambda)}$$

Just as in section 3.2.

Appendix E: Three-dimensional population changes.

Theorem 5: Starting with any initial conditions, $f_0^I(\lambda, q, r, s)$, and $f_0^S(\lambda, q, r, s)$; for any $i > 0$,

$$f_i^I(\lambda, q, r, s) = C_i^I \left(e^{-\lambda p \sum_{j=0}^{i-2} \widetilde{\beta_j}} - e^{-\lambda p \sum_{j=0}^{i-1} \widetilde{\beta_j}} \right) f_0^S(\lambda, q, r, s) \tag{E.1}$$

$$f_i^S(\lambda, q, r, s) = C_i^S \left(e^{-\lambda p \sum_{j=0}^{i-1} \widetilde{\beta_j}} \right) f_0^S(\lambda, q, r, s) \tag{E.2}$$

Where C_i^I and C_i^S are normalizing constants and $\widetilde{\beta_i} = \frac{N_i^I E_i^I(\lambda r)}{N_i E_i(\lambda)}$.

Proof:

We proceed by induction as in Theorem 1. Suppose that for some $i > 0$,

$$f_i^I(\lambda, q, r, s) = C_i^I \left(e^{-\lambda s \sum_{j=0}^{i-2} \widetilde{\beta_j}} - e^{-\lambda s \sum_{j=0}^{i-1} \widetilde{\beta_j}} \right) f_0^S(\lambda, q, r, s)$$

$$f_i^S(\lambda, q, r, s) = C_i^S \left(e^{-\lambda s \sum_{j=0}^{i-1} \widetilde{\beta_j}} \right) f_0^S(\lambda, q, r, s)$$

Then,

$$\begin{aligned} f_{i+1}^I(\lambda, q, r, s) &= \frac{p_i(I|\lambda, q, s) * f_i^S(\lambda, q, r, s)}{\int_0^1 \int_0^1 \int_0^1 \int_0^\infty p(I|\lambda, q, s) * f_i^S(\lambda, q, r, s) d\lambda \, dq dr ds} \\ &= C_{i+1}^I \left(e^{-\lambda s \sum_{j=0}^{i-1} \widetilde{\beta_j}} \right) \left(1 - e^{-\lambda s \widetilde{\beta_i}} \right) f_0^S(\lambda, q, r, s) \\ &= C_{i+1}^I \left(e^{-\lambda s \sum_{j=0}^{i-1} \widetilde{\beta_j}} - e^{-\lambda s \sum_{j=0}^{i} \widetilde{\beta_j}} \right) f_0^S(\lambda, q, r, s) \end{aligned}$$

And

$$\begin{aligned} f_{i+1}^S(\lambda, q, r, s) &= \frac{\left(1 - p_i(I|\lambda, q, s)\right) * f_i^S(\lambda, q, r, s)}{\int_0^1 \int_0^1 \int_0^1 \int_0^\infty \left(1 - p(I|\lambda, q, s)\right) * f_i^S(\lambda, q, r, s) d\lambda \, dq dr ds} \\ &= C_{i+1}^S \left(e^{-\lambda s \sum_{j=0}^{i-1} \widetilde{\beta_j}} \right) \left(e^{-\lambda s \widetilde{\beta_i}} \right) f_0^S(\lambda, q, r, s) \\ &= C_{i+1}^S \left(e^{-\lambda s \sum_{j=0}^{i} \widetilde{\beta_j}} \right) f_0^S(\lambda, q, r, s). \end{aligned}$$

Proceeding to the base case,

$$\begin{aligned} f_1^I(\lambda, q, r, s) &= \frac{p_1(I|\lambda, q, s) * f_0^S(\lambda, q, r, s)}{\int_0^1 \int_0^1 \int_0^1 \int_0^\infty p(I|\lambda, q, s) * f_i^S(\lambda, q, r, s) d\lambda \, dq dr ds} = C_1^I \left(1 - e^{-\lambda s \widetilde{\beta_0}} \right) f_0^S(\lambda, q, r, s) \\ &= C_1^I \left(e^{-\lambda s \sum_{j=0}^{-1} \widetilde{\beta_j}} - e^{-\lambda s \sum_{j=0}^{0} \widetilde{\beta_j}} \right) f_0^S(\lambda, q, r, s) \end{aligned}$$

And

$$f_1^S(\lambda, q, r, s) = \frac{\left(1 - p_1(I|\lambda, q, s)\right) * f_0^S(\lambda, q, r, s)}{\int_0^1 \int_0^1 \int_0^1 \int_0^\infty \left(1 - p(I|\lambda, q, s)\right) * f_0^S(\lambda, q, r, s) d\lambda \, dq dr ds} = C_1^S \left(e^{-\lambda s \widetilde{\beta_1}} \right) f_0^S(\lambda, q, r, s)$$

We would like to prove results similar to those in Theorem 2 in a population with all three types of heterogeneity. The populations still change with time, but this time whether a person is infected depends on his susceptibility as well as his activity level. A quick analysis of equation (2b) shows that people get infected proportionally to the value of the product of their parameters λ and s. We introduce a new variable $\psi = \lambda s$ with day i distributions $h_i^S(\psi)$ and $h_i^I(\psi)$ in the susceptible and infectious subpopulations respectively. Note that

$$h_i^S(\psi) = \int_0^1 \int_0^1 \int_0^\infty f_i^S\left(\lambda, q, r, \frac{\psi}{\lambda}\right) d\lambda\, dq\, dr$$

And

$$h_i^I(\psi) = \int_0^1 \int_0^1 \int_0^\infty f_i^I\left(\lambda, q, r, \frac{\psi}{\lambda}\right) d\lambda\, dq\, dr$$

Theorem 6: Stochastic ordering in the variable Ψ=Λ S

Let Ψ_i^S be the random variable distributed as $h_i^S(\psi)$, and similarly Ψ_i^I as $h_i^I(\psi)$.

Then, as long as $f_0^S(\lambda, q, r, s), f_0^I(\lambda, q, r, s)$ are non-deterministic, for any $i > 0$,

$$\Psi_{i+1}^S <_{ST} \Psi_i^S \quad \text{(E.3)}$$
$$\Psi_i^S <_{ST} \Psi_{i+1}^I \quad \text{(E.4)}$$
$$\Psi_{i+1}^I <_{ST} \Psi_i^I \quad \text{(E.5)}$$

Where $X <_{ST} Y$ implies that X is strictly smaller than Y in the usual stochastic order.

Proof:

We first show the relationships between $h_i^S(\psi)$, $h_{i+1}^S(\psi)$, $h_i^I(\psi)$ and $h_i^I(\psi)$. After that step the proof of this theorem is exactly analogous to the proof of Theorem 2.

$$h_{i+1}^S(\psi) = \int_0^1 \int_0^1 \int_0^\infty f_{i+1}^S\left(\lambda, q, r, \frac{\psi}{\lambda}\right) d\lambda\, dq\, dr = \int_0^1 \int_0^1 \int_0^\infty f_i^S\left(\lambda, q, r, \frac{\psi}{\lambda}\right) \frac{\left(e^{-\lambda s\widetilde{\beta_i}}\right)}{E_i^S(e^{-\lambda s\widetilde{\beta_i}})} d\lambda\, dq\, dr$$
$$= \frac{\left(e^{-\psi\widetilde{\beta_i}}\right)}{E_i^S(e^{-\psi\widetilde{\beta_i}})} \int_0^1 \int_0^1 \int_0^\infty f_i^S\left(\lambda, q, r, \frac{\psi}{\lambda}\right) d\lambda\, dq\, dr = \frac{\left(e^{-\psi\widetilde{\beta_i}}\right)}{E_i^S(e^{-\psi\widetilde{\beta_i}})} h_i^S(\psi)$$

Similarly,

$$h_{i+1}^I(\psi) = \int_0^1 \int_0^1 \int_0^\infty f_{i+1}^I\left(\lambda, q, r, \frac{\psi}{\lambda}\right) d\lambda\, dq\, dr = \int_0^1 \int_0^1 \int_0^\infty f_i^S\left(\lambda, q, r, \frac{\psi}{\lambda}\right) \frac{\left(1 - e^{-\lambda s\widetilde{\beta_i}}\right)}{E_i^S(1 - e^{-\lambda s\widetilde{\beta_i}})} d\lambda\, dq\, dr$$
$$= \frac{\left(1 - e^{-\psi\widetilde{\beta_i}}\right)}{E_i^S(1 - e^{-\psi\widetilde{\beta_i}})} \int_0^1 \int_0^1 \int_0^\infty f_i^S\left(\lambda, q, r, \frac{\psi}{\lambda}\right) d\lambda\, dq\, dr = \frac{\left(1 - e^{-\psi\widetilde{\beta_i}}\right)}{E_i^S(1 - e^{-\psi\widetilde{\beta_i}})} h_i^S(\psi)$$

Finally,

$$h_i^I(\psi) = \int_0^1 \int_0^1 \int_0^\infty f_i^I\left(\lambda, q, r, \frac{\psi}{\lambda}\right) d\lambda\, dq\, dr$$
$$= \int_0^1 \int_0^1 \int_0^\infty C_i^I \left(e^{-\psi \sum_{j=0}^{i-1} \widetilde{\beta_j}} - e^{-\psi \sum_{j=0}^{i} \widetilde{\beta_j}}\right) f_0^S\left(\lambda, q, r, \frac{\psi}{\lambda}\right) d\lambda\, dq\, dr =$$
$$= \frac{C_i^I}{C_i^S}\left(e^{\psi \widetilde{\beta_{i-1}}} - 1\right) \int_0^1 \int_0^1 \int_0^\infty f_i^S\left(\lambda, q, r, \frac{\psi}{\lambda}\right) d\lambda\, dq\, dr = \frac{C_i^I}{C_i^S}\left(e^{\psi \widetilde{\beta_{i-1}}} - 1\right) h_i^S(\psi)$$

Furthermore,

$$h_{i+1}^I(\psi) = \frac{\left(1 - e^{-\psi \widetilde{\beta_i}}\right)}{E_i^S(1 - e^{-\psi \widetilde{\beta_i}})} h_i^S(\psi) = \frac{\left(1 - e^{-\psi \widetilde{\beta_i}}\right)}{E_i^S(1 - e^{-\psi \widetilde{\beta_i}})} \frac{C_i^S}{C_i^I\left(e^{\psi \widetilde{\beta_{i-1}}} - 1\right)} h_i^I(\psi) = \frac{\left(\frac{\left(1 - e^{-\psi \widetilde{\beta_i}}\right)}{\left(e^{\psi \widetilde{\beta_{i-1}}} - 1\right)}\right)}{E_i^I\left(\frac{\left(1 - e^{-\psi \widetilde{\beta_i}}\right)}{\left(e^{\psi \widetilde{\beta_{i-1}}} - 1\right)}\right)} h_i^I(\psi)$$

The rest of the proof proceeds exactly like that of Theorem 2.

Corollary 2: For $i > 0$, and $f_0^S(\lambda, q, r, s), f_0^I(\lambda, q, r, s)$ nondeterministic

$$E_i^S(\psi) < E_{i+1}^S(\psi) \quad \text{(E.6)}$$
$$E_i^I(\psi) < E_{i+1}^I(\psi) \quad \text{(E.7)}$$

Proof: The proof follows immediately from the stochastic ordering of Ψ_i^S and Ψ_i^I.

Chapter III: Herd immunity in a heterogeneous population

1. Introduction

The concept of herd immunity is as important in epidemiology as the concept of R_0. The two parameters are two sides of the same coin, characterizing the transmissibility of a given virus and the difficulty in containing it within a community. Herd immunity is formally defined as the fraction of the population that needs to be immune to the disease so that $R(t) = 1$ [**1**]. When a population is homogeneous, with each individual being statistically identical to all others, the calculation for herd immunity is immediate. As we've mentioned in Chapter 1, the homogeneous value for herd immunity can be simply calculated as $1 - \frac{1}{R_0}$. Herd immunity in a heterogeneous setting is not as straight forward to find. In this chapter we will focus on herd immunity and its relevance to vaccination programs within a heterogeneous community.

The purpose of a vaccination program is to prevent the spread of the epidemic and ease the burden on the healthcare infrastructure. In 2009, countries around the world ordered hundreds of millions of vaccine doses to cover their populations and combat the novel H1N1 virus. In most countries, however, a big portion of these vaccines went unused. The CDC in the United States shipped over 118 million (38% of the population) vaccines to US states, with approximately 66 million being administered to patients [**2**]. This comprises approximately 20% of the US population and 56% of all vaccines ordered in the country. While a 56% rate of vaccine administration might seem little, this is higher than the percentage of vaccines used in France, where approximately 94 million vaccines were ordered for a population of just under 65 million people. Even if, as was initially believed, two doses were required for each individual to be vaccinated, this order would cover over 72% of the French population. Of those, only 5 million actual vaccines were delivered [**3**, **4**] despite a national vaccination campaign. Similarly, the Netherlands ordered enough vaccine to vaccinate 100% of its population, ending up forced to sell 19 million vaccines by January of 2010 [**5**].

A more cost effective vaccination campaign would take into account two factors against vaccinating 100% of the population. Firstly, the policy-makers must plan for the fact that in a voluntary vaccination program not all people would be willing to accept a vaccine. Individuals often weigh the benefits of the vaccine against the perceived threat of the actual outbreak and the potential dangers of the vaccine itself [**6**]. Even if the uptake of vaccines were perfect, unless the vaccination is done prior to the beginning of the outbreak, there would be no need to vaccinate all individuals as some of them would have contracted the disease and recovered on their own.

In fact this question of how many individuals need to be vaccinated is very complex. If we were to vaccinate enough people to reach herd immunity prior to the beginning of the outbreak, the outbreak would never take off. Thus, the number of vaccines necessary to prevent the outbreak would be much lower than 100%. Of course, if the vaccines arrived later and herd immunity occurred when the outbreak was well on the way, there could be millions of infectious individuals still in the population. After herd immunity is achieved, the outbreak should naturally die out on its own without the need for external interventions but this might take a long time. Even though the epidemic curve will decrease, vaccination will help eradicate the virus from the population faster than it would have happened naturally. Even so,

the proportion of people vaccinated would still be well below 100%. In planning for an outbreak, healthcare officials must determine when to stop administering vaccines after herd immunity has been reached.

In this chapter we will focus on the number of vaccines necessary to reach herd immunity. In particular, we will discuss the various targeting strategies that could decrease that number by vaccinating those individuals in the population that contribute the most to spreading the virus. The insights gained from our methodology should help with planning and preparing for an upcoming epidemic as well as for targeting strategies while the epidemic is taking place. When planning to reach herd immunity in a homogeneous population, it is reasonable to plan to vaccinate a proportion of $1 - \frac{1}{R_0}$ of the population. In a heterogeneous population – and all real populations are heterogeneous – the answer depends on who gets the vaccine. It might be useful to target certain subsections of the population over others. Indeed, the problem of vaccine allocation has been analyzed in the context of distribution of vaccine to various characteristic groups. The general public policy consensus has been to first vaccinate high-risk groups including pregnant women, health care workers, and those at risk of complications from influenza [**7, 8**]. Some authors, however, claim that in the event that large stockpiles of vaccine are available, the vaccines should be distributed to the "drivers" of infection, such as schoolchildren and other high-activity individuals [**9**]. The natural implication is that people at the "right hand tails" of the activity, susceptibility, and infectivity spectra contribute the most to spreading or getting the infection. Thus, targeting them will result in fewer doses of vaccine needed to reach herd immunity than vaccinating people at random. In reality, the objectives are nuanced, balancing the need to prevent patient deaths, stop infection spread, and ease the economic burden on the community. Such a complex objective requires the balancing of vaccination targets in the community.

In addition to needing to select the most effective group of people to vaccinate and reach herd immunity, policy-makers run into the problem of certain flu-related population characteristics changing with time. We know from Chapter 2, Section 3 that the flu-related attributes of a heterogeneous population may change significantly as an epidemic continues. A random susceptible person at the beginning of the outbreak is statistically different (e.g. more socially active and susceptible to infection) from a random susceptible person at the end of the outbreak. Due to this effect, the proportion of the population that needs to be vaccinated to reach herd immunity depends on the flu-related characteristics of the population that is still susceptible.

Herd immunity in a heterogeneous population is discussed by Hethcote and van Ark. They mention the homogeneous-fallacy of herd immunity and the fact that herd immunity is quite different for heterogeneous and homogeneous populations [**10**]. They develop a discrete time differential equation model for herd immunity that occurs naturally in areas with spatial heterogeneity, where a few distinct subpopulations experience influenza with different transmission parameters. They conclude that the number of vaccines needed to reach herd immunity is higher for a community with multiple subpopulations, that is, groups that interact much more closely with people inside the population than those outside. Hill and Longini also perform extensive analysis of a single community with multiple mixing groups, each with different activity level characteristics [**11, 12**]. They find the number of vaccines needed to stop the epidemic under the assumption that vaccines are administered prior to the start of the epidemic. We extend this analysis to a generalized population in which individuals have

different activity, susceptibility, and infectivity characteristics using the model from Chapter 2 and consider the evolution of the epidemic with time.

In this chapter we will talk about the proportion of the population that would need to be vaccinated to reach herd immunity using various targeting techniques at different points of the outbreak. Of course, perfect targeting is impossible, so the values we will develop can serve as upper and lower bounds for the actual values to reach herd immunity. Still, these bounds should give guidelines for officials in charge of planning vaccination programs in their communities.

2. Definitions

In a heterogeneous community, we define two types of herd immunity in a situation in which $R_0 > 1$. Recall from Chapter 2, Appendix D, that

$$R_0 = \frac{E_0^S(p_0(I|\lambda,s) * \lambda r)}{E_0^S(p_0(I|\lambda,s))} * \frac{E_0^S(s\lambda)}{E_0^S(\lambda)},$$

where $p_0(I|\lambda,s) = 1 - e^{-\frac{\lambda s N_0^I E_0^I(\lambda r)}{NE(\lambda)}}$.

1. *Vaccine-induced herd immunity* at time t, $HI_v(t)$: Vaccine-induced herd immunity occurs when we take an outbreak that evolved naturally until time t and then instantly vaccinate that portion of the susceptible population required to yield $R(t)$ dropping to 1. $HI_v(t)$ is that *fraction* of the population that *needs to be* immune to the virus (be it because of naturally acquired immunity or through vaccination) so that $R(t) = 1$. For this theoretical concept we assume that the vaccines become effective immediately.
2. *Naturally occurring herd immunity*, HI: Naturally occurring herd immunity occurs when the *fraction* of the population that *has become naturally immune* to the outbreak yields $R(t) = 1$, without any outside interventions. Let us call the time when this occurs t^*.

In both cases, once herd immunity has been achieved, we assume that $R(t)$ stays below 1, and the epidemic dies out. There are two immediate observations:

- $HI_v(t)$ exists and makes sense only when $t \leq t^*$. For $t > t^*$ we no longer need to use outside interventions and vaccinate anyone in the population to reach herd immunity as it has already been naturally achieved. In fact, for $t > t^*$, $R(t)$ is strictly less than 1, so *more* people are naturally immune that would be necessary for the epidemic curve to start declining. If it were possible, we could even make some immune people in the population susceptible again without raising $R(t)$ above 1

- $HI_v(t^*)$ must equal HI. At $t = t^*$, no vaccination is necessary to reach herd immunity. Due to the natural course of the outbreak itself, enough people are already immune to make $R(t) = 1$.

Given any epidemiological infectious disease model, and using our modeling assumptions HI can be simply calculated. In addition, the model can be used to find the time t^* for which $R(t) > 1$ for $t < t^*$ and $R(t^*) = 1$. While different models yield different results for t^* and HI, the answers should be close for those models that closely represent the community and virus characteristics under investigation.

Finding $HI_v(t)$ is more complicated as there are many different ways in which vaccines can be distributed to the population. Each distribution method results in a different amount of vaccine necessary to reach herd immunity. To solve this problem we use the heterogeneous model from Chapter 2. In the next two sections, we will derive some bounds on the number of vaccines necessary to reach herd immunity in a heterogeneous population with three types of heterogeneity: susceptibility level s, infectivity level r, and activity level λ.

Before moving on to the calculating herd immunity bounds, we summarize the assumptions we make in this chapter:

- Once vaccines have been administered they take effect immediately. The vaccinated person becomes completely immune to infection.
- Only susceptible individuals may receive vaccines at time t.
- There are no deaths associated with the outbreak.
- There are no mass non-pharmaceutical interventions in place. The behavior of the population does not change in response to the news of an ongoing epidemic.

With these assumptions in mind, we may calculate the possible values for herd immunity given different targeting strategies; but first let us look closer at the common notion that herd immunity equals $1 - \frac{1}{R_0}$.

3. Some intuition

3.1 Random targeting

We try to get some intuition on herd immunity at the beginning of an outbreak by examining an average infectious person in the population. At any point in time, this average infectious person has some number of potentially infectious contacts we call $C(t)$. A potentially infectious contact is a contact that would result in an infection *given that this contact occurs between an infected and a susceptible person*. The value of $C(t)$ must include all components of transferring infection from an infected person to another. That is, $C(t)$ is a function of λ, r, and s.

We can think of $C(t)$ roughly as the number of contacts made by an infectious person scaled by $r, s,$ and the probability that these contacts are with a susceptible individuals. We know that for very small t, $C(t) \approx R_0$, because nearly every person in the population is susceptible. Remember from Chapter 2 that we must be careful in calculating R_0 to account for heterogeneity in the population.

Now suppose at this time t close to 0, we vaccinate some portion of the population $HI_v(t)$ at random without giving any preference to any subgroup of the population. We do this in such a way that this fraction of the population becomes immediately immune. The number of susceptible people in the population becomes $\big(1 - HI_v(t)\big)N$ where N is the total size of the population. Since t is very small and we did not give preference to any subgroup, the distributional characteristics of the susceptible population are approximately the same as those of the entire population. Thus, the probability that a given contact by an infectious person occurs with a susceptible person is simply $\big(1 - HI_v(t)\big)$, and the average number of people now infected by a given infectious person at time t is $R(t) = \big(1 - HI_v(t)\big) * R_0$.

Setting $R(t) = 1$ and solving for $HI_v(t)$, we get that $HI_v(t) \approx 1 - \frac{1}{R_0}$ for small t.

It seems that the common definition of herd immunity holds for low t and a vaccination strategy that gives no preference to any group. What happens to this equation when t is larger? From Chapter 2, Theorem 3 and Corollary 1 we know that the average activity level, λ decreases in the susceptible and infectious populations as time goes on. The more active and more susceptible individuals will become sick and naturally recover first, leaving the less active members of the population to become infected. Thus, the value of $C(t)$ will likely decrease. With that, the calculation above will also decrease the value of herd immunity.

Similarly, changing the targeting strategy will change the make-up of the resulting susceptible population, and with it the probability that a potentially infectious contact is with a susceptible person. That in turn changes the value of $HI_v(t)$. The key question is, how can we choose a targeting strategy that is most effective in bringing down $HI_v(t)$?

3.2 Selective targeting thought experiment

Suppose that we have a small community where everyone is susceptible (for now) and in which we know for sure every person's values of λ, s, and r. We would like to vaccinate some number of these people, say V, to be most effective in lowering $R(t)$. How do we select, based on each person's triple (λ, s, r), which people should receive the vaccine?

One natural conclusion would be to notice that while s only affects a person's ability to contract the infection and r only affects a person's ability to spread the infection, a person's λ value affects both. A high-activity person is both more likely to catch a disease through his numerous contacts and spread the disease to more people once she is infectious. We would expect that high-λ individuals contribute the most to spreading the disease, and we should simply vaccinate the V people with the highest activity levels.

While this strategy is very effective, it is not necessarily the optimal one. Consider a situation in which most people in the community have very similar values of λ. However, the values of r and s are spread out. That is, some people are very infectious, while others hardly shed any virus even if infected. Similarly some people have immune system deficiencies while others claim to never get sick at all. In that case, targeting high-λ individuals will have little overall effect on the spread of disease, while taking out people that are extremely infectious or extremely susceptible to infections should lower $R(t)$ quite a bit.

So far, we mentioned two key elements to keep in mind while choosing a vaccine targeting strategy. A good targeting strategy should give preference to:

- High activity-level individuals.
- People with high values of some heterogeneity parameter that has high variance.

As we will discuss further in Section 4, there are other subtle considerations we need to keep in mind when choosing the individuals with triples (λ, s, r) to vaccinate with the highest effect. Before we get to those, however, we first derive some bounds on herd immunity from simpler strategies using the

heterogeneous influenza spread model from Chapter 2. We develop these bounds for the random vaccination strategy and the strategies for targeting the highest values of λ, r, and s respectively.

4. Derivation of Herd Immunity Bounds

For initial simplicity, we assume, again, that we do not target any type of individual in particular but distribute vaccines to the susceptible population, with each person being equally likely to receive a vaccine. We will relax this assumption later in this chapter. Suppose now that at time $t < t^*$, we want to immediately reach herd immunity. Recall from Chapter 2, that we can define the state of the outbreak in the community at a time t using four quantities, $N_t^S, N_t^I, f_t^S(\lambda, r, s), f_t^I(\lambda, r, s)$. The equations for how to proceed to time $t + 1$ and further without any vaccine interference are repeated below.

$$\beta_t(r) = \frac{N_t^I E_t^I(\lambda|r) f_t^I(r)}{NE(\lambda)} \tag{1}$$

$$p_t(I|\lambda, s) = 1 - e^{\frac{\lambda s N_t^I E_t^I(\lambda r)}{NE(\lambda)}} \tag{2}$$

$$N_{t+1}^I = N_t^S \int_0^1 \int_0^\infty p_t(I|\lambda, s) * f_t^S(\lambda, s) d\lambda \, ds \tag{3}$$

$$N_{t+1}^S = N_i^S - N_{t+1}^I \tag{4}$$

$$f_{t+1}^I(\lambda, r, s) = \frac{p_t(I|\lambda, s) * f_t^S(\lambda, r, s)}{\int_0^1 \int_0^\infty p_t(I|\lambda, s) * f_t^S(\lambda, s) d\lambda \, ds} \tag{5}$$

$$f_{t+1}^S(\lambda, r, s) = \frac{\big(1 - p_t(I|\lambda, s)\big) * f_t^S(\lambda, r, s)}{\int_0^1 \int_0^\infty \big(1 - p_t(I|\lambda, s)\big) * f_t^S(\lambda, s) d\lambda \, ds} \tag{6}$$

4.1 No targeting, upper bound

Suppose we start on day t with some known quantities $N_t^S, N_t^I, f_t^S(\lambda, r, s)$, and $f_t^I(\lambda, r, s)$. We now want to find the quantity $v_v(t)$, defined as the total proportion of the population that needs to be vaccinated at time t so that that $R(t)$ decreases to below one, assuming that we only vaccinate susceptibles. As long as $t \neq 0$, some fraction $g(t)$ of the population is already immune having been sick and those who are currently ill. Combining these two concepts, $HI_v(t) = g(t) + v_v(t)$. Since our vaccination strategy is random, the only change to the four quantities from our model that occurs after the vaccination is the change to N_t^S. Let us denote this change by calling the number of susceptible people in the population at time t after the vaccination effect

$$\widetilde{N_t^S} = \big(1 - v_v(t)\big) N_t^S. \tag{7}$$

Now we want to find $v_v(t)$ such that $R_t = \frac{N_{t+1}^I}{N_t^I} = 1$.

Let us now use equations (1)-(6) to calculate R_t in terms of $v_v(t)$.

$$\beta_t(r) = \frac{N_t^I E_t^I(\lambda|r) f_t{}^I(r)}{NE(\lambda)} \tag{8}$$

$$p_t(I|\lambda, s) = 1 - e^{-\frac{\lambda s N_t^I E_t^I(\lambda r)}{NE(\lambda)}} \tag{9}$$

$$N_{t+1}^I = \widetilde{N_t^S} \int_0^1 \int_0^\infty p_t(I|\lambda, s) * f_t^S(\lambda, s) d\lambda \, ds \tag{10}$$

Now that we have N_{t+1}^I, we can determine R_t from the quantities we already know.

$$R(t) = \frac{N_{t+1}^I}{N_t^I} = \frac{(1 - \nu_v(t)) N_t^S \int_0^1 \int_0^\infty p_t(I|\lambda, s) * f_t^S(\lambda, s) d\lambda \, ds}{N_t^I} = 1 \tag{11}$$

Solving for $\nu_v(t)$:

$$\nu_v(t) = \frac{N_t^I}{N_t^S \int_0^1 \int_0^\infty p_t(I|\lambda, s) * f_t^S(\lambda, s) d\lambda \, ds} \tag{12}$$

This value can be calculated at any point t to find the fraction of the population that needs to be vaccinated in order for herd immunity to be reached. We did not target any particular individuals, so it is safe to say that this value for $HI_v(t)$ is an upper bound for the fraction of people that need to be immune to bring down the value of $R(t)$ to 1. Similarly, the value of $v_v(t) = HI_v(t) - g(t)$, is the upper bound on the proportion of the population that needs to be vaccinated at time t to reach herd immunity. We can also likely obtain lower values of $\nu_v(t)$ and $HI_v(t)$ if we target those individuals who are more likely to cause the virus to spread.

4.2 Susceptibility targeting

If we want to have the highest possible effect through the vaccination program, a logical targeting strategy is to target those individuals that contribute the most to spreading the infection in the population. We will focus only on one type of heterogeneity at a time. We start with the susceptibility level heterogeneity, s, for now. Suppose that we will only deliver vaccines to people with the highest values of s in the population. The effect of the vaccination is two-fold. Firstly, the vaccinations will decrease the susceptible pool making it less likely that an infectious individual contracts and spreads infection to a susceptible one. Secondly, the susceptible population as a whole is now less likely to contract the infection during a contact with an infectious individual.

If the highest s −valued individuals are vaccinated, $N_t^S$and $f_t^S(\lambda, r, s)$ change to reflect the vaccination, while N_t^I and $f_t^I(\lambda, r, s)$ remain unchanged. Let us call $f_t^S(\widetilde{\lambda, r, s})$ the distribution after the vaccines have taken effect and let s^* be the value of s, such that $\int_{s^*}^\infty \widetilde{f_t^S(s)}$ dr $= \nu_v(t)$. That is, s^* is the cutoff value for vaccination targeting, where we vaccinate everyone whose value of s is greater than the cutoff:

$$\widetilde{f_t^S(\lambda,r,s)} = \begin{cases} \dfrac{f_t^S(\lambda,r,s)}{1-\nu_v(t)} & for\ s \le s^* \\ 0\ otherwise \end{cases} \tag{13}$$

We make s^* our decision variable keeping in mind that we can calculate $\nu_v(t)$ given s^*.

$$\widetilde{N_t^S} = \big(1-\nu_v(t)\big)N_t^S \tag{14}$$

as before.

We now modify the steps from the previous section to find $R(t)$ as a function of s^* keeping in mind the above changes.

$$\beta_t(r) = \frac{N_t^I E_t^I(\lambda|r) f_t^{\ I}(r)}{NE(\lambda)} \tag{15}$$

$$p_t(I|\lambda,s) = 1 - e^{\frac{\lambda s N_t^I E_t^I(\lambda r)}{NE(\lambda)}} \tag{16}$$

$$N_{t+1}^I = \widetilde{N_t^S}\int_0^1\int_0^\infty p_t(I|\lambda,s) * \widetilde{f_t^S(\lambda,s)}d\lambda\, ds \tag{17}$$

$$R(t) = \frac{N_{t+1}^I}{N_t^I} = \frac{\big(1-\nu_v(t)\big)N_t^S\int_0^1\int_0^\infty p_t(I|\lambda,s) * \widetilde{f_t^S(\lambda,s)}d\lambda\, ds}{N_t^I} = 1 \tag{18}$$

Note that this time both $\nu_v(t)$ and $\widetilde{f_t^S(\lambda,s)}$ depend on s^*, so we can no longer find a closed form solution for herd immunity. However, we can use any available optimization program to find the smallest value of s^* such that $R(t) = 1$. Once s^* is found, we can calculate the corresponding values for $\nu_v(t)$ and $HI_v(t)$.

The methodology for finding optimal infectivity targeting is almost identical to this case. Since both of these parameters correspond to the probability of spreading infection, we would expect the values of $\nu_v(t)$ and $HI_v(t)$ to very similar when we target susceptibility and infectivity individually as long as the initial distributions for the two sources of heterogeneity are the same.

4.3 Activity level targeting

Recall that unlike susceptibility levels and infectivity levels, activity levels determine both a person's relative infectiousness and his likelihood to get infected. As such, the steps we take to find herd immunity are slightly different. We start by using the same steps as in Section 4.2. Given that we vaccinate $\nu_v(t)$ people at time t we change $\widetilde{f_t^S(\lambda,r,s)}$ to be

$$\widetilde{f_t^S(\lambda,r,s)} = \begin{cases} \dfrac{f_t^S(\lambda,r,s)}{1-\nu_v(t)} & for\ \lambda \le \lambda^* \\ 0\ otherwise \end{cases} \tag{19}$$

and

$$\widetilde{N_t^S} = (1 - v_v(t))N_t^S \tag{20}$$

We continue, to find

$$p_t(I|\lambda, s) = 1 - e^{-\frac{\lambda s N_t^I E_t^I(\lambda r)}{NE(\lambda)}}$$

Note that in this case, the calculation for $p_t(I|\lambda, s)$ depends on the distribution of λr in the infectious population at time t. Since we only vaccinated members of the susceptible population, this distribution is unchanged, but it will change drastically in the next generation due to the vaccination of highly active individuals. The fact that that we are only taking into account the effect of activity levels on the ability to contract the infection but not the effect on the ability to spread the infection means that we will underestimate the total effect of taking out the highly active people from the population. As a result we will likely overestimate the total number of vaccines needed to reach herd immunity.

In order to correct for this error, we need to substitute the probability density function corresponding to the infectious population on day t with another function that takes into account the effect of the vaccination on a person's ability to spread the infection. We do this by letting $\widetilde{f_t^I(\lambda, r, s)} = f_{t+1}^I(\lambda, r, s)$,

where $f_{t+1}^I(\lambda, r, s)$ is the distribution function for the infectious population on day $t + 1$ that we obtain by using equations (18) and (19), followed by equation (5) to get

$$f_{t+1}^I(\lambda, r, s) = \frac{p_t(I|\lambda, s) * \widetilde{f_t^S(\lambda, r, s)}}{\int_0^1 \int_0^\infty p_t(I|\lambda, s) * \widetilde{f_t^S(\lambda, s)} d\lambda \, ds} \tag{21}$$

In other words, we calculate the distribution in the infectious population during the next generation, and then we use this new distribution function as an approximation of the infectious distribution function at time t.

We can now make one more iteration to complete our calculations:

$$\beta_t(r) = \frac{N_t^I \widetilde{E_t^I(\lambda|r)} f_t^I(r)}{NE(\lambda)} \tag{22}$$

$$p_t(I|\lambda, s) = 1 - e^{-\frac{\lambda s N_t^I \widetilde{E_t^I(\lambda r)}}{NE(\lambda)}} \tag{23}$$

$$N_{t+1}^I = \widetilde{N_t^S} \int_0^1 \int_0^\infty p_t(I|\lambda, s) * \widetilde{f_t^S(\lambda, s)} d\lambda \, ds \tag{24}$$

$$R(t) = \frac{N_{t+1}^I}{N_t^I} = \frac{(1 - v_v(t))N_t^S \int_0^1 \int_0^\infty p_t(I|\lambda, s) * \widetilde{f_t^S(\lambda, s)} d\lambda \, ds}{N_t^I} = 1 \tag{25}$$

This time the quantities affected by λ^*, are $v_v(t)$, $\widetilde{f_t^S(\lambda, s)}$, and $p_t(I|\lambda, s)$. Once again, there is no closed form solution for this problem but we can easily solve for λ^* using an optimization algorithm.

4.4 Three-dimensional targeting

While understanding the relationship between any one heterogeneity parameter and $v_v(t)$ is useful for approximating the possible numbers of vaccines needed to reach herd immunity with targeted interventions, the true lower bound on the herd immunity would involve optimizing with respect to all three dimensions at the same time. In effect, we want to find an indicator function $h(\lambda, r, s)$ that equals 0 when a person with parameters (λ, r, s) should be vaccinated and 1 otherwise. We want to find the optimal $h(\lambda, r, s)$ such that $R(t) = 1$ after the vaccination and

$$\int_0^1 \int_0^1 \int_0^\infty \big(1 - h(\lambda, r, s)\big)\, d\lambda\, dr\, ds = v_v(t)$$

is minimized.

We can be tempted to say that the optimal $h(\lambda, r, s)$ has the following structure:

$$h(\lambda, r, s) = \begin{cases} s \le s^*, r \le r^*, \lambda \le \lambda^* \\ 0 \; otherwise \end{cases}.$$

That is, we vaccinate those people whose λ, r, or s value is higher than some predetermined cutoff values, which we can use as decision variables to find herd immunity. However, this is too simplistic. Let us inspect, closely, the key element we have been using in calculating $R(t)$ -- the probability that a given susceptible person with parameters λ and s becomes infected on day $t + 1$:

$$p_t(I|\lambda, s) = 1 - e^{-\frac{\lambda s N_t^I E_t^I(\lambda r)}{NE(\lambda)}}.$$

This function depends on the overall ability to contract an infection by a susceptible person - (λs) - as well as the overall ability to spread the infection by an infectious person - $E_t^I(\lambda r)$. This suggests that an optimal targeting strategy would be more intricate than simply removing the highest values of the three sources of heterogeneity.

While a true optimal strategy is very difficult to find due to the nonlinear nature of influenza models, we propose, as a heuristic, the strategy of targeting those individuals with maximum values of the product $\lambda^2 rs$. That is, we reduce the problem to one with a single variable w^*, and we let

$$h = \begin{cases} 1 \; for \; \lambda^2 rs \ge w^* \\ 0 \; otherwise \end{cases}. \tag{26}$$

This strategy minimizes the value of $p_t(I|\lambda, s)$ and should be very close to the optimal strategy. We modify equations (1)-(4) and follow similar steps to the ones in previous sections:

$$\widetilde{N_t^S} = \big(1 - v_v(t)\big) N_t^S \tag{27}$$

$$\widetilde{f_t^S(\lambda, r, s)} = \begin{cases} \dfrac{f_t^S(\lambda, r, s)}{1 - v_v(t)} & for \; \lambda^2 rs \le w^* \\ 0 \; otherwise \end{cases} \tag{28}$$

$$f_t^I(\widetilde{\lambda, r, s}) = f_{t+1}^I \tag{29}$$

Once again,

$$R(t) = \frac{N_{t+1}^I}{N_t^I} = \frac{\left(1 - \nu_v(t)\right) N_t^S \int_0^1 \int_0^\infty p_t(I|\lambda, s) * f_t^{\widetilde{S}}(\lambda, s) d\lambda \, ds}{N_t^I} = 1 \tag{30}$$

but this time $\nu_v(t)$, $p_t(I|\lambda, s)$, and $f_t^{\widetilde{S}}(\lambda, s)$ are all dependent on the parameter w^*, so we solve for the value of w^* that minimizes $\nu_v(t)$ while keeping $R(t) \leq 1$ using an optimization algorithm. Using our heuristic, this method gives us the proportion of the population that needs to be vaccinated at time t to reach herd immunity.

5. Illustration

5.1 Independent sources of heterogeneity

We use an illustrative example for a heterogeneous community with 100,000 people and an influenza virus with $R_0 = 1.5$, a typical influenza pandemic parameter. The community's 3 sources of heterogeneity are independently distributed with the following truncated Gaussian distribution functions:

- Activity level: $f_0^S(\lambda) = C_1 e^{-\frac{(\lambda - 10)^2}{2*3^2}}$ for $0 \leq \lambda$
- Infectivity: $f_0^S(r) = C_2 e^{-\frac{(r - .37)^2}{2*.1^2}}$ for $0 \leq r \leq 1$
- Susceptibility: $f_0^S(s) = C_3 e^{-\frac{(s - .37)^2}{2*.1^2}}$ for $0 \leq s \leq 1$

where C_1, C_2, and C_3 are normalizing constants.

Figure 1 contains graphs for $\nu_v(t)$, the proportion of the population that needs to be vaccinated at time t to reach herd immunity and herd immunity itself, $HI_v(t) = \nu_v(t) + g(t)$, for five different vaccine targeting strategies.

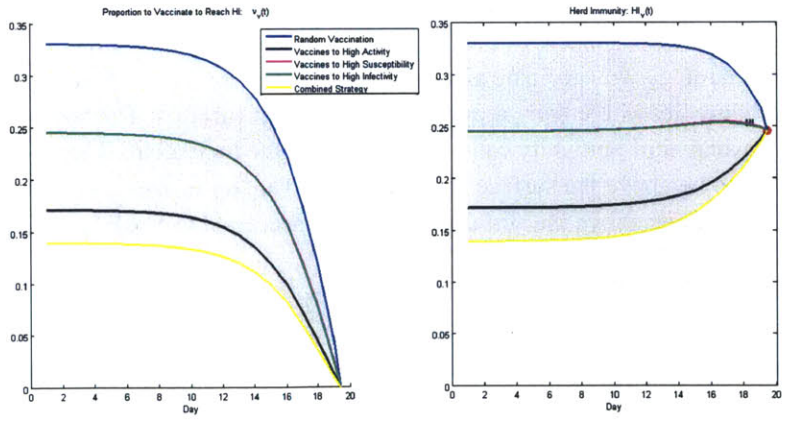


Figure 1: Values of herd immunity using 5 different targeting strategies. The shaded area represents the region possible values for $\nu_v(t)$ and $HI_v(t)$ respectively using a reasonable targeting strategy.

As expected from Section 4, the value of $HI_v(t)$ on day 1 for the random vaccination targeting strategy is approximately $1 - \frac{1}{R_0} = .33$. The high-susceptibility and high-infectivity targeting curves are identical, and vaccinating highly active individuals seems to be the best single-heterogeneity strategy. Using the combined strategy from Section 4.4, however, herd immunity can be achieved by vaccinating only 14% of the population prior to the beginning of the outbreak.

Moving on to Figure 1b, at first glance it might seem counter-intuitive that the high-activity vaccine-induced herd immunity actually increases with time. Further thought, however, implies that this is correct. The reason that the blue curve (random vaccination) decreases with time, is that the population of infective people becomes progressively less and less active, and so we need fewer people immune to stop the epidemic. This is also the case with the black and yellow curves. However, whereas in the beginning it was possible to vaccinate a significant number of high-activity people, toward the end the only people left to vaccinate are low-activity. Thus, even though there exists some proportion of people of $HI_v(1)$ who would have been enough to stop the outbreak had we chosen our vaccinated group at the very beginning, at the end of the outbreak, the immune group must include some low-activity people who have already become immune through natural means. This causes the group needed for herd immunity to become larger. This loss of "vaccine-power" is so strong that it even outweighs the benefits of a less powerful infectious population. For populations where λ, r and s have a lower variance, the change in the susceptible population would become less drastic, and the black and yellow curves would actually decrease.

Does a decreasing value of $HI_v(t)$ with time imply that we should wait to vaccinate people until late in the outbreak or not at all? This is not necessarily the case because even though the value of herd immunity changes people still get infected during that time. Waiting simply allows high-activity people to get infected, and then, since only low activity people are left, less effort is necessary to control the infection. Instead of waiting, we would like to find ways to make interventions more effective through targeted vaccination. Most importantly, the blue curve must always be above the other targeting strategies we have employed, and all curves meet when naturally-occurring herd immunity is reached. In fact, the shaded area in Figure 1 corresponds to the range of all reasonable values of herd immunity.

We also refer the reader to Figure 2, which describes the decisions that comprise the combined targeting strategy. Figure 2a is the summary of the decision variable w^* as a function of time. In other words, at a given time t under this strategy we vaccinate all people for whom $\lambda^2 rs \geq w^*$. The resulting cutoff values can be visualized in Figure 2b, where there are two example cutoff surfaces. The bottom surface corresponds to day 1, when herd immunity can be reached if all the individuals whose heterogeneity parameters cause them to lie above the surface are vaccinated. The top surface corresponds to day 19, where, similarly all individuals above the surface have to be vaccinated to reach herd immunity.

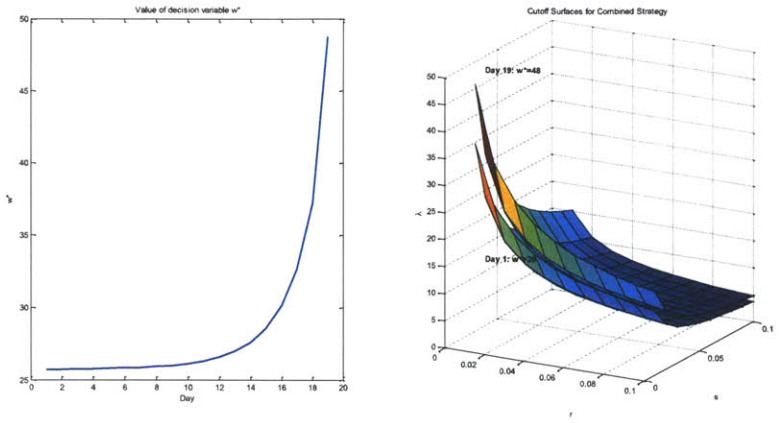


Figure 2: The value of w^* for the first example and the surfaces in (λ, s, r) space corresponding to the w^*values of Day 1 and Day 19.

5.2 Dependent sources of heterogeneity

We conclude this section with another example. So far we have assumed that initially, the three types of heterogeneity are mutually independent. However, this is likely an unrealistic assumption. Susceptibility and infectiveness have many common causes – people with poorly functioning immune systems or other preexisting conditions are likely to be both highly infectious and highly susceptible. School children are also a cause of correlation between activity levels and susceptibility. Little research has been done in the quantification of sources of heterogeneity aside from activity levels, so it is difficult to project more appropriate distributions. For the sake of completeness, we perform the same analysis as above with a population where all three sources of heterogeneity are positively correlated. We would expect that in this situation the combined targeting strategy would be particularly effective. It should be even more effective than in the first example since it takes into account the joint distribution of λ, s and r.

In the new example the community's three sources of heterogeneity initially have a truncated multinomial Gaussian distribution function with the vector $[\lambda, r, s]$ having mean $M = [10\ .35\ .35]$ and covariance matrix

$$\Sigma = \begin{bmatrix} 9 & .09 & .09 \\ .09 & .01 & .003 \\ .09 & .003 & .01 \end{bmatrix}.$$

Where $0 \leq \lambda, 0 \leq r \leq 1$, and $0 \leq s \leq 1$.

The R_0 for this initial distribution is still 1.5. The resulting values for $HI_v(t)$ and $v_v(t)$ are given below.

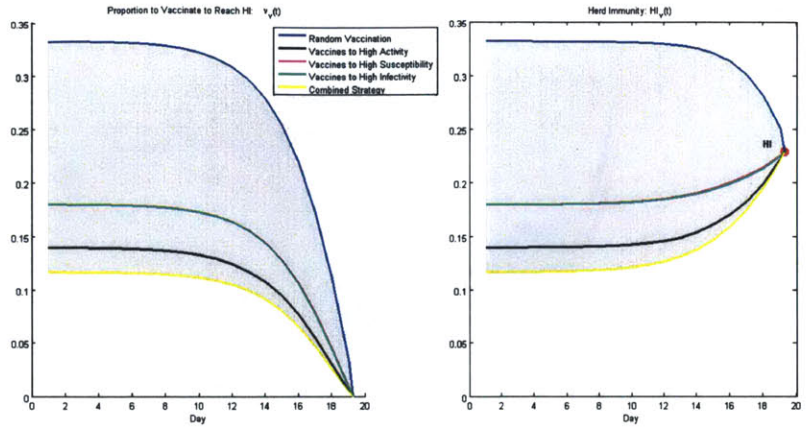


Figure 3: Values of herd immunity using 5 different targeting strategies in example 2. The shaded area represents the region possible values for $v_v(t)$ and $HI_v(t)$ respectively using a reasonable targeting strategy.

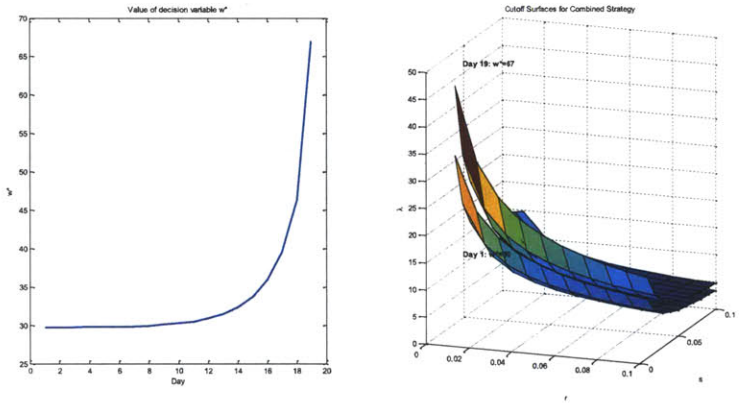


Figure 4: The value of w^* for the second example and the surfaces in (λ, s, r) space corresponding to the w^*values of Day 1 and Day 19.

As expected, all targeted curves are lower than they were in the first example. The combined strategy now requires only 12% of the population to be vaccinated to reach herd immunity prior to the beginning of the outbreak. The other noticeable change due to the newly introduced correlation is the fact that w^* is now higher, confirming that fewer people need to be vaccinated to achieve the same results of reaching herd immunity.

6. Discussion and caveats

Looking back at the range of $HI_v(t)$ values in the first example, we can conclude that in a community with the characteristics defined in Section 5 and an R_0 of 1.5, as few as 14% of the people in the population need to be vaccinated prior to the outbreak to reach herd immunity. With time, the value of herd immunity using the combined targeting strategy increases, but it is always below the traditional definition of $1 - \frac{1}{R_0} = 33.3\%$. Of course, perfect targeting is unrealistic for many reasons. Firstly, it is extremely difficult for a person to know exactly their three heterogeneity parameter values. Even if the

parameters could be estimated, it would be impossible to actually implement a campaign that would try to estimate each person's value of $\lambda^2 rs$. Finally even if these values were known perfectly, compliance is never guaranteed due to the perceived dangers of influenza vaccines [**6**].

Daily contact rates are the easiest source of heterogeneity that can be objectively measured by an individual. This makes activity level targeting the most feasible strategy in real life. In the first example, this would imply herd immunity of 17%, a difference worth just 3% of the population from the most successful strategy. Even when the targeting strategy cannot be ascertained prior to the outbreak, this methodology determines the spectrum in which the value of herd immunity falls. Knowing that the number of vaccines necessary at the beginning of the outbreak is enough to cover between 14 and 33 percent of the population healthcare officials in the community can plan for the upcoming demand in the future.

As a final note, we would like to revisit the caveats and assumptions that we made in this analysis.

- **100% effective vaccines working immediately:** This is an unrealistic assumption that is necessary for the results in Section 4. All of the results in this chapter show theoretical bounds occurring due to immediate immunity at that moment. In reality, it would take some time for the vaccines to come into effect. By the time the vaccinated individuals become immune, the value of herd immunity will have changed. A good approximation for the realistic case in which vaccines take approximately 10 days to become effective is to use the bounds from a few (about 3) generations after t to estimate the amount of vaccine necessary to reach herd immunity. Modifying the vaccine effectiveness so that vaccines are not 100% effective in the model is easily incorporated into the above methodology by appropriately scaling $\nu_v(t)$ in equation (7).
- **Vaccine distributed only to susceptibles:** A more realistic assumption would be that only *asymptomatic* individuals receive vaccines. While checks are performed during vaccine administration for recent flu like symptoms, a person may already be infected and spreading the virus prior to receiving the vaccine. The number of such people should be fairly low, however, and most likely does not interfere with the insights gained by calculating the bounds in Section 4.
- **No mortality:** Accounting for deaths adds unnecessary complexity and is unlikely to change the bounds from Section 4 by a significant amount if the objective is to limit the number of infections. A more complex objective that includes minimizing morbidity from an outbreak will need to take these into account, however.
- **No NPIs or behavior changes:** It is likely that the public will be able to respond to the news of an outbreak by changing their behavior to minimize the spread of infection. Behavioral changes like increased hand-washing and telecommuting to work can have a significant impact on the spared of infections, and they would increasingly complicate our computations. In fact, we devote Chapter 6 to analyzing the relationship between different types of non-pharmaceutical interventions and vaccine programs.

The methodology in this chapter can be extended to include different objectives or even threshold bounds other than $R(t) = 1$. For example, a goal of vaccinating enough people to have $R(t) = .5$ will ensure that the outbreak will die out very quickly but would require more vaccines than if we only needed to reach herd immunity. The key is to note that in the case of an outbreak with a value of R_0 similar to those we

have seen in the last century, it is very unlikely that more than 50% of the population would need to be vaccinated to manage the spread of infection.

We have emphasized the fact that the values for herd immunity we derived here are just bounds for what can actually happen in a real pandemic, and they should not be used as the exact numbers of vaccines necessary for a community. The main insights to be drawn here is that the number of vaccines needed to control an outbreak is dependent on the time when these vaccines can be used and the people who will wind up receiving those vaccines. If vaccines expected to arrive late, fewer vaccines will be needed as, unfortunately, a large proportion of the population will become immune through contracting influenza. Also, targeting those individuals who are most susceptible, infectious, active, or better yet have the highest combination of those factors will be more effective in stopping the disease than a random vaccination program.

Our methodology should be used as a guide to planning the total number of vaccines to order for a community and the targeting strategy for those vaccines. In the short period between the time when a novel virus has been identified but has not yet entered a community, the healthcare officials in that community can use this methodology to estimate the values of herd immunity (or another threshold) and order appropriate amounts of vaccine. Such estimates can avoid unnecessary orders that cause economic strain on the community and often result in wasted doses at the end of the outbreak. As an added benefit to society, communities that do not order excess vaccine avoid denying other communities the resources in case of a deficit.

References

1. Anderson R. The concept of herd immunity and the design of community-based immunization programmes. *Vaccine*. 1992;10:928-35.
2. Centers for Disease Control and Prevention. 2009 H1N1 Vaccine Doses Allocated, Ordered, and Shipped by Project Area. February 2, 2010. Available at: http://www.cdc.gov/h1n1flu/vaccination/vaccinesupply.htm. Accessed January 20, 2011.
3. Bone A, Guthmann J, Nicolau J, Levy-Bruhl D. Population and risk group uptake of H1N1 influenza vaccine in mainland France 2009-2010: Results of a national vaccination campaign. *Vaccine*. November 2010;28(51):8157-8161.
4. Balmer C. France wants to sell millions of surplus flu shots. *Reuters*. January 2010.
5. Dutchnews.nl. Netherlands tries to sell excess flu vaccine. *Dutchnews.nl*. January 2010:2010.
6. Harris K, Maurer J, Kellermanmn A. Influenza Vaccine - Safe, Effective, and Mistrusted. *N Engl J Med*. 2010;363:2183-2185.
7. Centers for Disease Control and Prevention. Seasonal Flu Shot Questions & Answers. Available at: http://www.cdc.gov/FLU/about/qa/flushot.htm. Accessed January 6, 2011.
8. Longini I, Halloran E. Strategy for Distribution of Influenza Vaccine to High-Risk Groups and Children. *Am J Epidemiol*. 2005;161(4):303-306.
9. Patel R, Longini I, Halloran M. Finding optimal vaccination strategies for pandemic influenza using genetic algorithms. *Journal of Theoretical Biology*. May 2005;234(2):201-212.
10. Hethcote H, van Ark J. Epidemiological models for heterogeneous populations; proportionate mixing, parameter estimation, and immunization programs. *Mathematical Biosciences*. February 1987;84(1):4.
11. Hill A, Longini I. The critical vaccination fraction for heterogeneous epidemic models. *Mathematical Biosciences*. January 2003;181(1):85-106.
12. Britton T. Epidemics in heterogeneous communities: estimation of R0 and secure vaccination coverage. *Journal of Royal Statistics Society*. April 2002;63(4):705-715.

Chapter IV: Modeling the effects of H1N1 vaccine distribution in the US

1. Introduction

In the last three chapters of this work we will switch from modeling techniques to more applied, realistic decisions that policy makers face during an influenza pandemic. To get some motivation, in this chapter we will focus on the efficacy of vaccines during the 2009-2010 H1N1 "swine flu" pandemic as they were administered in the United States. We mentioned some background of this epidemic in Chapter 1. The virus surfaced in the US and Mexico in early April of 2009, becoming a worldwide pandemic by June. The US Center for Disease Control (CDC) ordered the development of vaccines in April, and by October of 2009 vaccines started shipping to the 50 constituent states [**1**]. The vaccine had varying ameliorative effects in different regions of the US. Due in part to early school openings in the Southeast [**2**], the major wave of the outbreak hit the southeastern states earlier than the northern ones. Nonetheless, vaccines as produced were delivered to states strictly on a per-capita basis [**3**], with each state receiving vaccine in direct proportion to its population. As a result, some states were able to vaccinate a significant portion of their population prior to the major flu wave hitting, while others did not receive vaccines until the major flu wave had already passed and interest in vaccinations had waned. Vaccine distribution within individual states was significantly more complicated as states had relative freedom in utilizing the vaccines shipped from the CDC [**4**].

In this chapter, we present analysis of specific states' epidemic curves and the relative effectiveness of vaccine programs. The essential analytical issue we face is estimating the epidemic flu curve *in the absence of vaccine* while the information we have is limited and available only in the presence of vaccine. For most states readily available information contains only the percentage of all hospitalizations and outpatient visits that were caused by influenza-like illness (ILI) [**5**]. The results of our analysis in this chapter suggest the importance of administering vaccination as early as possible, even if the numbers of vaccines administered are small. Early administration not only serves to decrease the peak of the epidemic curve, but also encourages higher participation rates from the population.

As we have discussed in previous chapters, there is extensive literature that discusses targeting different demographic groups with vaccine programs [**6, 7**]. Chowell *et al.* have approached the problem of measuring vaccine effectiveness by modeling a possible outbreak similar to H1N1 based on Mexican demographics. They concluded that as much as a 37% reduction in the number of hospitalizations could have been achieved with an "adaptive" vaccination strategy based on demographics and information obtained during the outbreak [**8**]. The authors suggested that as the outbreak progresses, vaccines should be allocated to different age groups in proportion to the then-reported influenza-like symptoms and hospitalizations from within those age groups.

While these papers analyze effectiveness of vaccination programs on such parameters as the transmissibility of the flu and the demographics of the population, we choose to consider another issue, *the timing of the vaccines administered.* Timing of vaccine distribution has been explored by some authors in the context of both general influenza models [**9, 10**] and specific outbreaks [**11, 12**]. We, in turn, analyze the effect of vaccine distribution timing on the spread of the 2009-2010 H1N1 virus in

different US states. The size of the country allows us to be able to compare the spread of the same strain of influenza as it infects different communities at different times.

Specifically, for each of several US states, we examine two timing relationships, the time of onset and growth of the flu wave and the timing of vaccinations. For each state, we apply simple temporal models to estimate the number of flu infections averted by administration of vaccine.

The rapid and massive spread of H1N1 in Mexico City in the spring of 2009 motivated numerous papers estimating parameters of the epidemic and characterizing the effects of H1N1. The most important parameter in epidemiology, the reproductive number, R_0, is defined to be the average number of new infections caused by a "typical" infectious individual in a fully susceptible population. Early estimates in Mexico estimated R_0 between 1.4 and 1.6 [**13**], implying a potentially dangerous progression for other countries. This in turn led to predicted estimates of $R_0$of 1.3-1.8 for the U.S., and an estimated 60% national infection rate [**14, 15**]. The CDC later re-estimated the range of actual USA infections to be between 43 million and 88 million [**16**], representing only 14% to 29% of the population.

A study in Ontario, Canada estimated R_0 there to be about 1.31 [**17**]. The authors attribute the lower-than-predicted estimates to heightened public awareness and better adherence to hygienic behavior that reduces virus transmission. In our modeling of US outbreak progression, we also encounter relatively low values for R_0 but note that these values agree with findings in Ontario and the relatively low reported infection rates in the US.

In a recent paper, Sander *et al*, used an agent-based model to depict the spread of H1N1 in Ontario. The authors asserted that the immunization program in the province was a "cost-effective" means to prevent H1N1 cases [**12**]. Ontario's program started administering vaccines 2 weeks before the peak of infections in Ontario. Since most of the U.S. states were hit by the outbreak earlier, half of the states started administering vaccine after the peak has already passed [**18**]. Not only were vaccines delivered late, but the demands for vaccines declined significantly by the time vaccines became available [**19**]. Eventually, in the US, as many as 70 million doses were left unadminstered [**20**]. Surveys showed that low participation rates were partially a result of concerns for the safety of vaccines and the relative innocuousness of the H1N1 virus [**19, 21**]. Moreover, by December of 2009, when vaccines became widely available to everyone, the proportion of people in the population who were concerned about the dangers of H1N1 had dropped significantly from between 51-59% to about 40% [**21**].

For each state, Figure 1 displays in color codes the timing of shipment of first vaccine with respect to the peak of the infection in that state [**3**]. Six southeastern states received their first vaccines more than four weeks *after* the peak of their respective H1N1 outbreaks. The infection then spread northwest, with the northeastern states having the most time to prepare and vaccinate their populations. Maine and isolated Hawaii had more than four weeks of vaccine dispensation prior to the peak outbreak in their states.

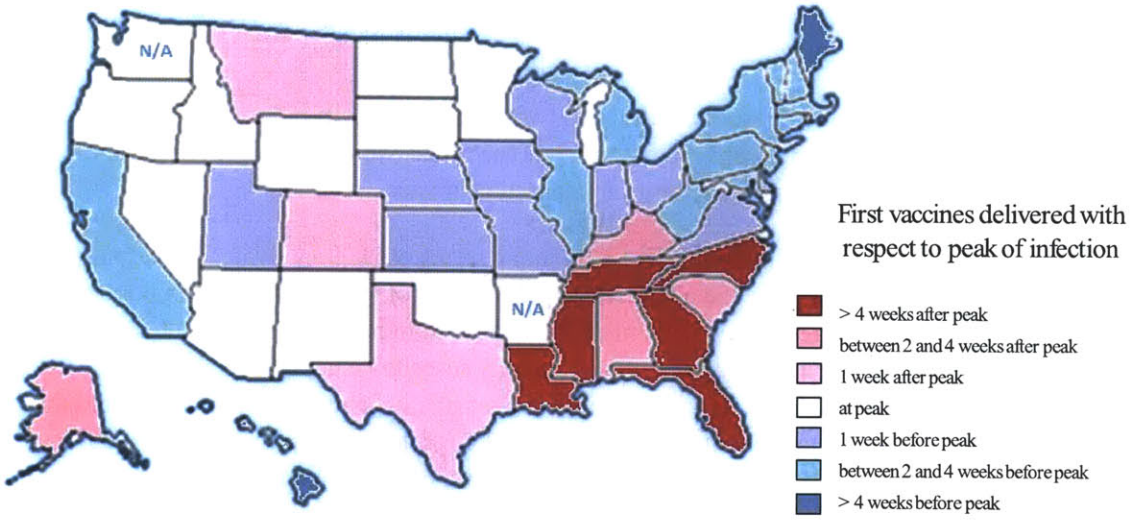


Figure 1: Timing of shipment of first vaccines with respect to the peak of the infection in US states.

This chapter is a natural follow-on to the Finkelstein *et al* paper as we demonstrate the potential loss of vaccine effectiveness with late shipments. We evaluate the effect of the timing of vaccination programs in 11 states, and suggest some lessons learned in the event of future pandemics and immunization efforts.

2. Methods

To estimate the number of infections averted from the various vaccine programs, we first use available data to estimate the epidemic curve during the period in the fall of 2009 when H1N1 was most prevalent. Once we obtain an estimated epidemic curve, we use it in conjunction with reported vaccine administration data to fit the observed epidemic curve to the one generated by a mathematical model based on difference equations. We use a discrete time version of the standard Kermack-McKendrick model to estimate infection spread within each state [**22, 23, 24**]. In the model calibration process, we estimate the relevant parameters such as R_0within each state. We estimate the R_0 for each state individually, since different states have different demographic, geographic, and cultural attributes. Moreover, states experienced the H1N1 outbreak at different times and implemented their vaccination programs in different ways [**4**], so the extent of the infection varied markedly.

We then estimate a different, *non-observable* flu wave curve, one assuming no available vaccine. This multi-step process provides a data-informed, model-supported basis for estimating the positive effects, if any, of the vaccine as administered in each of the states.

Before describing the methodology in full, it is worth noting some common drawbacks to using a discrete-time model that implies homogeneous mixing in the population. Each state's population is segmented into groups of varying susceptibility, infectivity, and activity levels, each of which may strongly influence the progression of the epidemic [**25, 26**]. While all models are imperfect, however, each region we investigate is large enough that homogeneous mixing provides a reasonable estimate for modeling the epidemic curve. This methodology has been used throughout the H1N1 outbreak [**15, 27**]. In addition, while epidemics occur in continuous time, our use of a discrete-time model both simplifies

calculation and allows us to incorporate the fact that during the 2009-2010 H1N1 pandemic vaccines were shipped on a weekly basis and vaccine stocks were increased at discrete time intervals.

2.1 Epidemic curve estimation

We first estimate the true epidemic curve within the different states, where the flu curve as experienced by states includes the effects of vaccinations. Since this virus was so prevalent in the US population, it was impossible to record an accurate epidemic curve in each state. There is no direct way to know the total number of people infected since 1) many such individuals did not present themselves to medical authorities, and 2) for those who did visit a physician, confirming tests for H1N1 were not routinely ordered. Instead, the data released by the states' health departments include the weekly percentage of all hospitalizations and outpatient visits resulting from influenza-like infections (%ILI) over the 2009-2010 flu seasons [5]. From that, we estimate the number of H1N1 infections for each week. We assume that the total number of cases is directly proportional to the total number of H1N1 hospitalizations and outpatient visits. While the %ILI curve provides good intuition for the spread of infection, we did not use it as the epidemic curve because the total number of hospitalizations and outpatient visits associated with all medical conditions changes throughout the year. Each point on the %ILI temporal curve represents the percentage of the total number of hospitalizations and outpatient visits that is specific to H1N1. Since the flu wave first grows and then declines, this total number is not uniform throughout the observation period. We expect a higher number of total hospitalizations and outpatient visits during the peak of H1N1, with the number of non-H1N1 hospitalizations and outpatient visits remaining relatively stable throughout the several-month observation period. Using that assumption, we perform a simple transformation on the % ILI curve to obtain a new curve that we consider to be directly proportional to the experienced flu wave curve.

We now go into some important mathematical details. To carry out the required transformation, we split our timeline into discrete "flu generation periods" with generations represented by $t = 1,2,3 \ldots t_{max}$. In our calculations, we let a single generation of the flu be equal to $\frac{10}{3} = 2.33$ days. Assuming just one number for the duration of the flu is a simplification. While actual generation of the flu varies with different viruses as well as on an individual level, we make this simplification to keep our discrete-time epidemic curve model tractable. The value we chose is consistent with analysis of H1N1 in Mexico done in May 2009 [**14, 17**].

Let

$C(t)$ = number of non-flu hospital patients in generation t.

Invoking our assumption of a constant number of such patients over time, we can write

$C(t) = C_0$, for $t = 1,2,3...$

Now consider the unobservable number of flu patients seeking hospital medical advice,

$N(t)$ = number of patients with H1N1 who visit a hospital in generation t, for $t = 1,2,3...$

What we observe in reported data is $F(t)$, the fraction of all hospital visits that are attributable to H1N1,

$$F(t) = \frac{N(t)}{N(t) + C_0}$$

Solving for $N(t)$, we can write

$$F(t)(N(t) + C_0) = N(t),$$

or

$$F(t)C_0 = N(t)[1 - F(t)],$$

or $N(t) = \frac{C_0 F(t)}{(1 - F(t))}$.

Invoking the assumption that the total number of H1N1 cases is proportional to the number of H1N1-related hospital visits, we deduce that the shape of the epidemic curve is proportional to $N(t)$. Since we have access to $F(t)$ only the value of C_0, the total number of non-H1N1 hospital visits per generation, is unknown. While we do not know that quantity, we can approximate the value of C_0 since we know the total number of H1N1 visits during our time period. Specifically, let I be the total number of H1N1 visits in a US state. Then,

$$\sum_{t=1}^{t_{max}} (N(t) + C_0) = \sum_{t=1}^{t_{max}} (\frac{C_0 F(t)}{1 - F(t)} + C_0) = I.$$

Consequently, we can write

$$C_0 = \frac{I}{t_{max} + \sum_{t=1}^{t_{max}} \frac{F(t)}{1 - F(t)}}$$

We approximated I using information from the CDC regarding the total number of infections in the United States.

This transformation gives us a closer estimate of the actual epidemic curve than the %ILI curve by itself. We are still faced with the fact, however, that people present to doctors and hospital at different rates throughout the pandemic. One might conjecture that the rate of hospital *visits* would be higher at the beginning of the pandemic, when media reports implied a real danger from flu-like symptoms, and that that rate would decrease toward the end of the epidemic, when reports confirmed that the consequences of H1N1 were relatively mild; but the rate of hospital *admissions* would not be affected by such psychological factors. Furthermore, there are still other factors, as parents of symptomatic children may

not wish to present their children to medical facilities for fear of them contracting the illness, should they not already have it. One can keep conjecturing ad infinitum. It is difficult to quantify these myriad possible effects so we do not include them in the transformation.

According to the CDC, the percentage of people infected with H1N1 in the US ranged from 14% to 29% of the population [**16**]. We tried a range of values for the total number of infections in the region and fit our model to each one by finding the parameters that minimize the mean square error between the epidemic curve and the resulting modeled curve. To find the best fit we modified three parameters: the total number of infections, the number of infections at the start of the epidemic ("Patient Zeros") and R_0. For each fit we took the resulting fitted value of R_0. The values were usually close together, and we chose the value of R_0 that best fit the exponential growth of cases near the onset of the epidemic.

2.2 A Set of difference equations

We use a set of difference equations that corresponds to a discrete time version to a classic SIR model [**22, 23, 24**].

For each state we adjust model parameters to obtain a best fit between the reported data, as transformed above, to the model-generated flu wave epidemic curve. This process leads to a direct estimation of R_0.

Define $R(t) \equiv$ mean number of new infections generated by a randomly infectious (asymptomatic) patient at generation t of the epidemic. $t = 1, 2, 3, \ldots$

2.2.1 The epidemic curve

Suppose the number of infected people at "Generation Zero" is 1 person. That is, there is one "Patient Zero." Then, the mean number of infected in Generation 1 is R_0 persons, each of whom subsequently infect on average $R(1)$ persons in Generation 2, and so on. So, in Generation t, the mean number of people infected $MI(t)$ will be the product $R_0 * R(1) * R(2) * \ldots * R(t\text{-}1)$. More compactly, assuming that $MI(0) = 1$, we can write

$$MI(t) = R_0 \prod_{n=1}^{t-1} R(n) \quad t = 1,2,\ldots \tag{2}$$

Where MI(t) = mean number of infections at time t.

2.2.2 Determining R(t)

Suppose the total population is N persons, all susceptible at the beginning of the epidemic and all mixing homogeneously – whether or not they are susceptible or have recovered and have immunity. Then suppose at Generation t we have $I(t)$ people recovered and immune, re-circulating in the population. Due to homogeneous mixing, R_0 is now modified by a factor of $\frac{[N - I(t) - MI(t)]}{N}$, to become

$$R(t) = R_0 \frac{N - I(t) - MI(t)}{N} \quad t = 1,2,... \tag{3}$$

2.2.3 Determining I(t)

At Generation t we assume that there are a total of $I(t)$ individuals immune to the disease. We can generalize this immunity, as it can derive from a vaccine immunity or from being recovered from the disease and now immune. For simplicity in this model we assume that individuals infected in Generation (t-1) become recovered and immune in Generation t. We will also assume that during the time of Generation (t-1), $V(t\text{-}1)$ individuals received vaccine that will make then immune for the first time in Generation t. Thus we can write a difference equation for determining the value of $I(t)$:

$$I(t) = I(t-1) + MI(t-1) + V(t-1) \quad t = 1,2,... \tag{4}$$

We now have a model consisting of the following 3 equations where V(t) is known for all t.

$$MI(t) = R_0 \prod_{n=1}^{t-1} R(n) \quad t = 1,2,...$$

$$R(t) = R_0 \frac{N - I(t) - MI(t)}{N} \quad t = 1,2,...$$

$$I(t) = I(t-1) + MI(t-1) + V(t-1) \quad t = 1,2,...$$

These equations can be solved iteratively, via recursion starting with the boundary conditions $MI(0)$, $I(0)$, and $R(0) = R_0$.

2.3 Estimating the effects of vaccine

2.3.1 The non-observable curve

Suppose, in the absence of vaccines, we have a pandemic infection curve $C(t)$. Here,

$C(t)$ = Number of new flu infections reported during generation t of the infection period, assuming no vaccine. $t = 1, 2...$

We call $C(t)$ the *base curve*, that is, the infection wave that occurs in the absence of vaccines. Also, for the sake of simplicity we will assume that human behavior as illustrated for instance by hygienic steps and other non-pharmaceutical interventions (NPIs) remains unchanged during the course of the infection period. This base curve is unobservable because the state-specific data reporting numbers of people infected included the effects of administered vaccines.

2.3.2 Timing of vaccinations and subsequent immunity

To depict the effect of vaccinations on the population, we recall the definition,

$V(t)$ = number of individuals who have been vaccinated and first acquired immunity during generation t, $t = 1, 2, \ldots$

We assume that only susceptibles receive vaccinations.

On average, vaccines take effect about 2 weeks after they have been administered. In our approximations we compare the numbers calculated in the case where the vaccines take effect immediately as well as the case where vaccines take effect after 2 weeks. Similarly we consider vaccine effectiveness (i.e., creating immunity to H1N1) to be between 75% and 100%. According to preliminary studies, the effectiveness of H1N1 vaccine was well over 90% in 10 days [**28**], so the actual effect of the vaccines should be well within the bounds of our analysis.

2.3.3 The observable curve

Now we have another pandemic infection curve,

$C(t, \vec{V}) \equiv$ Number of new flu infections reported during generation t of the flu, given that immunities due to vaccinations occur according to the known time vector $\vec{V}$, $t = 1, 2\ldots$

Our state-derived data depict the vaccine-affected flu wave, $C(t, \vec{V})$ and we wish to infer the (unobservable) vaccine-free flu wave $C(t)$. To estimate the effect of the vaccinations, we need to estimate $C(t)$ and then compare to $C(t, \vec{V})$. The difference in the areas under the respective curves represents our estimate of the number of infections averted due to the vaccine.

To estimate R_0, we fit the model-generated $C(t, \vec{V})$ curve to the empirically estimated epidemic curve, which includes the effects of vaccinations. The influenza cases during the H1N1 epidemic were severely underreported. The underreporting appeared to be especially significant and inconsistent at the beginning and end of the outbreak [**16, 29**]. To avoid this 'statistical noise" effect we fit the model-based epidemic curve to the data surrounding the peak of the epidemic, specifically to the part of the empirical curve that contains 75% of the cases. To find a best reasonable fit, we used the difference equations (2), (3), and (4), and manipulated the parameters R_0 and $MI(0)$. The best fit was determined to be one for which the peak of the model-determined epidemic curve coincided with the estimated real epidemic curve and exponential growth of the early stages of the outbreak coincided with the real epidemic curve. The R_0 obtained using this method might be a slight underestimate of the true R_0 for the H1N1 pandemic as it is not fitted at the very beginning of the outbreak when data are very noisy, but only once clear exponential growth is established. This methodology, however, is consistently used in the literature and provides a reasonable estimate for the basic reproductive number [**13, 30**].

3. Results

3.1 Oklahoma in detail

Consider as an illustrative example the estimation process for Oklahoma. Figure 2 contains 1) the empirical based estimated epidemic curve for Oklahoma using the transformation described above, and 2)

the time-sequenced vaccine administration data reported by the state. In Figure 3 we again include the empirical epidemic curve, and three model-generated epidemic curves:

- The curve generated by using the Oklahoma reported vaccine administration data, fitted to correspond best to the empirical epidemic curve.
- The curve generated in the hypothetical case where vaccines were not administered at all.
- The curve generated in the hypothetical case where vaccines were administered 2 weeks earlier than had actually occurred.

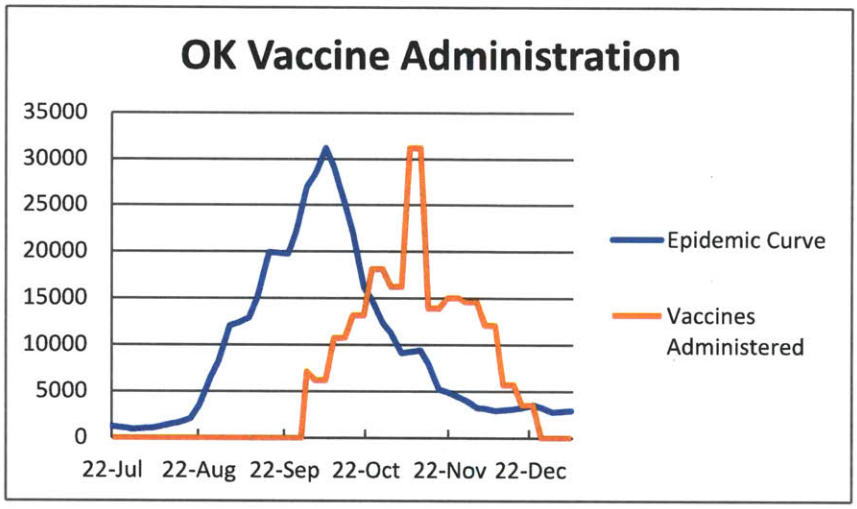


Figure 2: The estimated Oklahoma epidemic curve compared to vaccines as they were administered in the state

The total number of infections caused by the outbreak in Oklahoma is calculated by the area under an epidemic curve. The effect of the vaccines administered in Oklahoma can be determined by calculating the area between "actual" model-generated curve, and the "no-vaccine" model-generated curve (Figure 4).

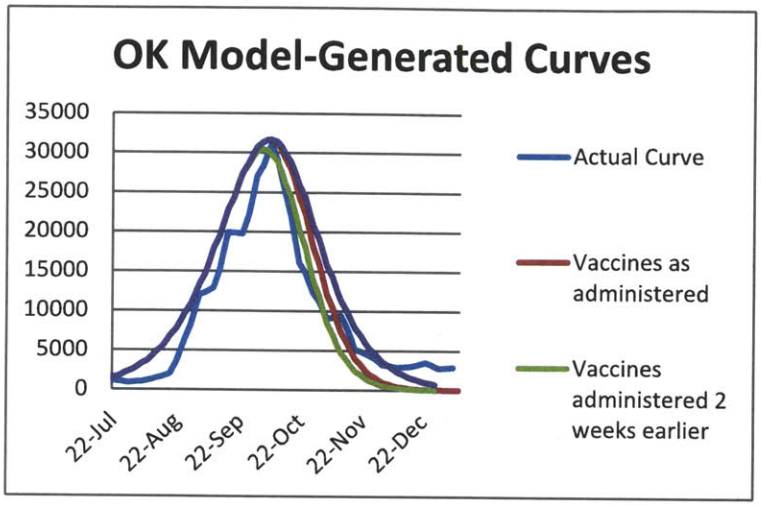


Figure 3: The estimated epidemic curve along with the model-generated curves with and without vaccines. The best fit for Oklahoma was found with parameters $R_0 = 1.14, MI(0) = 1500$, and $I(0) = 0$.

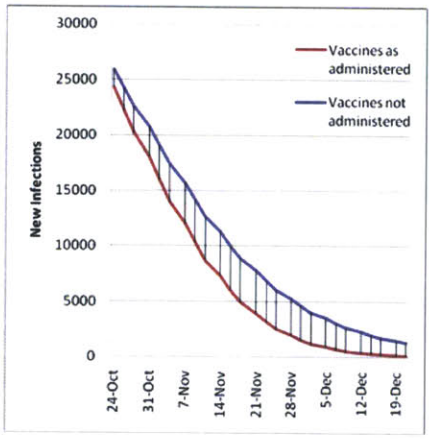


Figure 4: A closer look at Figure 3. The shaded region represents the difference between the estimated number of H1N1 infections that would have occurred without intervention of vaccines and the estimated number of infections that actually occurred with the vaccines.

We analyzed 11 states in detail, and inferred the total number of infections that have been prevented as a result of their respective immunization programs. The states Illinois, Indiana, Massachusetts, Mississippi, Montana, New Jersey, New York, North Dakota, Oklahoma, South Carolina, and Virginia graciously provided us with precise data on vaccines as they were dispensed throughout the outbreak. Here we display two cases for each state:

(1) The optimistic case, in which all vaccines are effective immediately and are 100% effective, and

(2) The pessimistic case, in which vaccines are effective 2 weeks after administration and are effective for only 75% of the individuals receiving the vaccine.

The actual effect of the vaccine should lie within the range specified by these two cases.

It is possible that some ILI-related cases were reported with a delay of one or two weeks with respect to symptom onset, since infected individuals may be more likely to visit a medical professional after symptoms become severe. If that is the case, our estimated epidemic curve might be shifted to the left by a few generations. If true, the infections occurred even earlier than what we had assumed, and so vaccine arrival occurred even later with respect to the peak of infections. This would imply that our results are an optimistic estimate of the effectiveness of states' vaccine programs.

3.2 Summary

In Table 1 we present the results of analysis of 11 states with detailed vaccination information. While we only have detailed data for these 11 states, we note that range of estimated values for R_0 is fairly narrow with higher values falling on more populated states such as New York and Illinois, and lower values corresponding to states with lower density such as Oklahoma and Montana. Geographic proximity also seems to result in similar R_0 values.

State	Estimated R_0	% Population Vaccinated	Optimistic Infections Averted (in % of total population)	Pessimistic Infections Averted (in % of total population)
Illinois	1.21	9	3.43	1.2
Indiana	1.15	20	4.28	1.81
Massachusetts	1.16	29	13.71	6.84
Mississippi	1.16	8	0.13	0.05
Montana	1.15	20	2.81	1.04
New Jersey	1.20	12	3.36	1.1
New York	1.20	14	3.23	1.12
North Dakota	1.16	27	2.95	1.06
Oklahoma	1.14	13	2.29	0.93
South Carolina	1.16	8	0.4	0.12
Virginia	1.19	22	1.77	0.52

Table 1: Summary of best fitted values for R_0 and model-determined effects of vaccines as they were distributed. IN the optimistic scenario vaccines are 100% effective and take effect immediately. In the pessimistic scenario vaccines are 75% effective and take effect 2 weeks after being administered.

4. Discussion

4.1 Vaccine effectiveness

While examining the estimated numbers of infections averted, we can identify two major contributing factors. The first is the total number vaccines administered to the general population. The second is the timing of the vaccine administration with respect to the peak of the infections. Once the H1N1 virus had been identified as a potentially devastating pandemic in the spring of 2009, the CDC worked to develop and distribute H1N1 vaccines. These vaccines were sent out to the individual states at the same time, and first doses were administered on October 5, 2009. These vaccines, however, had varying effects as the peak of the outbreak in different states occurred at markedly different times.

The peak of infection usually occurs when herd immunity occurs, that is the time when $R(t) = 1$, and every infectious person at that time infects on average just one other person [**31**]. So at the time of herd immunity, the number of infections in the next generation is approximately the same as it was in the previous one. We say 'approximately' due to statistical fluctuations in the actual number of susceptible people that any one newly infectious person will infect. Soon afterwards, infectious people no longer replace themselves in society, and so the number of infected and infectious people in each generation decreases. Early administration of vaccines decreases the number of people who still need to be infected to reach herd immunity, and so decreases the height of the peak. Late administration of the vaccine has almost no effect on the dynamics of the outbreak, and has little benefit to the society other than immunizing the people who received the vaccine. Such late immunizations may be important if the flu were to return later in a new wave.

Consider again the southeastern states of the US, the first region to report infection peaks. As early victims they received vaccines after the worst of the infection had already passed. Louisiana, Indiana, and South Carolina had not started administering vaccines until *after* the peak of outbreak. And these states were least successful in averting infections. On the other hand, Massachusetts and Virginia started administering their vaccines five and three weeks, respectively, before their respective peaks. These two states enjoyed a particularly good impact from their vaccination programs. In addition to having five weeks of vaccinations prior to the H1N1 peak in Massachusetts, that state vaccinated 29% of its population, the most of any state in our sample. As a result, as much as 7-14% of the population may have been spared infection and possible complications from influenza. While Massachusetts and Virginia had effective experiences with their vaccinations, most states did not. On average for our limited sample, vaccines were delivered just before the peak of the states' outbreaks.

To quantify the effect of time in averting infection we consider one of the states that vaccinated almost 20% of its population, Indiana. The hypothetical case where the same number of vaccines is delivered just two weeks earlier resulted in averting more than twice the number of infections than the vaccines as administered.

With a more granular approach, we considered the marginal benefit of administering just one vaccination at a given time. We mapped the total projected number of infections that could be averted if just one vaccination were to be administered to a susceptible person at different times during the outbreak. That is, we calculated the total number of infections that would occur in Oklahoma if exactly one vaccine were administered at different points in time, and compared that number to the total number of infectious that would happen if no vaccines were administered at all. The differences are presented in Figure 5.

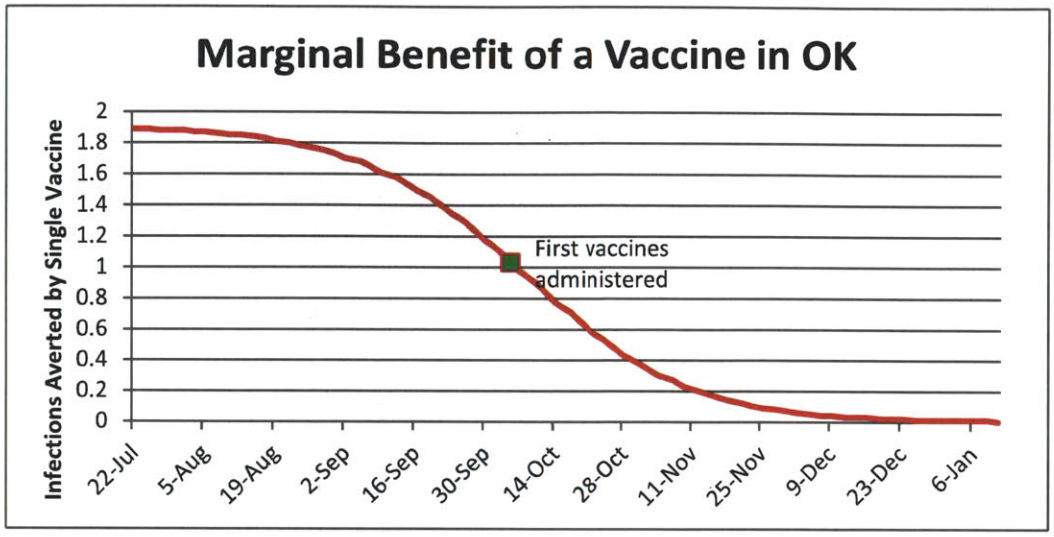


Figure 5: The number of infections averted by administering exactly one vaccination at different points throughout the outbreak in Oklahoma.

As expected, administered vaccines have a monotonically decreasing benefit with respect to time. A striking feature of Figure 5 is the fact that one vaccination to a susceptible person well before the flu wave starts averts almost two infections in the population, even with a low value for R_0 (1.14) and even considering the fact that the vaccinated person has a greater than even chance of never becoming infected assuming no vaccination. Clearly, vaccines administered well before the peak carry the added benefit of diluting the susceptible population with immune people and are particularly useful in mitigating the spread of infection.

Another insightful feature of this graph is the slope of the marginal benefit curve, which represents the time-dependence of effective vaccines. While starting vaccine administration in Indiana in July would be most effective for Indiana, the effect of those vaccines would not change significantly until the beginning of September. That is, if vaccines were to be available in July, Indiana could have waited to receive its share until September with minimal losses. Similarly, vaccines received after December will have the same (minimal) effect whether they are administered in December or February. The effectiveness of vaccines, however, is extremely time-sensitive from the end of September to mid-November, where each week results in a significant loss of effectiveness. Vaccines that become available during this critical period need to be administered as soon as possible. These results encourage us to recommend a more detailed cost-benefit analysis of trying to get some vaccine, even if in much smaller quantities, to the states at the beginning of this "critical period," when the population is particularly sensitive to the timing of vaccination. A small amount of vaccine delivered early should have a more significant effect on the total number of infections than a batch delivered just a few weeks later.

Furthermore, we conjecture that the timeliness of the vaccines is also closely related to the total amount of vaccine accepted by the population. While the CDC distributed its vaccines proportionally to the population of each region, states vary in the amount of vaccine that was actually used. For instance, Mississippi used less than 40% of its allocated vaccine, most likely due to "flu fatigue". While the media are particularly helpful at warning the public of an ongoing pandemic and encouraging the use of NPI's,

they can also give the impression that the outbreak is over, or has been blown out of proportion. Once the peak of the outbreak had passed and H1N1 had been determined to be less dangerous than originally feared, the populations of the "early victim" states would be less likely to spend their time and risk perceived possible side effects getting a flu shot.

Looking at all 50 states and comparing the percentage of vaccinated population by the end of the outbreak to the week of vaccine delivery [**32**], we notice in Figure 6 a weak negative correlation between the timing of vaccine delivery with respect to the peak, and the total percentage of the population that accepted vaccinations. These results are consistent with the position of Harris *et al*, that had the vaccines been delivered earlier to the states, more people would have been encouraged to accept a vaccination [**19**].

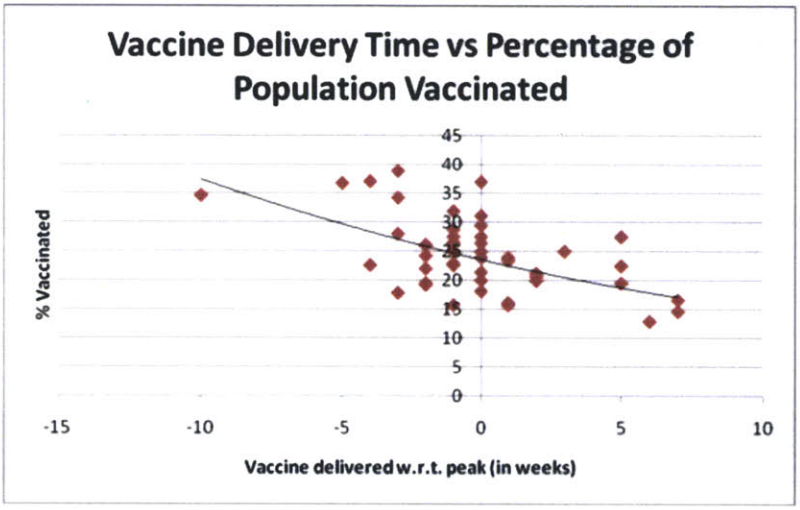


Figure 6: The relationship between the timing of vaccine delivery with respect to the peak and the percentage of the state population that received the H1N1 vaccine for all 50 states.

This effect would be particularly relevant to South Carolina, which started administering its vaccines one week after the peak of infection and subsequently managed to vaccinate only 8% of its population. If those vaccinations were to start earlier, before the peak, not only could we hypothesize an increase in effectiveness from timing alone, but a higher participation rate in the vaccination program. So, early administration is particularly important in that it increases the efficacy of vaccine along with encouraging people to actually accept vaccination.

4.2 Vaccine allocation

As shown in Figure 7, in the first few weeks of vaccine distribution, when demand for vaccine clearly exceeded supply, the CDC allocated vaccine to states proportionally to their populations [**3**]. Particularly in early October, this simple distribution scheme ensured that all states received amounts that could be used to immunize the same proportion of the population. Come November, those states that saw little demand started placing fewer orders for influenza vaccine, while those with later epidemics like Massachusetts and Virginia were still experiencing high demand and were shipped larger quantities of vaccine, confirming the intuition from Figure 6.

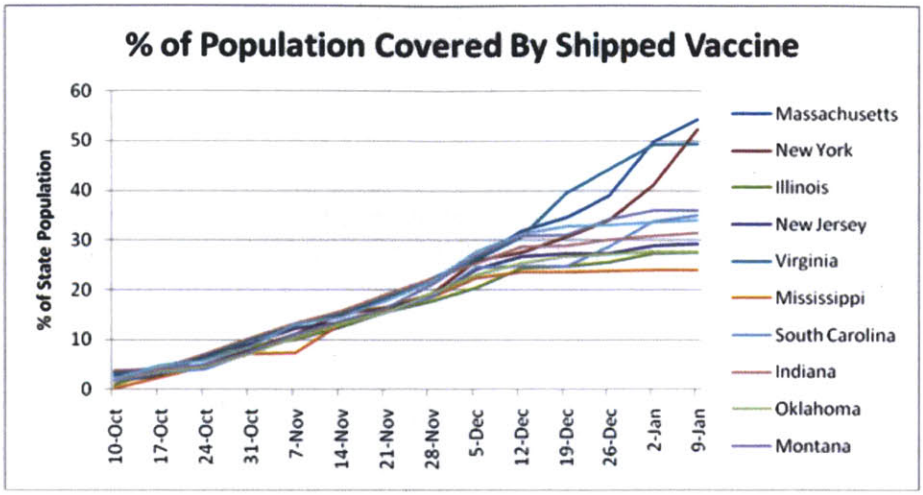


Figure 7: During the early stages of the outbreak when the demand for vaccines outweighed the supply, the CDC distributed vaccines proportionally to the population of the states.

The previous sections imply that the same vaccines administered in states that already experienced the peak of the infection at the time vaccines started arriving were much less effective than those administered in states that had not yet experienced the peak. Moreover, the states that were past the peak saw less demand for vaccines, and thus used only a small fraction of its allocated vaccine. Consider a side-by-side analysis of Mississippi and North Dakota in Table 2.

North Dakota	Mississippi
• Started administration 1 week before peak	• Received vaccines 6 weeks after peak
• A hypothetical single batch of 32,342 vaccines (5% of the population) administered on October 10, averts 33,745 infections (5.2% of the population)	• A hypothetical single batch of 147,599 vaccines (5% of the population) administered on October 10, averts 27,173 infections (0.9% of the population)
• In the first 4 weeks administered 62% of available vaccines	• In the first 4 weeks administered less than 25% of available vaccines

Table 2: Comparison of vaccination programs in North Dakota and Mississippi.

It is clear that vaccines administered in North Dakota were significantly more effective than those in Mississippi. In fact, the Mississippi vaccines had almost no effect because the infection was barely spreading by the time vaccines became available. Coupled with this, North Dakota was experiencing more demand for the vaccines at the beginning of its program. Motivated by this we believe that there is a need for more effective procedures of allocating vaccines to US States.

Naturally, allocating all of Mississippi's vaccine to North Dakota would not only be unethical but politically infeasible. Instead, as a thought experiment, suppose that just 20% of Mississippi's *unused* vaccine were to be transferred to North Dakota during the first 4 weeks of vaccine distribution. Suppose that with this addition, 60% of the new vaccines were actually administered. This additional vaccine

would decrease the total number of infections in North Dakota by 5%. That is a significant improvement for a relatively small cost. An adaptive decision like this can be made during the allocation process. We can form even approximate predictions about how much vaccine will actually be demanded by the state and how effective the extra vaccines would be in reducing infections. For example, by using data collected from our 11 states, we can weakly estimate that a state that had experienced peak infection six weeks earlier can be expected to vaccinate no more than 4% of its population within the first four weeks. In the first four weeks of the 2009 H1N1 outbreak, the CDC allocated to Mississippi enough vaccine to cover 7% of its population. With accurate data, some portion of that could have been redirected to states that are more likely to use and benefit from the vaccine.

5. Conclusions

Our analysis shows the importance of the timing of vaccinations for infectious respiratory diseases such as influenza. We emphasize the need to start administering vaccines well before the peak of an influenza outbreak. Moreover, when a governing body such as the CDC is faced with the allocation decision, it is important to take into consideration the stages of the outbreaks in different regions and to deploy vaccine with preference to the regions that are projected to administer more of the vaccine with greater beneficial effects.

In our analysis of individual state immunization programs we used a relatively simple model, assuming a homogeneous population and a deterministic model structure. As a result, our model, as most models, is not an exact picture of what happened in during the fall of 2009, but rather a tool to gain insight about strategies that could help the public mitigate the effect of influenza pandemics. While the exact numbers almost surely differ from the estimates, the relative results should hold under a range of assumptions about R_0, vaccine efficacy, and generation period of the flu.

Please note that the content in this chapter has been published in Value in Health [**33**].

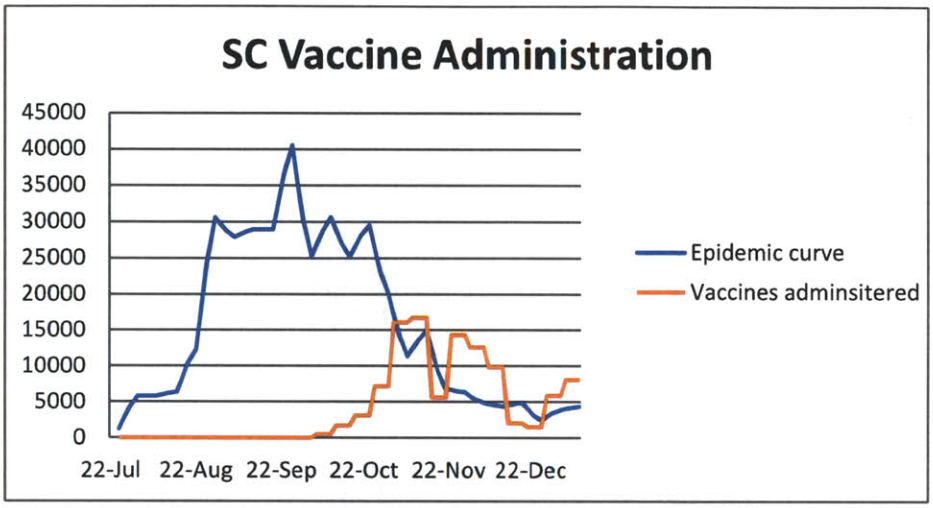
References

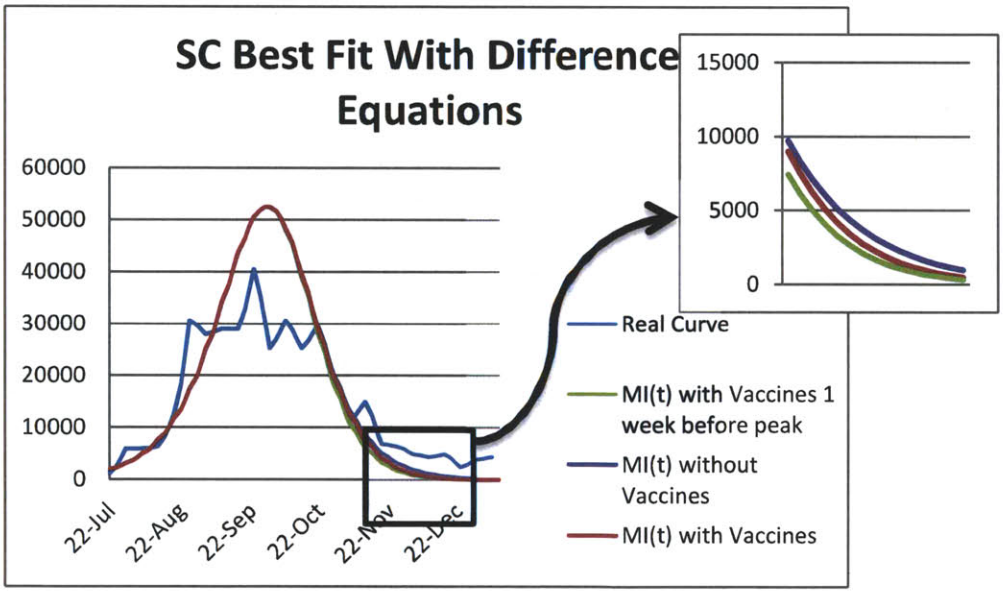
1. Centers for Disease Control and Prevention. The 2009 H1N1 Pandemic: Summary Highlights, April 2009-April 2010. June 16, 2010. Available at: http://www.cdc.gov/h1n1flu/cdcresponse.htm. Accessed January 12, 2011.
2. Centers for Disease Control and Prevention. Press Briefing Transcripts: Media Briefing: Update on 2009 H1N1 Flu. September 3, 2009. Available at: http://www.cdc.gov/media/transcripts/2009/t090903.htm. Accessed January 20, 2011.
3. Centers for Disease Control and Prevention. 2009 H1N1 Vaccine Doses Allocated, Ordered, and Shipped by Project Area. February 2, 2010. Available at: http://www.cdc.gov/h1n1flu/vaccination/vaccinesupply.htm. Accessed January 20, 2011.
4. Hopkins J. H1N1 After-action reports: lessons on vaccine distribution. *MIT ESD Working Paper.* 2011.
5. Centers for Disease Control and Prevention. *CDC, Seasonal influenza (flu), Weekly Report: Influenza summary update.* May 28, 2010. Available at: 2010. Accessed July 29, http://www.cdc.gov/flu/weekly.
6. Longini I, Halloran E. Strategy for Distribution of Influenza Vaccine to High-Risk Groups and Children. *Am J Epidemiol.* 2005;161(4):303-306.
7. Patel R, Longini I, Halloran M. Finding optimal vaccination strategies for pandemic influenza using genetic algorithms. *Journal of Theoretical Biology.* May 2005;234(2):201-212.
8. Chowell G, Viboud C, Wang X, Bertozzi S, Miller M. Adaptive vaccination strategies to mitigate pandemic influenza: Mexico as a case study. *PLoS One.* 2009;4(12).
9. Larson R, Nigmatulina K. Engineering Responses to Pandemics. *Info Knowl Sys Manag.* 2009;8:311-339.
10. Nigmatulina K, Larson R. Living with influenza: Impacts of government imposed and voluntarily selected interventions. *EJOR.* 2009;195(2):613-627.
11. Khazeni N, Hutton D, Garber A, Hupert N, Owens D. Effectiveness and cost-effectiveness of vaccination against pandemic influenza (H1N1) 2009. *Ann Intern Med.* 2009;151:829-831.
12. Sander B, Bauch C, Fishman D, et al. Is a mass immunization program for pandemic (H1N1) 2009 good value for money? Early evidence from Canadian experience. *Vaccine.* 2010;28:6210-6220.
13. Fraser C, Donnelly CCS, Hanage W, et al. Pandemic potential of a strain of influenza A (H1N1): early findings. *Science.* 2009;1557-1561:324.
14. White L, Wallinga J, Finelli L, Reed C. Estimation of the reproductive number and the serial interval in early phase of the 2009 influenza A/H1N1 pandemic in the. *Influenza and Other Respiratory Viruses.* September 2009;3(6):267-276.
15. Towers S, Feng Z. Pandemic H1N1 influenza: predicting the course of a pandemic and assessing the efficacy of the planned vaccination programme in the United States. *Euro Surveill.* 2009;13(41).
16. Centers for Disease Control and Prevention. CDC Estimates of 2009 H1N1 Influenza Cases, Hospitalizations and Deaths in the United States, April 2009 – March 13, 2010. April 19, 2010. Available at: http://www.cdc.gov/h1n1flu/estimates/April_March_13.htm. Accessed January 20, 2012.
17. Tuite A, Greer A, Whelan M, et al. Estimated epidemiologic parameters and morbidity associated with pandemic H1N1 influenza. *CMAJ.* 2010;182(2):131-136.
18. Finkelstein S, Hedberg K, Hopkins J, Hashmi S, RC L. Vaccine availability in the United States during the 2009 H1N1 outbreak. *Am J Disaster Med.* Jan-Feb 2011;6(1):23-30.
19. Harris K, Maurer J, Kellermanmn A. Influenza Vaccine - Safe, Effective, and Mistrusted. *N Engl J Med.* 2010;363:2183-2185.

20. Moyer C. H1N1 vaccine: what physicians can do with leftover doses. July 19, 2009. Available at: http://www.ama-assn.org/amednews/2010/07/19/prsb0719.htm. Accessed March 18, 2012.
21. SteelFisher G, Blendon R, Bekheit M, Lubell K. The public's response to the 2009 H1N1 influenza pandemic. *N Engl J Med.* 2010.
22. Kermack W, McKendrick A. A contribution to the mathematical theory of epidemics. *Proc. Roy. Soc. Lond.* 1927;115:700-721.
23. Keeling N, Rohani P. *Modeling infectious diseases in humans and animals.* Princeton, NJ: Princeton Univ Press; 2007.
24. Murray J. *Mathematical Biology.* 2 ed. Berlin: Springer; 1989.
25. Larson R. Simple models of influenza progression within a heterogeneous population. *Operations Research.* 2007;55(3):399-412.
26. Nigmatulina K. Modeling and Responding to Pandemic Influenza: Importance of Population Distributional Attributes and Non-Pharmaceutical Interventions (Doctoral dissertation).
27. Coburn B, Wagner B, Blower S. Modeling influenza epidemics and pandemics: insights into the future of swine flu (H1N1). *BMC Medicine.* 2009;7(30).
28. Greenberg M, Lai M, Hartel G, Wichems C. Response after one dose of a monovalent influenza A (H1N1) 2009 vaccine - preliminary report. *N Engl J med.* 2009;361:2405-2413.
29. Centers for Disease Control and Prevention. Underreporting of 2009 H1N1 influenza cases. Available at: http://www2c.cdc.gov/podcasts/media/pdf/EID_12-09_FluEstimates.pdf. Accessed June 10, 2011.
30. Mills C, Robins J, Lipsitch M. Transmissibility of 1918 pandemic influenza. *Nature.* 2004;432:904-906.
31. Anderson R. The concept of herd immunity and the design of community-based immunization programmes. *Vaccine.* 1992;10:928-35.
32. Centers for Disease Control and Prevention. Morbidity and Mortality Report. Interim results: State-Specific Influenza A (H1N1) 2009 Monovalent Vaccination Coverage -- United States, October 2009 -- January 2010. Available at: http://www.cdc.gov/mmwr/preview/mmwrhtml/mm5912a2.htm. Accessed January 20, 2011.
33. Larson R, Teytelman A. Modeling the effects of H1N1 influenza vaccine distribution in the United States. *Value in Health.* January 2012;15(1):158-166.

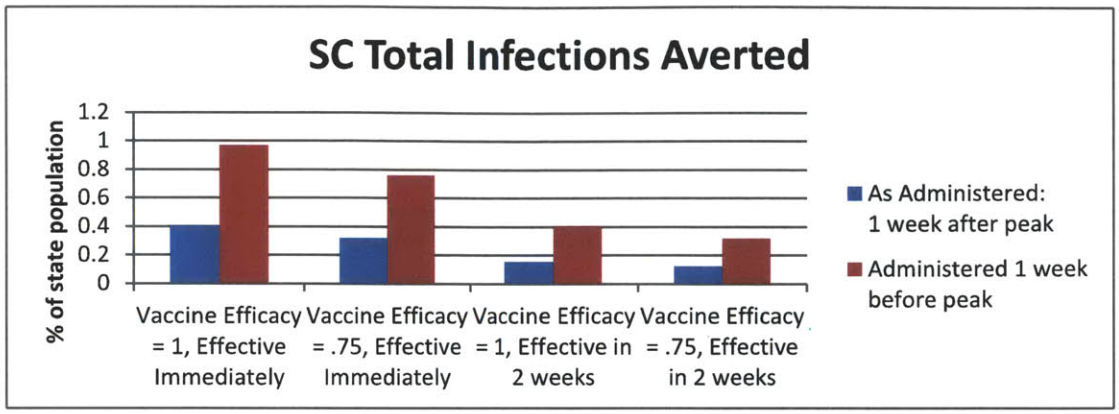
Appendix A: Detailed state curves
(all data used are valid as of August 2010)

South Carolina:

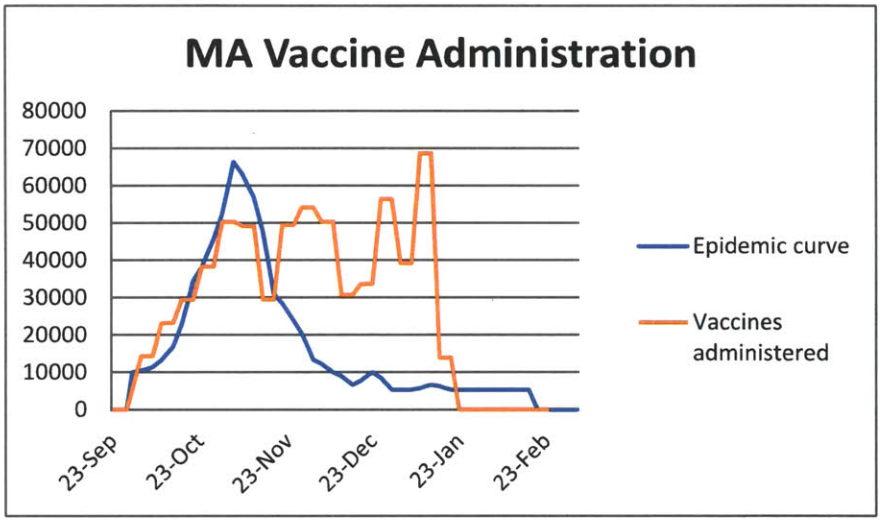


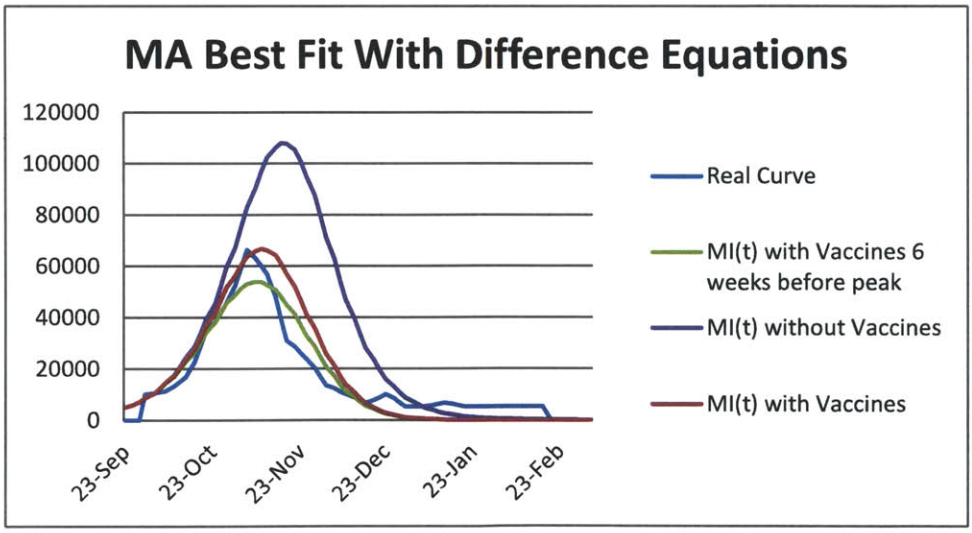


% of Population Vaccinated	Vaccination start with respect to the peak (in weeks)	Exponential Growth R_0
8	1	1.17

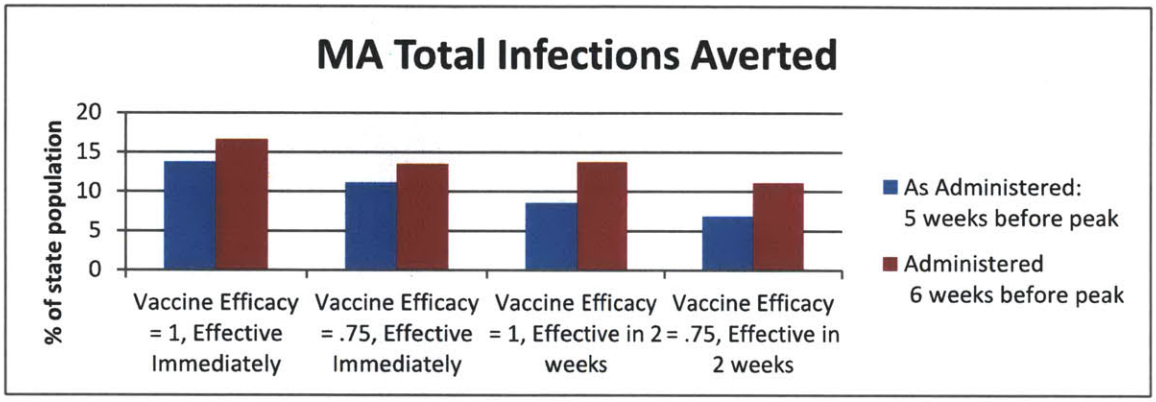


Massachusetts:

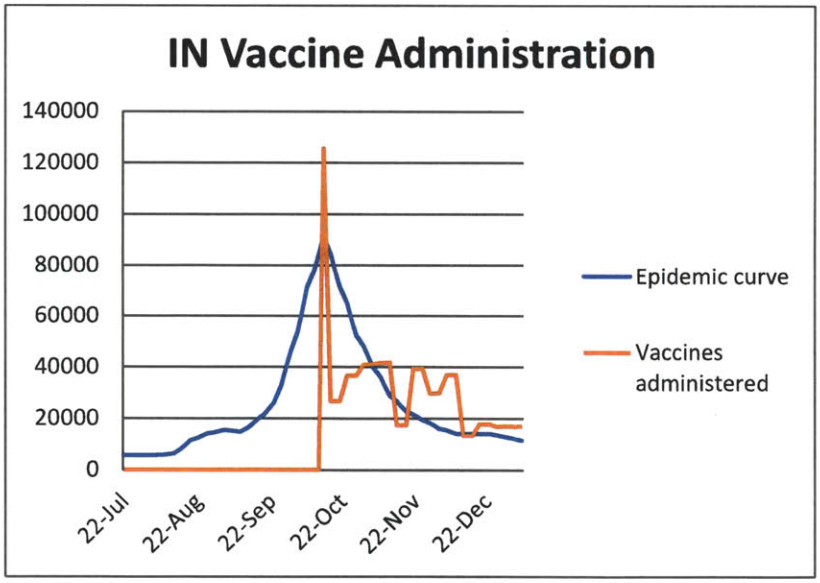


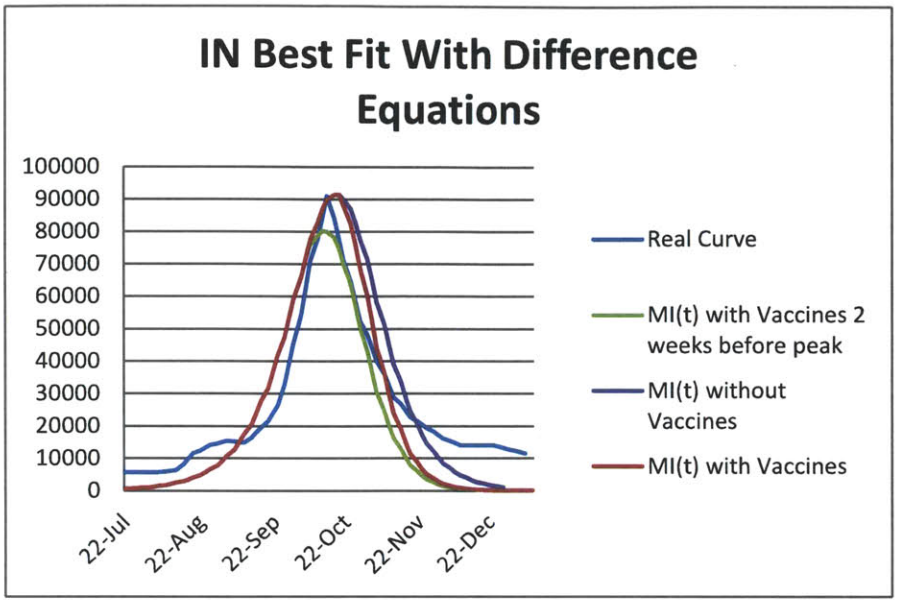


% of Population Vaccinated	Vaccination start with respect to the peak (in weeks)	Exponential Growth R_0
29	-5	1.20

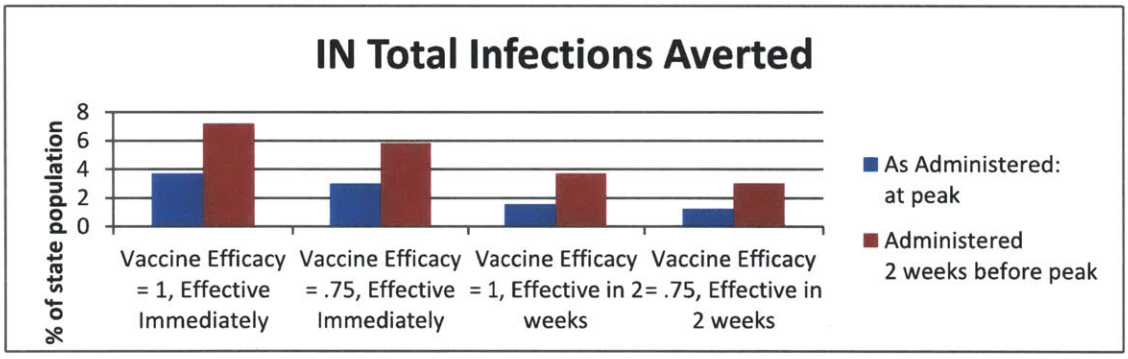


Indiana:

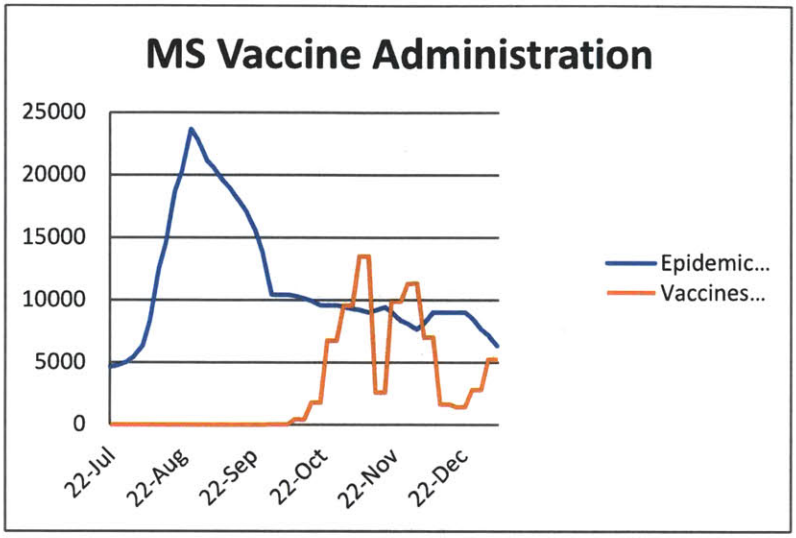


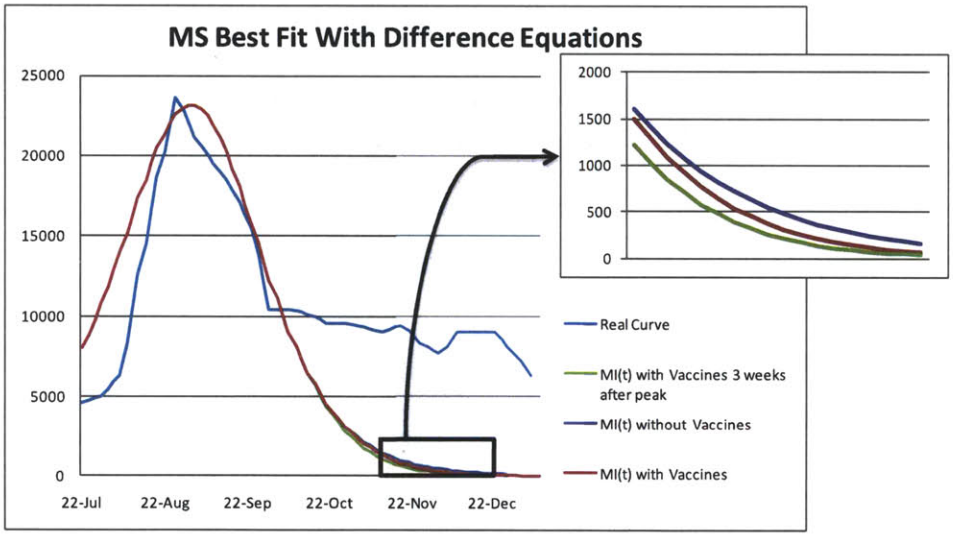


% of Population Vaccinated	Vaccination start with respect to the peak (in weeks)	Exponential Growth R_0
20	0	1.19

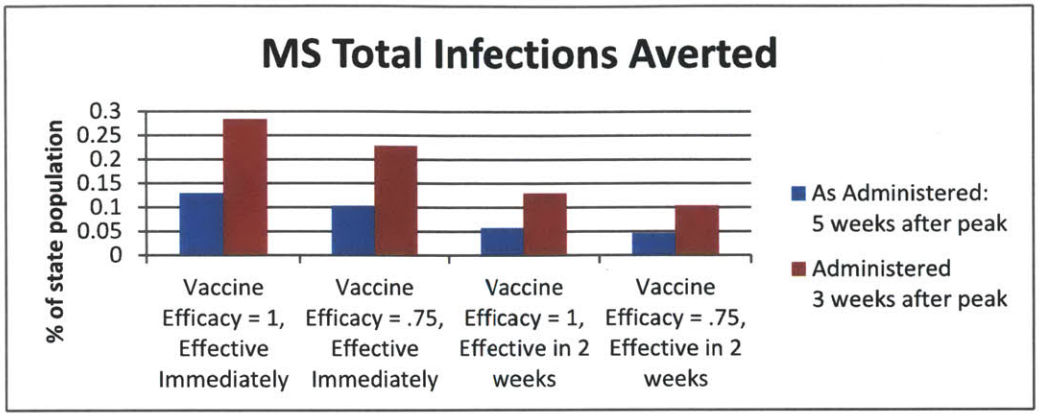


Mississippi:

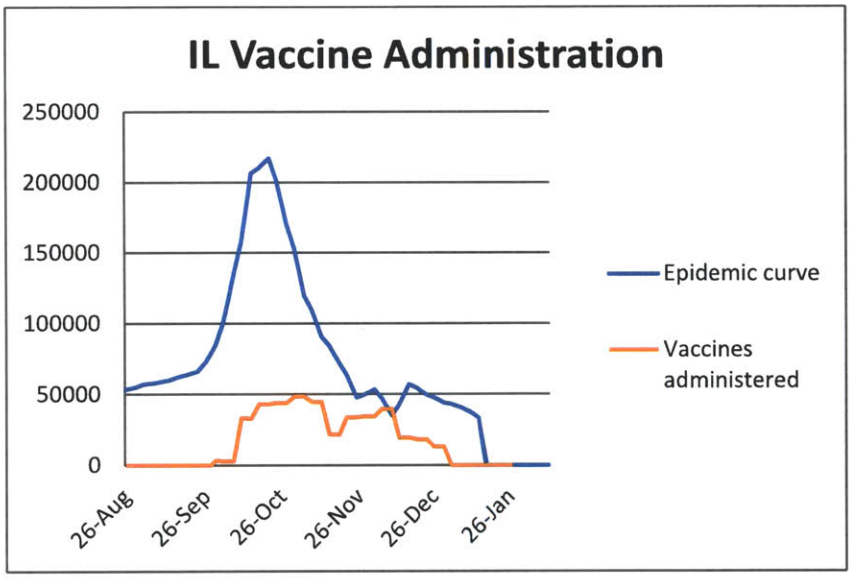


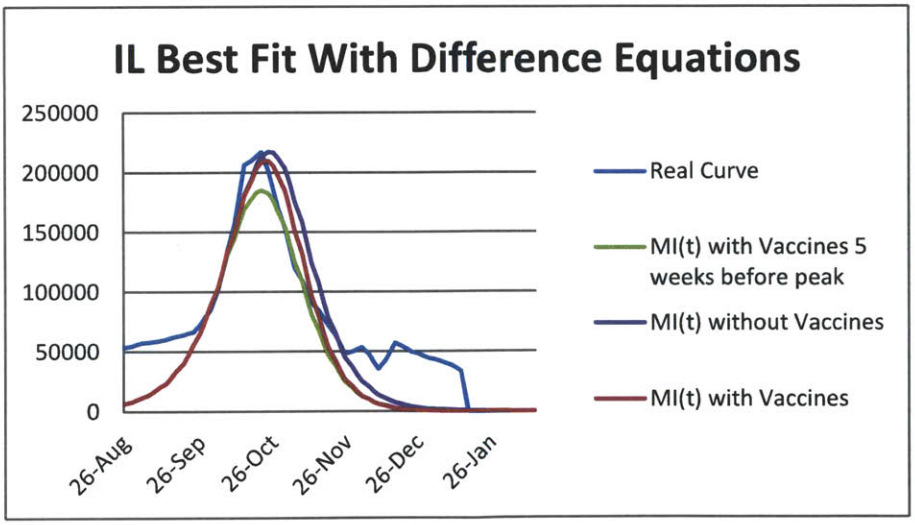


% of Population Vaccinated	Vaccination start with respect to the peak (in weeks)	Exponential Growth R_0
8	5	1.11

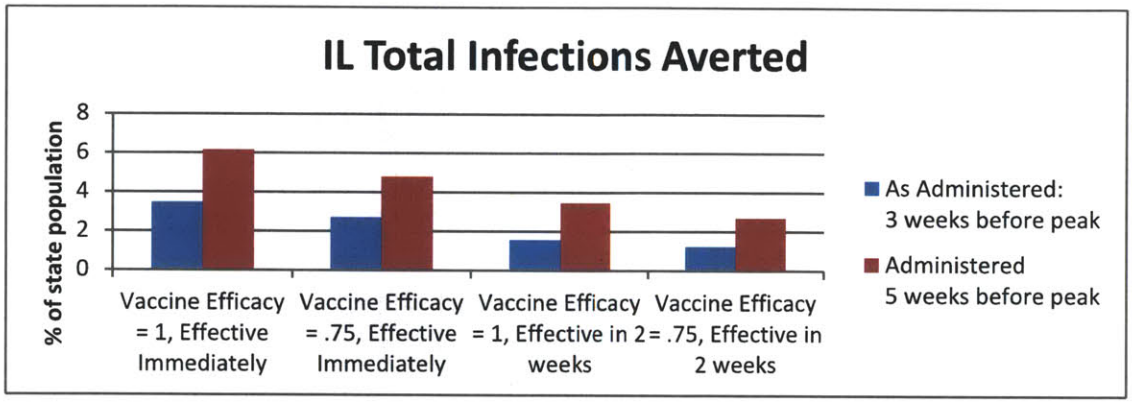

MS Total Infections Averted
% of state population
0.3
0.25
0.2
0.15
0.1
0.05
0
Vaccine Efficacy = 1, Effective Immediately
Vaccine Efficacy = .75, Effective Immediately
Vaccine Efficacy = 1, Effective in 2 weeks
Vaccine Efficacy = .75, Effective in 2 weeks
As Administered: 5 weeks after peak
Administered 3 weeks after peak

Illinois:

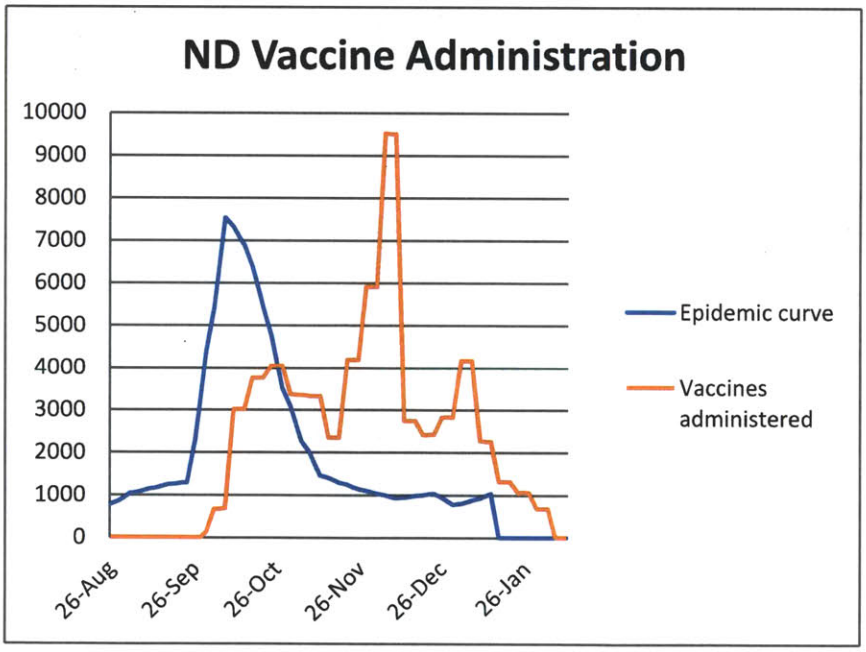

IL Vaccine Administration
250000
200000
150000
100000
50000
0
26-Aug
26-Sep
26-Oct
26-Nov
26-Dec
26-Jan
Epidemic curve
Vaccines administered

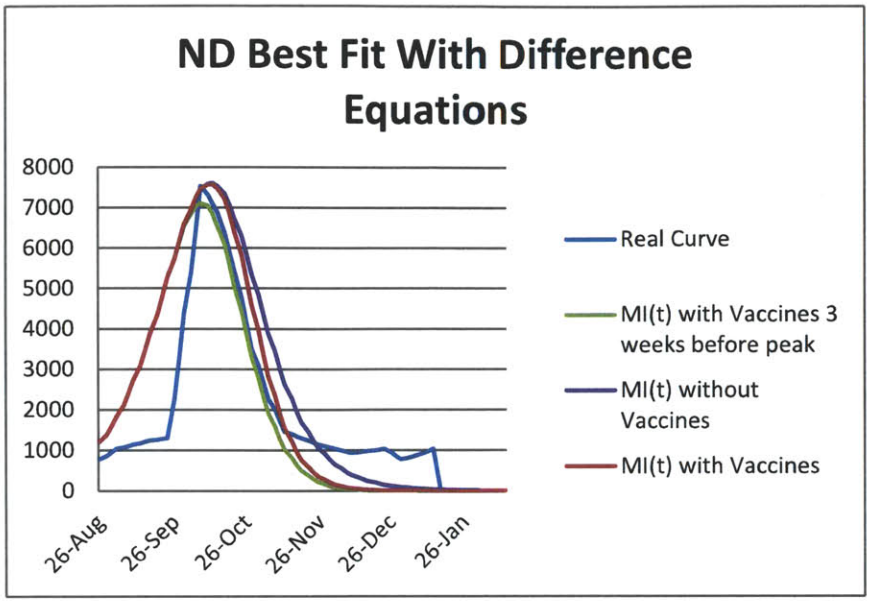

IL Best Fit With Difference Equations
250000
200000
150000
100000
50000
0
26-Aug
26-Sep
26-Oct
26-Nov
26-Dec
26-Jan
Real Curve
MI(t) with Vaccines 5 weeks before peak
MI(t) without Vaccines
MI(t) with Vaccines

% of Population Vaccinated	Vaccination start with respect to the peak (in weeks)	Exponential Growth R_0
9	-3	1.21

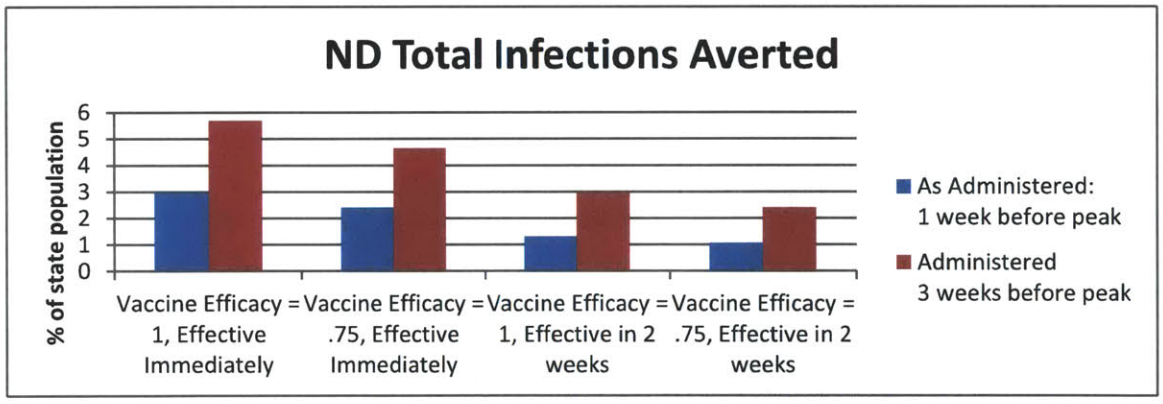


North Dakota:

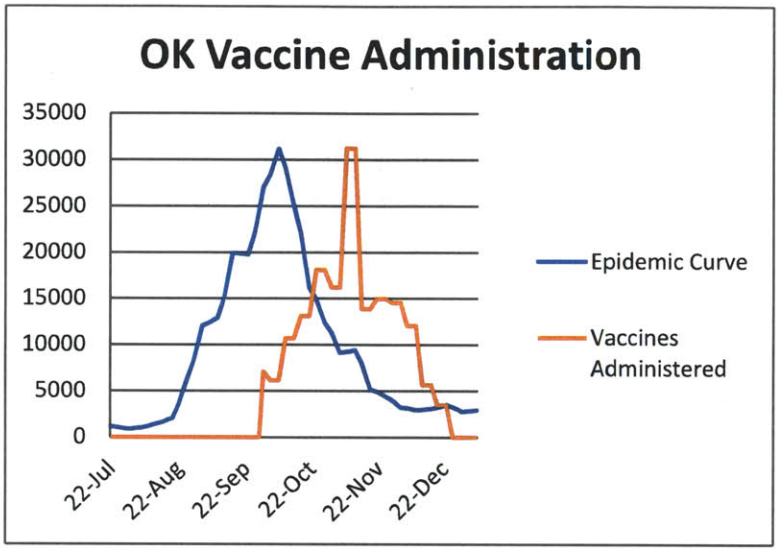


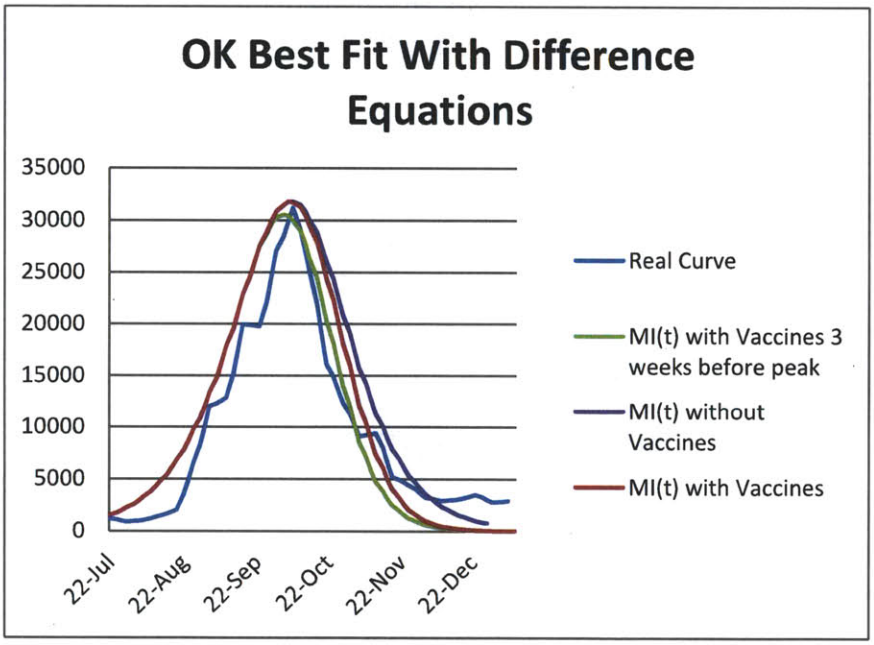


% of Population Vaccinated	Vaccination start with respect to the peak (in weeks)	Exponential Growth R_0
27	-1	1.16

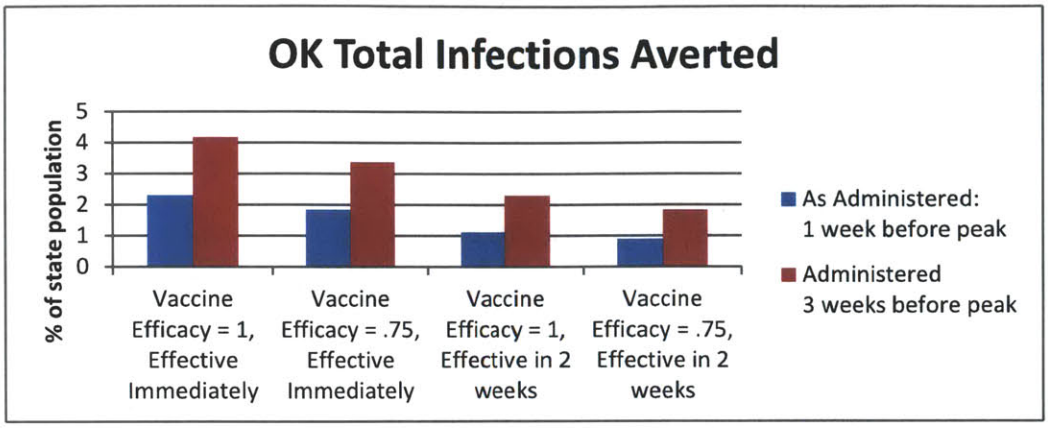


Oklahoma:

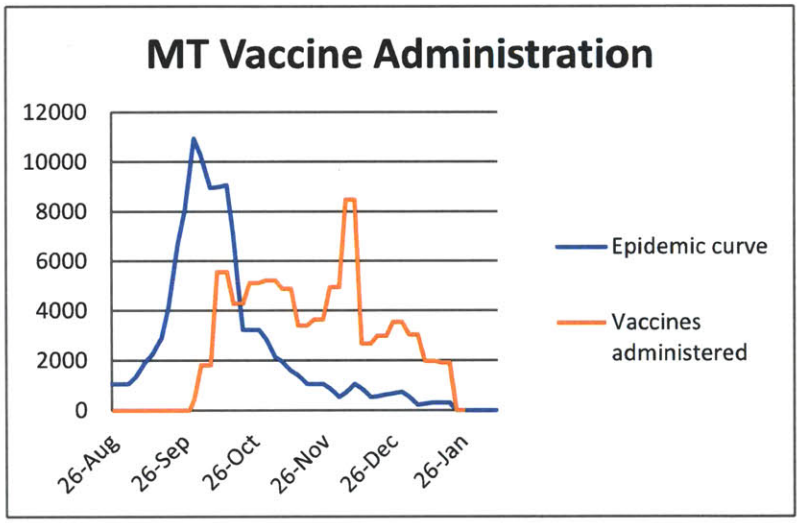


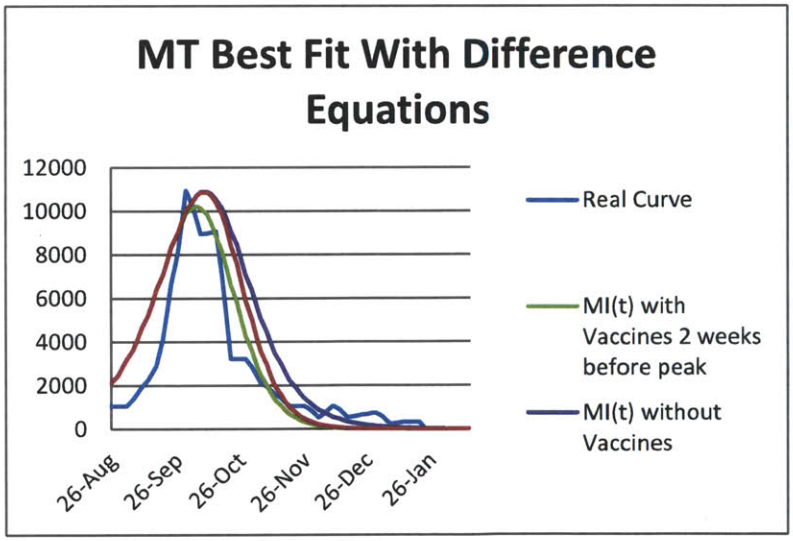


% of Population Vaccinated	Vaccination start with respect to the peak (in weeks)	Exponential Growth R_0
13	-1	1.14

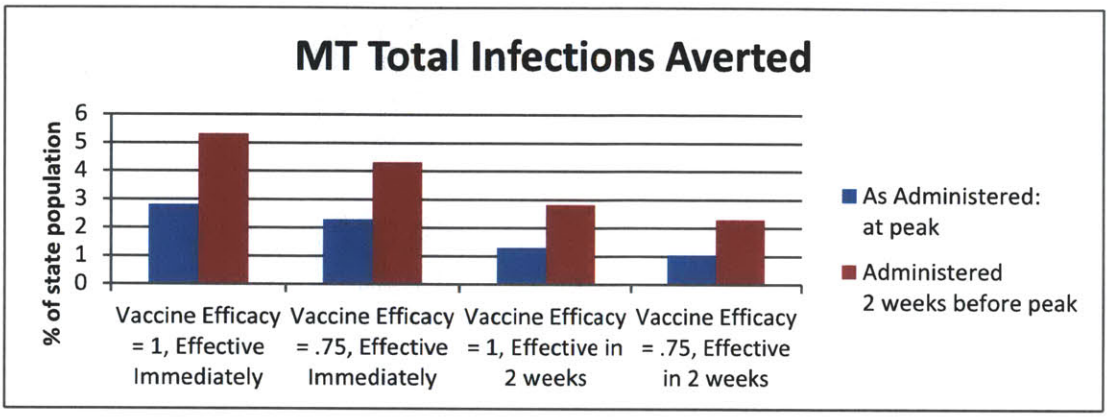

OK Total Infections Averted
% of state population
5
4
3
2
1
0
Vaccine Efficacy = 1, Effective Immediately
Vaccine Efficacy = .75, Effective Immediately
Vaccine Efficacy = 1, Effective in 2 weeks
Vaccine Efficacy = .75, Effective in 2 weeks
As Administered: 1 week before peak
Administered 3 weeks before peak

Montana:

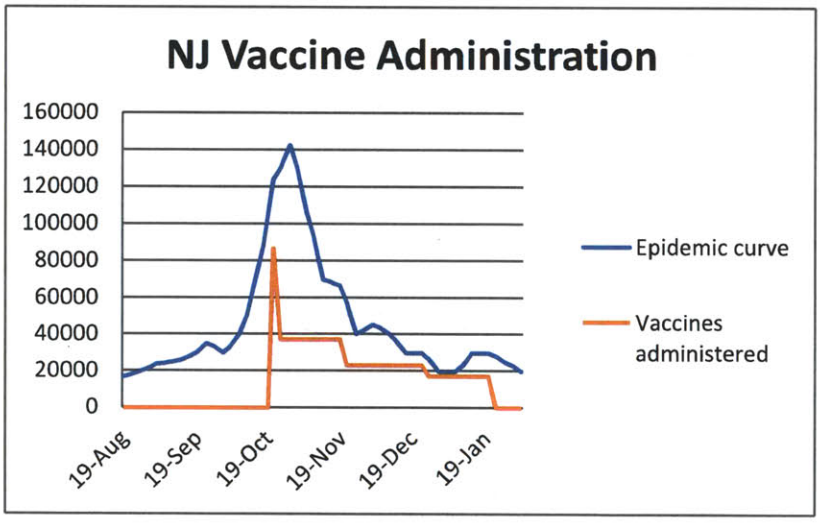

MT Vaccine Administration
12000
10000
8000
6000
4000
2000
0
26-Aug
26-Sep
26-Oct
26-Nov
26-Dec
26-Jan
Epidemic curve
Vaccines administered

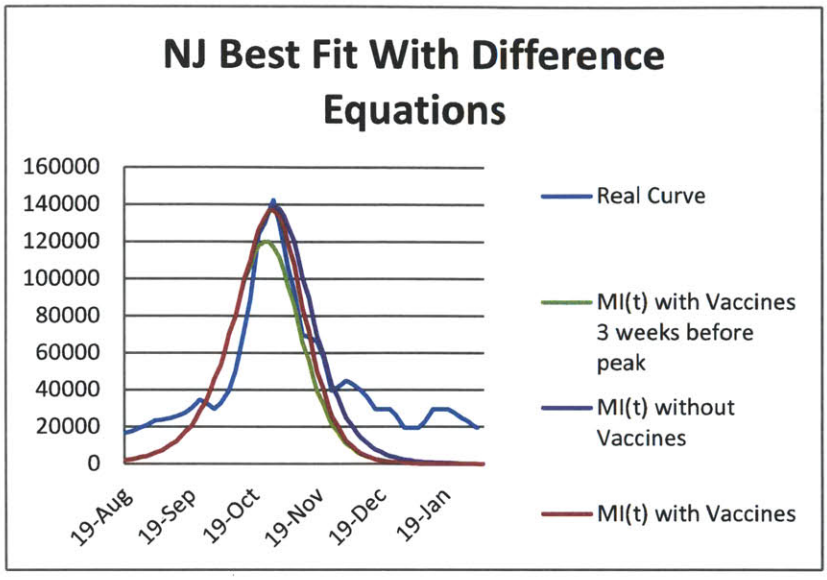

MT Best Fit With Difference Equations
12000
10000
8000
6000
4000
2000
0
26-Aug
26-Sep
26-Oct
26-Nov
26-Dec
26-Jan
Real Curve
MI(t) with Vaccines 2 weeks before peak
MI(t) without Vaccines

% of Population Vaccinated	Vaccination start with respect to the peak (in weeks)	Exponential Growth R_0
20	0	1.15

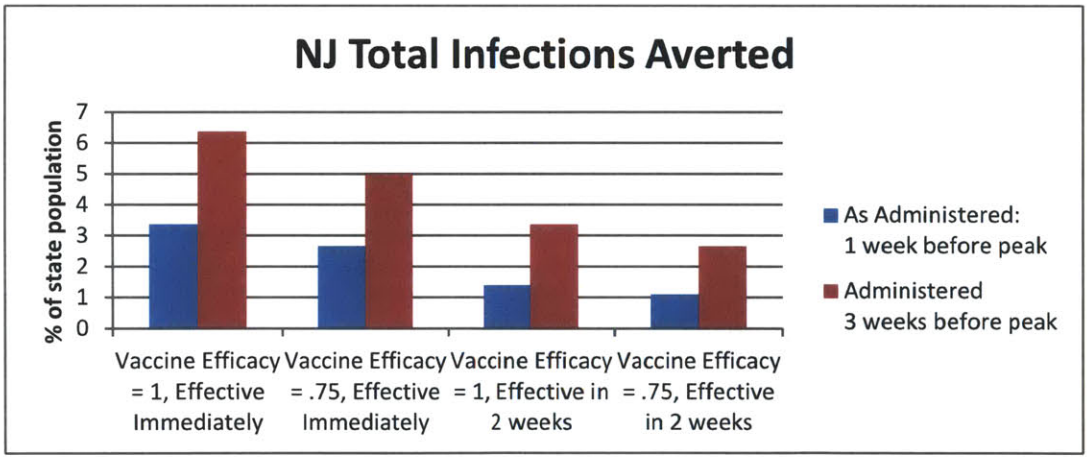


New Jersey:

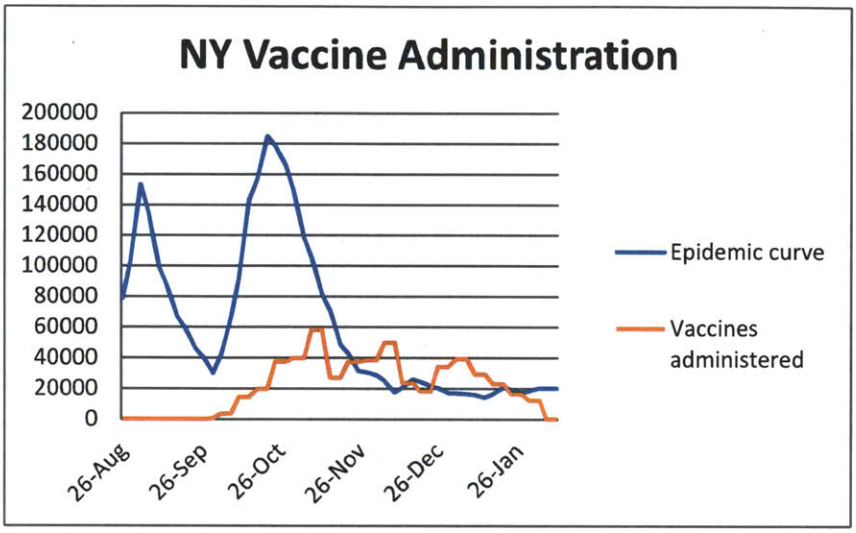


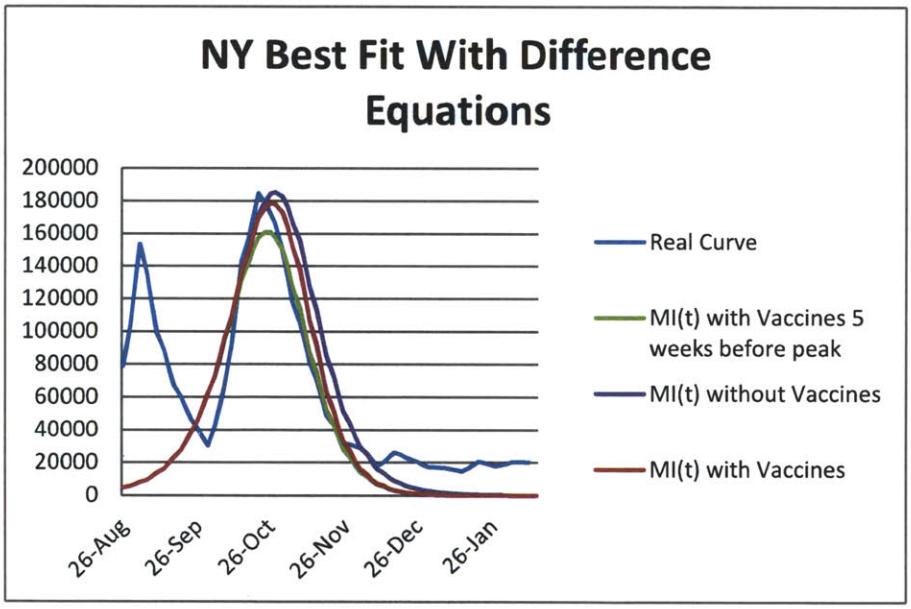


% of Population Vaccinated	Vaccination start with respect to the peak (in weeks)	Exponential Growth R_0
12	-1	1.20

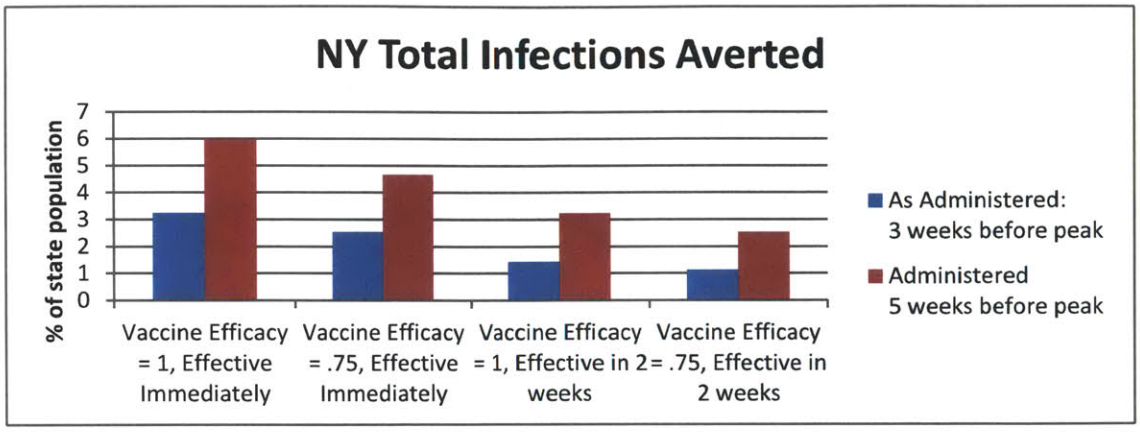


New York:

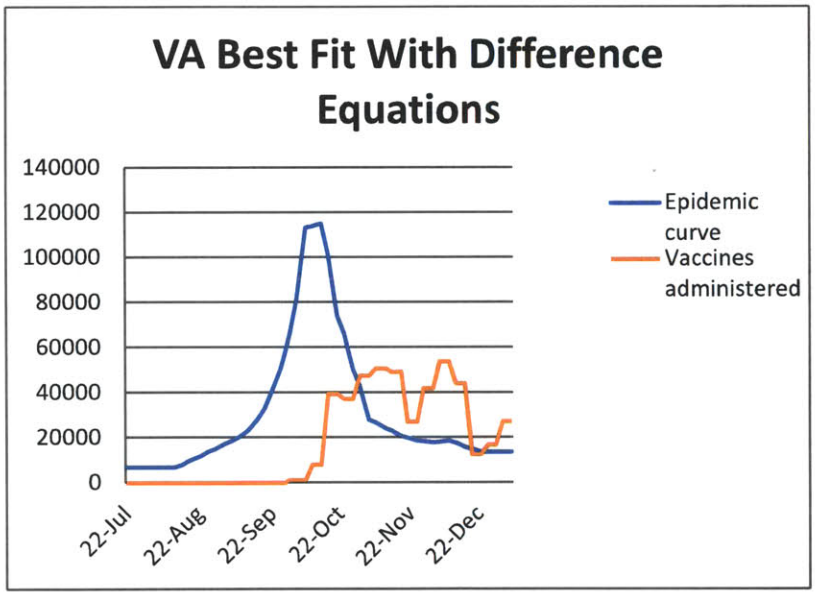


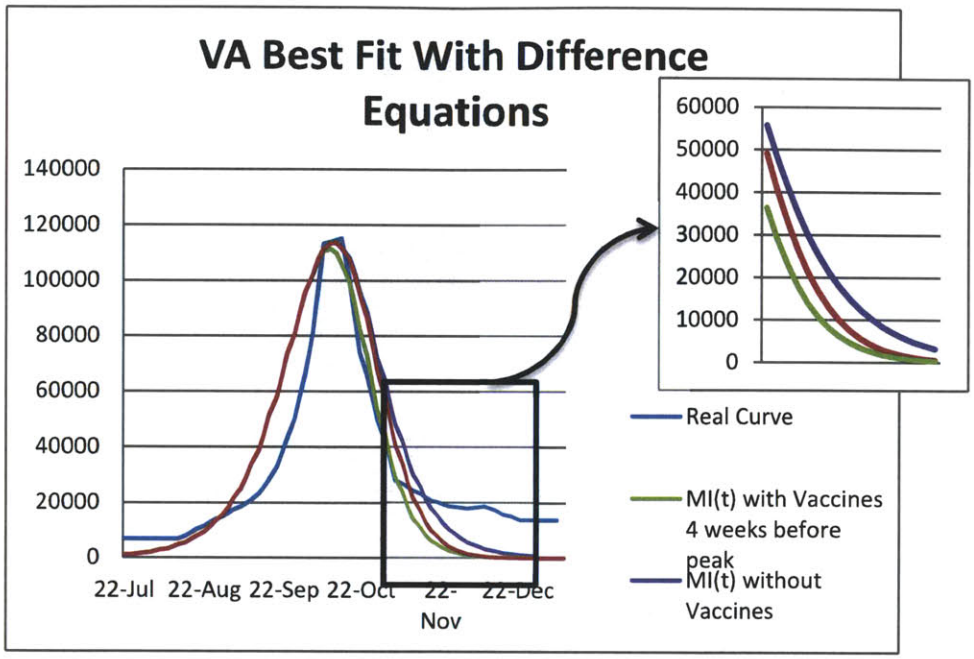


% of Population Vaccinated	Vaccination start with respect to the peak (in weeks)	Exponential Growth R_0
14	-3	1.20

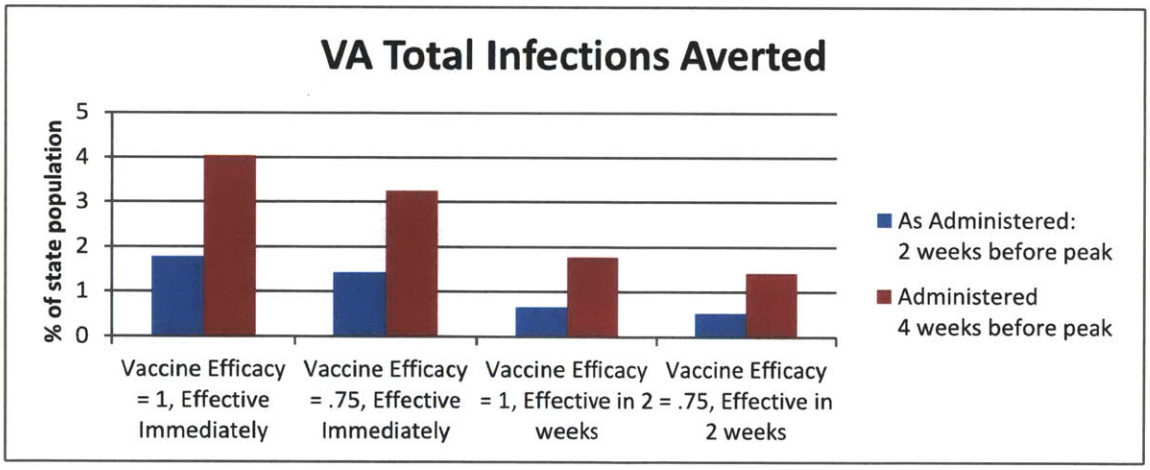

NY Total Infections Averted
% of state population
7
6
5
4
3
2
1
0
Vaccine Efficacy = 1, Effective Immediately
Vaccine Efficacy = .75, Effective Immediately
Vaccine Efficacy = 1, Effective in 2 weeks
Vaccine Efficacy = .75, Effective in 2 weeks
As Administered: 3 weeks before peak
Administered 5 weeks before peak

Virginia:


VA Best Fit With Difference Equations
140000
120000
100000
80000
60000
40000
20000
0
22-Jul
22-Aug
22-Sep
22-Oct
22-Nov
22-Dec
Epidemic curve
Vaccines administered



% of Population Vaccinated	Vaccination start with respect to the peak (in weeks)	Exponential Growth R_0
22	-2	1.19



Chapter V: Multi-regional dynamic vaccine allocation

In the previous chapter we discussed the importance of timing in vaccine effectiveness while looking individually at 11 US states. We would like to explore how we can use this information to solve a higher-level problem of distributing a single stock of available vaccines in an effective manner to many such regions. As mentioned before, in the United States the job of distributing vaccines is done by the CDC using the pro rata heuristic, where every constituent state gets the number of vaccines proportional to its population regardless of the state's demographics, R_0, or the state of the epidemic progression in the region. In this chapter we will show that other allocation heuristics that take into account the state of infection in each region can be significantly more effective in mitigating the spread of influenza.

1. Motivation

Recall the marginal benefit of a single vaccine if it were administered at different points in time in Oklahoma. The graph in Figure 1 perfectly illustrates that timing is particularly important in the effectiveness of a single vaccine.

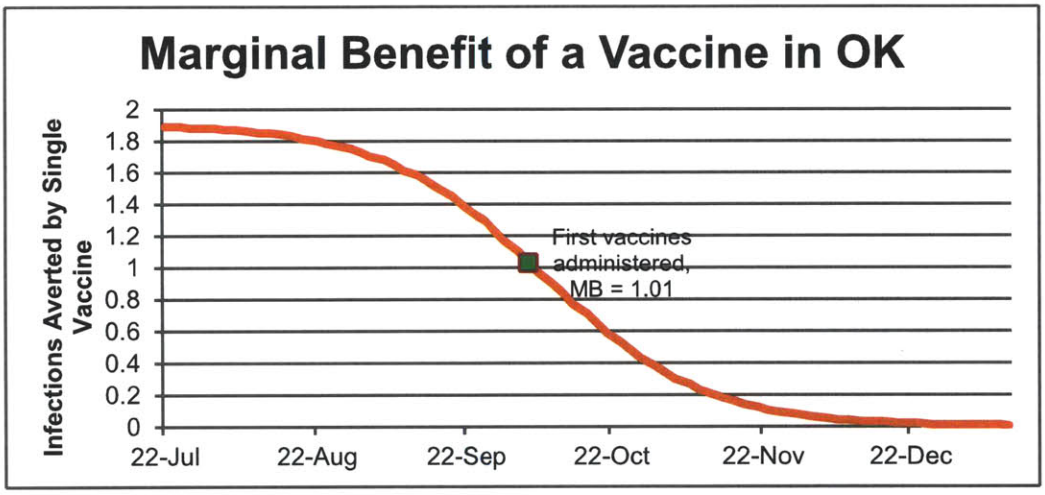


Figure 1: Marginal benefit of a single vaccine administered at different times in Oklahoma based on data from the 2009-2010 H1N1 pandemic.

If a region receives vaccines early, a single vaccine can have save almost two people, while if a region receives vaccines late, the vaccines are virtually wasted. Note also the slope of the curve, which represents the time-dependence of vaccine effectiveness. While the curve is strictly monotonically decreasing, it is relatively flat in July and early August as well as after the beginning of December. That is, the relative effectiveness of a vaccination would be virtually unchanged if we were to administer a vaccine in July or beginning of August. However, a week's delay in September would decrease the effectiveness of a vaccine significantly. We thus arrive at what we label a *critical period*, ranging from the middle of August until the end of November, when quick administration is of the essence for stopping Oklahoma's influenza spread.

In a multi-region problem, where different regions are at different stages of an influenza epidemic, we want to exploit this property and focus primarily on the regions where the effect of vaccines is particularly time-dependent. While states that are not within their respective critical periods are less likely

to suffer from waiting a few weeks for their vaccination, we must balance this with the fact that early administration is always preferable to later vaccinations.

2. Basic problem

The multi-region allocation problem has to be solved by a governing body several times throughout the outbreak. Each time a shipment of vaccines becomes available, the vaccines need to be allocated and shipped to different regions. At each of these decision points, the CDC needs to make a decision in two major steps.

STEP 1: The first stage is to gather the modeling information for each region. The CDC needs to use the available information such as the partially approximated epidemic curves for each region to fit the parameters that will be used by some underlying model to approximate the expected progression of the epidemic in each region. As long as a single consistent model is being used throughout, the allocation algorithm is can be independent of model choice.

STEP 2: The second stage of the decision making process is the allocation stage. In this stage, we assume that we already have a calibrated model for each region. Given some estimated models for n different regions, we must make a decision to allocate vaccines to obtain the greatest effect on the number of total infections.

Once the decision is made, the vaccines are shipped to the respective regions in accordance with the allocation scheme. After the vaccines arrive to a given region, the state takes it upon itself to organize the logistics of distributing the actual vaccines. Evidence from the H1N1 outbreak, particularly the analysis done by Hopkins [1], shows that far from all of the vaccines are actually used after they have been shipped. Moreover, once the vaccines are actually administered, there is a delay between the time of vaccination and when the vaccine actually takes effect by significantly lowering a person's susceptibility to infection.

3. Two sources of complexity

3.1 Limited observable information

In the US, the governing body like the CDC has access to very limited information. In other words, there is some information that is actually observable to the CDC; and there is other information that we can try to deduce from the observable information. We can then use this new, extrapolated information, to make a decision as to how to allocate vaccines to distinct regions. There are three main types of observable information.

1. Epidemic curve

Due to the inherent difficulties in tracking influenza spread, the only available information is the % ILI curve provided to the CDC by sentinel sites in individual states. This curve does not represent actual numbers of infections, so decision makers must form an approximation of a real available epidemic curve with the limited % ILI data available. Moreover, these data often contains inaccuracies due to initial underreporting of illness during the pandemic. Finally, this information is often delayed due to the time needed to actually gather reports by sentinel sites.

2. **Vaccine schedule**

Vaccines are developed through the CDC using an egg-based process which takes approximately 6 months to complete [2, 3]. The vaccine orders come in on a weekly basis and need to be shipped out to constituent regions [4]. While some forecasting can be done to predict the amount of vaccine that will become available in time, the schedule is not known exactly. Once the vaccines are shipped to the states, the CDC no longer regulates the logistics of actual vaccine administration. Some states experienced delays and general confusion. As much as 3 weeks delay in actual vaccination clinics from the time of shipments has been experienced by some states.

3. **Demand**

The CDC receives requests from different regions that give relay some information about the demand for vaccine experienced by each region. However, these requests might not actually represent real demand accurately. During the H1N1 pandemic many states continued to request vaccines well after they were no longer being used in large quantities. The CDC, however, still sent large shipments of vaccines to states like Mississippi for a while despite limited quantities actually being administered to patients.

3.2 Computational difficulties

The ultimate solution to this problem would be solved with stochastic dynamic programming as we are making decisions at multiple stages. However, the state and transition space in this problem is inherently too large to use dynamic programming because the problem is not Markovian. That is, any decision on day i, strongly and non-linearly affects what happens on all days $j > i$.

At a given decision point, we can try to analyze the marginal benefit of a single vaccine. That is **the number of infections that would be averted if a single vaccine were administered**. We can look at the marginal benefit as a function of two parameters, the time at which the vaccine actually takes effect, and how many vaccines have already been administered prior to the moment our test vaccine takes effect. We can get some intuition by looking at the marginal benefit of the i^{th} vaccine at different points in time. The marginal benefit of the i^{th} vaccine at time t is the difference between the number of infections that occur if we administered i and $i + 1$ vaccines at time t.

We ran a simple example in a community of 100,000 people. Figure 2 shows that the marginal benefit of a vaccine is actually not a "nice function". Note that in our motivational example in Figure 1, we calculated the marginal benefit of the first vaccine. We assumed that no vaccines were administered before or after the test vaccines. The equivalent of Figure 1 for the example in Figure 2 is the cross-section of the curve that corresponds to 0 vaccines being administered before the experiment. As a sanity check, these two curves are indeed similar. Now let us look at cross sections in the other dimension – time.

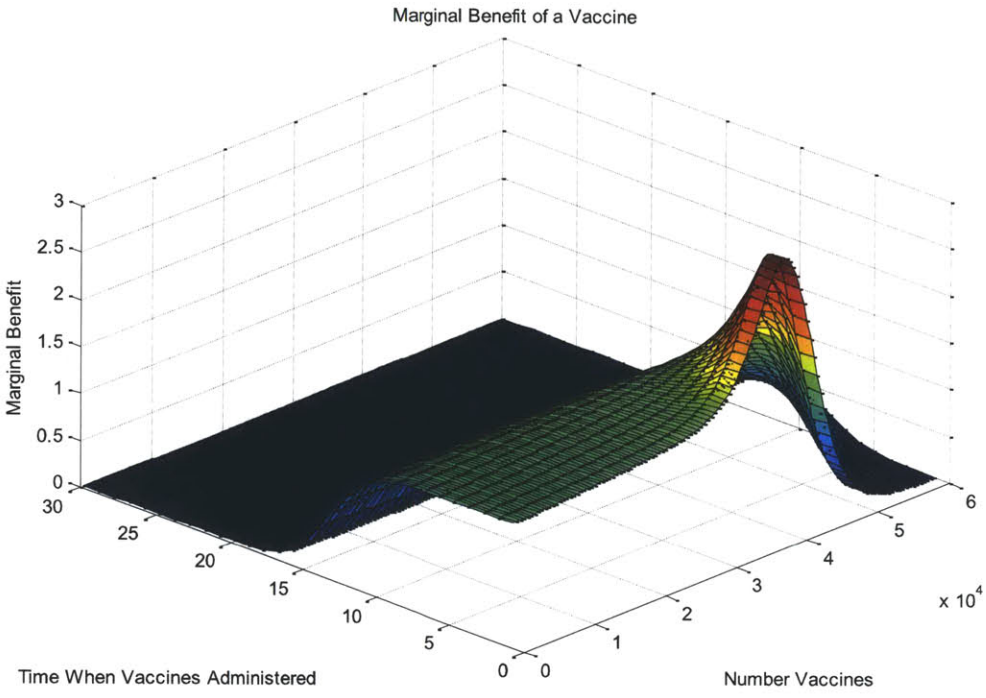


Figure 2: Marginal benefit of a single vaccine administered as a function of time and the number of vaccines that have already been administered in the community.

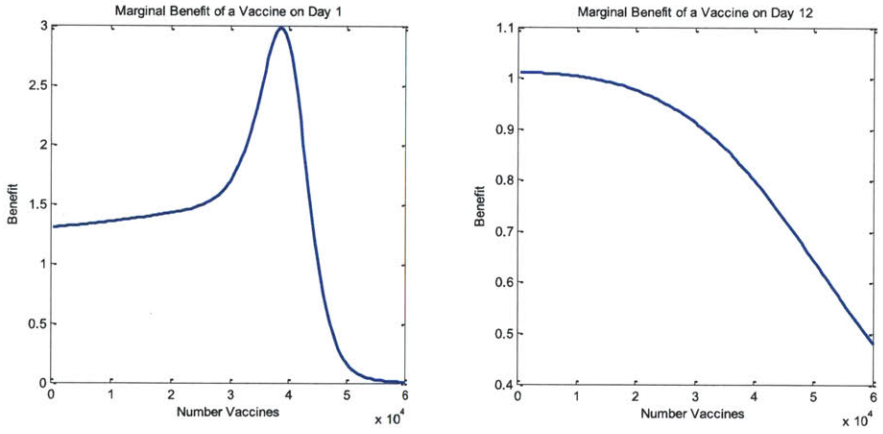


Figure 3: Marginal benefit functions of a single vaccination at different administered on Day 1 and Day 12, as a function of the total number of vaccines that have already been administered in the community.

Consider first what happens if we administer the vaccines on day 0 (Figure 3a). The marginal benefit of the vaccines with respect to the number of vaccines administered is convex at first, then turns concave, and finally convex again. The marginal benefit is highest when the number of vaccines administered is close to herd immunity for this community. In fact, the highest marginal benefit for a vaccine is reached by the vaccine that causes the population to attain herd immunity to the infection.

Now let us look at a later time, say day 12 (Figure 3b). At this point, the marginal benefit of a vaccine is monotonically decreasing. This occurs because herd immunity has already been reached through the natural course of the infection on day 12. Still, there are both convex and concave segments of this curve. The reason it would be too difficult to try to use optimization and approximate dynamic programming methods is not only because the marginal benefit function is complicated but also because this function changes based on decisions made in earlier stages.

Instead, we propose using one of several heuristic algorithms that should improve on the simple strategy of distributing vaccine on a per capita basis.

4. Problem formulation

We now formulate this problem mathematically. We say that we have R regions, each having population n_i. We also have a known vaccine schedule vector $\vec{v}$, so that $\vec{v}(k)$ is total number of available vaccines on day k. We must make an allocation decision on each of T decision days for which $\vec{v}(k) > 0$. As before, we are using the word day to mean generation of influenza.

A decision consists of R vectors $\overrightarrow{v_j}$ for $j = 1 \dots R$, such that $\sum_{j=1}^{n} \overrightarrow{v_j} = \vec{v}$. Each vector $\overrightarrow{v_j}$ represents the vaccines allocated to region j, and $\overrightarrow{v_j}(k)$ is the number of vaccines allocated to region j at time k. Let's denote $I(\overrightarrow{v_j})$ as the total number of infections in region j with the allocated vaccine schedule $\overrightarrow{v_j}$. Finally, let $\overrightarrow{e_k}$ be the vector such that $\overrightarrow{e_k}(m) = \begin{cases} 0 \; for \; m \neq k \\ 1 \; for \; m = k \end{cases}$.

For observed information, let v_i^* be the vector of all vaccines already shipped prior to decision day t, and let $E_i(t)$ be the number of infections on day t in region i, the observed value of the epidemic curve on day t.

Our goal is to make a decision $\overrightarrow{v_j}(t)$ for $j = 1 \dots R$ on each day t such that $\vec{v}(k) > 0$ in such a way that the sum of total infections over all regions, $\sum_{i=1}^{R} I(\vec{v})$ is minimized. In cases of particularly deadly epidemics, similar analysis may be performed with different objective functions such as the total number of deaths or the total economic impact of an epidemic.

Now that we have formalized the quantities needed for our decision making process we proceed with suggesting allocation heuristics to evaluate.

5. Proposed heuristics

5.1 Pro-rata heuristic

This is the current heuristic used by the CDC. Specifically,

$$v_i(t) = v(t) * \frac{n_i}{\sum_{i=1}^{R} n_i}$$

5.2 Pre-peak heuristic

A simple, "quick and dirty" way to increase the effectiveness of vaccines is to allocate all vaccines only to those states that have not yet reached the peak of infection. That is, to allocate to states for which the epidemic curve is currently increasing, distributing vaccines proportionally within that set of states:

$$v_i(t) = v(t) * \frac{n_i}{\sum_{k \in A} n_k} \; for \; i \in A$$

$$v_i(t) = 0 \; otherwise$$

$$\text{Where } i \in A \text{ iff } E_i(t) > E_i(t-1)$$

The problem with this heuristic is that it is too simplistic. States just past their epidemic curves may very well be benefited by some vaccine. Moreover, slight variations in the observed epidemic curve due to the stochastic nature of an epidemic would often cause states that have not yet reached the peak of the epidemic to lose allocated vaccines during the early stages of the outbreak when vaccines are most effective. Such mistakes can be very costly and are also bound to elicit strong disagreement from the states passed over for vaccine shipments.

5.3 Greedy heuristic

The most intuitively straight-forward approach to allocating vaccines is to allocate the vaccines one by one to the region with the highest marginal benefit for the next vaccine. To use the greedy heuristic on day t we ignore all information about any future vaccines, and make a decision based only on vaccines already allocated prior to day t and those available today. The following algorithm allocates the day t vaccines in a greedy manner.

Initialize with $v_i = v_i^*$ for $i = 1 \ldots R$;
Initialize a 1 x R vector A
for $i = 1$ *to* R
    for $j = 1$ *to* $v(t)$
        A = $I(\overrightarrow{v_i} + e_t)$
        let $i^* = argmin(A)$
        set $\overrightarrow{v_{i^*}} = \overrightarrow{v_{i^*}} + \overrightarrow{e_t}$
    end
end

5.3.1 Drawbacks of greedy algorithm

This algorithm allocates each available vaccine, one by one, to the region that will get the most marginal benefit from the single dose. This greedy algorithm performs better than the pro-rata heuristic, but the greedy property causes suboptimal solutions. Consider a two region problem, where region 1 is experiencing the peak of its epidemic early on, say on day 5, and the other is set to experience the outbreak much later, on day 25. Suppose, also, that only two shipments of vaccines are available; namely on day 5 and on day 20. Intuition tells us that we should give most of the first shipment to the first region and all of the second shipment to the second region. However, on day 5, the marginal benefit of most vaccines for the second region will be slightly higher than the marginal benefit for the first region, so the greedy algorithm will allocate the majority of the first shipment to the second region as well. The greedy algorithm does not take into consideration later interventions, and thus is likely to make some inefficient decisions early on.

5.4 Critical period

To illustrate the shortcoming of the greedy heuristic, consider Figure 4 where we superimpose the marginal benefit curves of two regions. The shape of the two curves is identical, but region 2 experiences the outbreak 3 weeks prior to region 1. Due to this, while the marginal benefit of a single vaccine is slightly higher in region 1, the derivative of the curve is also relatively high as compared to that of region

2. We can often afford to wait with region 1 and try to focus on region 2 while we are its highly time-sensitive period.

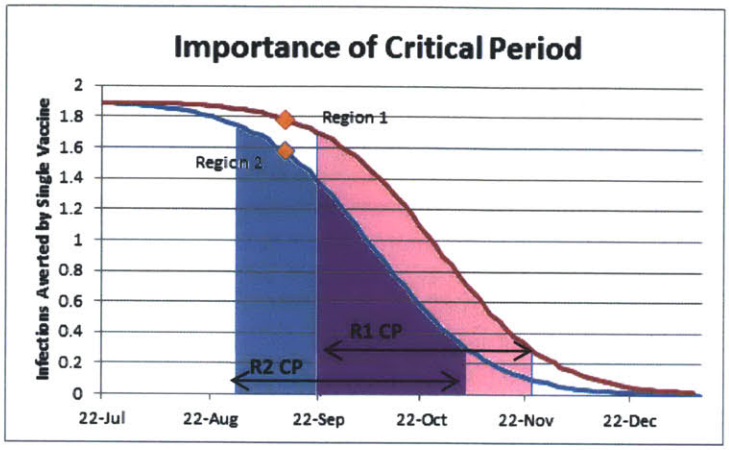


Figure 4: Critical periods of two regions based on the state of Indiana with epidemic starting at different times. The epidemic for Region 2 starts earlier than the epidemic in Region 1.

Note that the marginal benefit of a single vaccine in region i at time t is $MB_i(t) = I(\overrightarrow{v_i} + e_t) - I(\overrightarrow{v_i})$. We then say that a decision day t is in the Critical Period with threshold ε $(CP_{i,\varepsilon})$ if $-\frac{dMB_i(t)}{dt} > \varepsilon$. In our discrete model, this is equivalent to saying $MB_i(t) - MB_i(t+1) = I(\overrightarrow{v_i} + e_t) - I(\overrightarrow{v_i} + e_{t+1})$. Thus, for this heuristic we set some ε and allocate vaccines on day t as

$$v_i(t) = v(t) * \frac{n_i}{\sum_{k \in B} n_k} \; for \; i \in B$$

$$v_i(t) = 0 \; otherwise$$

where $i \in B$ iff $t \in CP_{i,\varepsilon}$

While this heuristic takes into account the future information available at time t, it is still somewhat simplistic in that it also follows the "all-or-nothing" approach. Some states are given significant preference over others even though the difference in the effect of a vaccine in those states might be fairly close. This is both hurtful to our objective function of minimizing the total number of infections and infeasible in a real-life situation as it introduces unreasonable inequality between regions. Instead, we propose an iterative algorithm which combines the benefits of both the critical period heuristic and the greedy heuristic by looking at the marginal benefit that can be attained at all possible decision points, including future ones.

5.5 Switching algorithm heuristic

To implement the switching heuristic algorithm on day t, we first need to make some assumption about the information that is available to decision makers on at this decision point. We assume that the following things are known or estimated:

- R_0 for all regions
- $E_i(k)$ for $i = 1 \ldots R, k \leq t$

- The function $I(\overrightarrow{v_i})$ for $i = 1 \dots R$
- The already allocated vaccine vector $\vec{v}_i^*$ for $i = 1 \dots R$
- And finally the forecasted vaccine schedule $\vec{v}$

Given this information, we make a decision for day t and estimate forecasted decisions for all subsequent decision days. That is, we calculate $v_i(t)$ and $v_i(k)$ for $k > t$. The algorithm follows the following logic. We start with a possible allocation (satisfying $v_i(t) > 0 \; \forall \, t, i$ and $\sum_{i=1}^{R} v_i = v$) for all future decision points. Then, for every decision point $k \geq t$, we calculate the cost of switching a batch of vaccines between every pair of regions. We then select the best switch over **all possible decision points**, and update the allocation accordingly. Note that this may happen in the future and not reflect any actual action on day t. We continue this until no more improvements in the objective function can be made.

By continually taking into account future decisions, we obtain a solution that maintains a balance between the myopic and "all-or-nothing" heuristics. Since our forecast into the future is bound to be flawed, however, we must repeat this procedure at each day $k > t$, continually updating the forecasted decisions for future time periods.

The formalized version of this heuristic is given below:

Initialize with $v_i(k) = v_i^*(k)$ for $k < t$ and $v_i(k) = v(k) * \frac{n_i}{\sum_{i=1}^{R} n_i}$ for $k \geq t$ for all i;
Initialize a R x R x $(T - t + 1)$ matrix A
while $\left(\exists \, i, j, k \mid I(v_i) + I(v_j) > I(\overrightarrow{v_i} + \overrightarrow{e_k}) + I(\overrightarrow{v_j} - \overrightarrow{e_k})\right)$
    for every triple (i, j, k) such that $\overrightarrow{v_j}(k) > 0$ and $k \geq t$ calculate
    $I(\overrightarrow{v_i} + \overrightarrow{e_k}) + I(\overrightarrow{v_j} - \overrightarrow{e_k})$

        Let $(i^*, j^*, k^*) = argmin(A)$
        $\overrightarrow{v_{i^*}} = \overrightarrow{v_{i^*}} + \overrightarrow{e_{k^*}}$
        $\overrightarrow{v_{j^*}} = \overrightarrow{v_{j^*}} - \overrightarrow{e_{k^*}}$
    end
end

In this algorithm, we start with some vaccine allocation and then move vaccines, one by one, from one region to another in manner that most effectively decreases the total number of infections in all regions. Since, during each iteration, the total number of infections decreases, this algorithm must terminate. However, it takes an intractably long time with complexity of $O(kVR^2)$.

We can make the algorithm run faster by trying to move higher numbers of vaccine at first. We also do not need to recalculate the matrix A for any regions except the ones that were changed on the previous iteration. The resulting algorithm is a modified gradient descent algorithm which converges to a near optimal solution quickly and efficiently. Please see Appendix A for implementation details and Appendix D for runtime details.

6. Testing strategy

Influenza progression in a community is made difficult by the fact that an epidemic curve rarely follows the epidemic curve generated in expectation by any deterministic flu spread model. In fact, day-to-day

progression is not only volatile but suffers from inaccuracy of available information. Patient reporting to doctors and hospitals is highly dependent on the public estimation of the risk associated with the current outbreak. To test the effectiveness of various allocation heuristics we must incorporate the maximal amount of uncertainty relevant to a possible outbreak. We test the heuristics by using Monte Carlo simulation to approximate the expected total number of infections that will occur with the governing body using a given heuristic. For each run of the simulation, we use a discrete time model, where for each day we perform the following four steps:

STEP 1: We first determine the starting time and the fundamental parameters of epidemic spread in each region. To put our results in context we used the data from the 2009-2010 H1N1 epidemic to infer these parameters. We used the H1N1 data provided by the 50 states to the CDC to fit value for R_0 for each region. We then used these estimated parameters as inputs to the stochastic version of the model submitted to EJOR to generate the next day of the model. See Appendix B and C for details of the stochastic implementation of the model and the summary parameters used. In Chapter 4, we estimated the value of R_0 from the real data of the H1N1 outbreak. We used these values with the benefit of hindsight and 2 years of data gathering and analysis. Naturally, this information was not available as the outbreak was happening, so in this case we assume that these parameters are hidden from decision makers, who experience only the epidemic curve. For testing purposes now, we assume that the epidemic curve information is observable to the CDC as it happens. However, more accurate modeling needs to take into account the delay between information available and real events.

Using the heterogeneous epidemic spread model from our EJOR paper, we represent the effect of a vaccine administration as a drastic reduction in susceptibility for a patient. While ideally we would hope that all allocated vaccines become available and administered immediately once they are shipped to their respective regions; the vaccine uptake percentage has been significantly lower than 100%. For the 11 states analyzed in the American Journal of Disaster Medicine paper, the percentage of all vaccines actually used by each state ranged from 20% to 40%. This number is lower than what we would expect, but it is not too discouraging if we understand that most of the unused vaccine was the vaccine that was shipped at the end of the vaccination program when most of the demand for vaccines has gone down. In the first few weeks of the program, the administered vaccine percentage came up to 60% in some states. For this testing framework we estimated the administered vaccine percentage for each region from the H1N1 data, and used this percentage to model a similar regional and public reaction to vaccine available. We assumed that this information was available to the CDC in their decision-making process as it can be easily estimated from historical data and using the information from the first few weeks of the epidemic. Appendix C contains the values of all parameters.

Once the next day is generated, if new vaccine becomes available to be shipped, we move on to testing the decision-making process.

STEP 2: The decision-maker must now estimate the model parameters to be used for the allocation algorithm. For our purposes this time we use the deterministic version of the EJOR model and fit the parameters to the available data generated in Step 1. Note that only the switching algorithm heuristic, the greedy heuristic, and the critical period heuristic require this fit, while the other heuristics require only the observed epidemic curve for the decision making process.

STEP 3: Armed with approximated parameters, the decision makers can now use one of the possible heuristics in Section 5 to output a set of values $v_i(t)$ which represent the amount of available vaccine that is to be distributed to each region.

STEP 4: Apply decision to each region by allocating the vaccines according to the decision in Step 3. This allocation will now be used in Step 1 when the next generation day is calculated.

Naturally, if there is no vaccine available, that is $v(t) = 0$, Steps 2 through 4 need not be performed, and we simply generate the next day in Step 1.

To make runtime for the testing module tractable we needed to make some small modifications to this process. Please see Appendix D for runtime details and the testing approximation. Note that during the actual decision making process in a real pandemic, we would only need to complete Steps 2 through 4 once for every available batch of vaccines. Thus, we can afford to run a fairly complex decision algorithm without any simplifications, as it is only needs to be executed approximately once every week.

7. Testing results

7.1 Recreating 2009-2010 H1N1 epidemic

We ran our Monte Carlo simulation using the parameters fit to the 2009-2010 H1N1 epidemic and other the characteristics of the 10 regions monitored by the CDC. As in 2009, most of the regions started experiencing signs of a massive epidemic in the summer, while the vaccines started arriving on October 10th. From that point on, decisions to ship the available vaccines were made on a weekly basis. Figure 5a contains the estimated epidemic curves for the 10 regions of the United States. The epidemic curves were obtained from the %ILI numbers reported by sentinel sites in the same manner as the Chapter 4. We also include the vaccines as they were made available to be shipped by the CDC in Figure 5b.

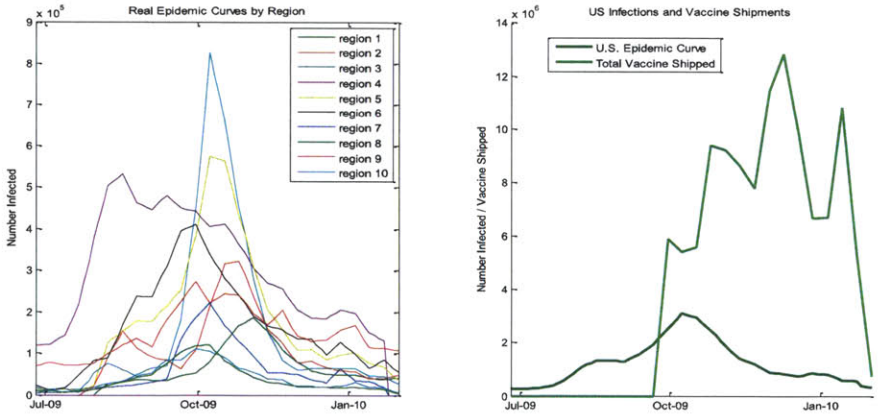


Figure 5: Figure 5a includes the real estimated epidemic curves by region from the 2009-2010 H1N1 epidemic. Figure 5b includes the vaccine shipped and the total epidemic curves for the US.

Note that the number of vaccines developed and shipped by the CDC is almost twice as large as the total number of infections. This is partly due to the less-transmissible-than-expected nature of the 2009 H1N1 virus and partly due to a knee-jerk reaction by the CDC to vaccinate as many people as possible. However, the vaccines started being shipped just before the peak of the infection over the country soon

becoming relatively ineffective and largely unused. Figure 6, in turn, contains the model-generated epidemic curves of the pandemic as well as the model-generated cumulative curve of the epidemic as it would have occurred **without** any available vaccines.

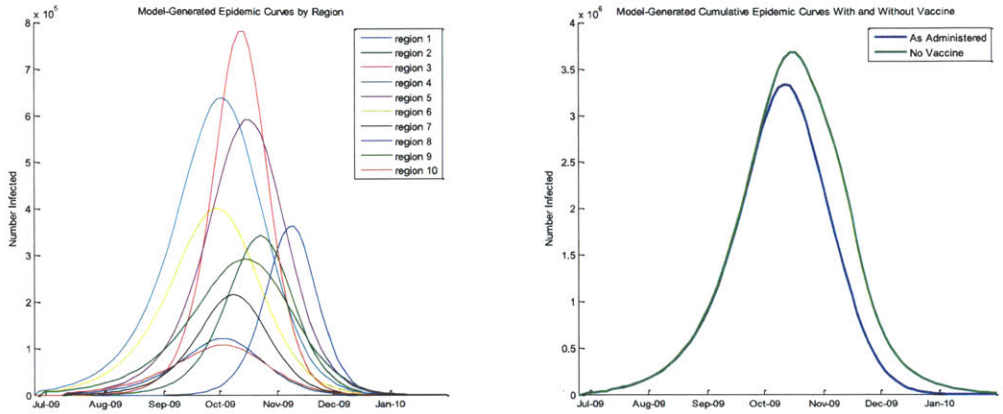


Figure 6: Figure 6a includes the model-generated epidemic curves by region. Figure 6b contains the US combined epidemic curve and the model-generated curve of what would have occurred if no vaccine were administered.

	Total Infections (Vaccines As Administered)	**Total Infections (No Vaccine)**
Region 1	4,464,000	6,988,000
Region 2	5,454,000	6,821,000
Region 3	10,311,000	11,627,000
Region 4	13,414,000	15,218,000
Region 5	11,116,000	14,464,000
Region 6	8,577,000	9,488,000
Region 7	3,659,000	4,238,000
Region 8	2,414,000	2,802,000
Region 9	7,002,000	9,866,000
Region 10	2,385,000	2,878,000
Total	**68,795,000**	**84,391,000**

Table 1: Model-generated total infection numbers by region with and without administered vaccines.

The vaccines made available by the CDC have certainly made a difference for the total number of infections in the US. The vaccine program was particularly effective in Region 1 with 36% improvement in the total number of infections and virtually ineffective in Region 6 with a 9% improvement. Let us look closer, in Figure 7, at the epidemic curves and allocation strategies if we use four of the heuristics from Section 5: pro-rata, pre-peak, critical period, and switching heuristics. Table 2 summarizes the total number of infections under each heuristic.

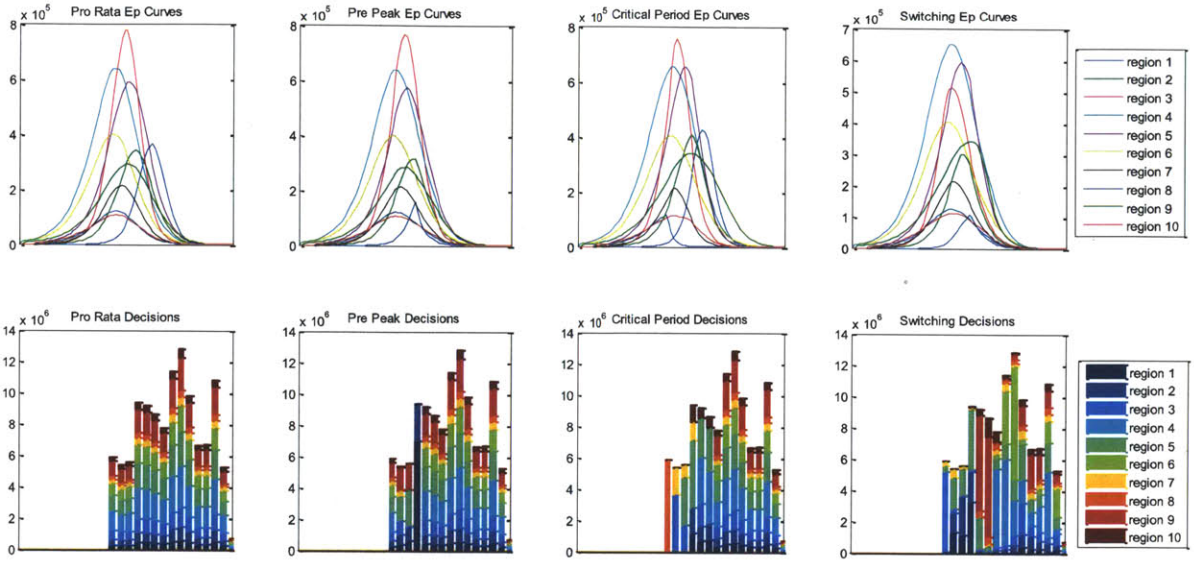


Figure 7: Epidemic curves by region and decisions made by using the pro-rata, pre-peak, critical period, and switching heuristics.

Strategy	Total Infections	95% CI	Difference in Infections Averted (%)
No Vaccine	84,391,000	(84,388,000, 84,394,000)	---
Pro-rata	68,795,000	(68,783,000, 68,807,000)	---
Pre-peak	65,798,000	(65,773,000, 65,822,000)	19.22%
Critical Period	67,536,000	(67,458,000, 67,613,000)	8.07%
Switching Algorithm	64,012,000	(63,999,000, 64,026,000)	30.67%

Table 2: The cumulative number of infections using the pro-rata, pre-peak, critical period, and switching algorithm using the parameters fit from the 2009-2010 H1N1 epidemic.

The three heuristics suggested in Section 5 performed better than the pro-rata heuristic in mitigating the total effects of the outbreak. The critical period heuristic is the least effective of the three due to the fact that this is an all-or-nothing heuristic and it depends very strongly on the correctness of the approximation algorithm that determines whether a region has entered its critical period. With the limited information available during the first few decision points, this heuristic often made entirely incorrect decisions. The best performing heuristic was the switching algorithm, saving roughly 20 million infections compared to the pro-rata heuristics' 15.5 million. This is a 31% improvement in averted infections over the currently used allocation algorithm. A closer look at the two heuristics is contained in Figure 9.

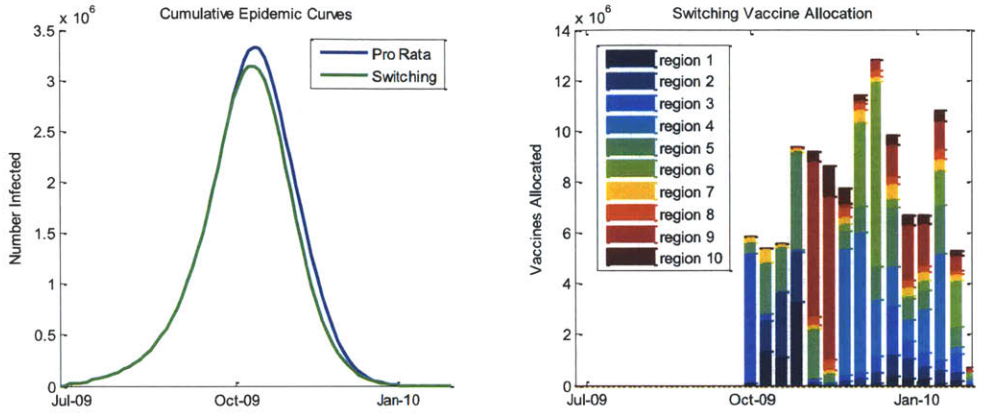


Figure 8: Cumulative epidemic curve for the pro-rata and switching heuristic and the decisions made through the switching heuristic.

The pre-peak heuristic, in which we gave preference to those states that were deemed to have not yet reached the peak of the epidemic also performed well, with a 19% improvement over the Pro Rata heuristic. However, this all-or-nothing approach is difficult to justify and implement with uncertain data in an actual real-life scenario, especially with the potential pitfalls evident with the critical period heuristic. The switching algorithm allows for a more gentle distribution of vaccines while maintaining high effectiveness.

7.2 Improving results

7.2.1 Earlier vaccine shipment

We must keep in mind that the numbers in Table 2 are a result of the modeling based on the 2009-2010 H1N1 scenario, in which the vaccines arrived very late and most of the states received their first vaccines after the peak of the outbreak has already occurred. It is possible [5], that when the next deadly outbreak occurs, it will come to the US from overseas, Asia or even Africa, and the CDC will have a little more time to develop and distribute influenza vaccines. A switch from the egg-based vaccine technology currently employed to the animal-cell research used in some European countries would ensure an even bigger preparation time to ship vaccines. The new technology would produce vaccines about 8 to 10 weeks faster than the egg-based methods [6, 7]. To examine the relative effect of timing on vaccine effectiveness we consider what would happen if the vaccines became available two or even four weeks earlier than in the case of 2009, in early September.

Strategy (2 weeks earlier)	Total Infections	95% CI	Difference in Infections Averted (%)
No Vaccine	84,391,000	(84,388,000, 84,394,000)	---
Pro-Rata	57,349,000	(57,333,000, 57,364,000)	---
Pre-Peak	54,817,000	(54,809,000, 54,824,000)	9.34%
Critical Period	55,748,000	(55,724,000, 55,771,000)	5.9%
Switching Algorithm	52,628,000	(52,611,000 52,644,000)	17.44%

Table 3: The cumulative number of infections using the pro-rata, pre-peak, critical period, and switching algorithm if vaccines were shipped 2 weeks earlier than in 2009.

Strategy (4 weeks earlier)	Total Infections	95% CI	Difference in Infections Averted (%)
No Vaccine	84,391,000	(84,388,000, 84,394,000)	---
Pro-Rata	43,379,000	(43,364,000, 43,394,000)	---
Pre-Peak	41,489,000	(41,482,000, 41,497,000)	4.61%
Critical Period	42,560,000	(42,533,000, 42,587,000)	2%
Switching Algorithm	39,776,000	(39,760,000, 39,791,000)	8.79%

Table 4: The cumulative number of infections using the pro-rata, pre-peak, critical period, and switching algorithm if vaccines were shipped 4 weeks earlier than in 2009.

With earlier vaccines, the number of infections using the pro Rata heuristic dropped 11 million with a two week headway and an entire 25 million people with a four week lead. In the case of the switching heuristic the respective increases in the number of infections averted was also 11 and 25 million. The extra weeks given to prepare significantly increase the effectiveness of available vaccines, but with earlier administration the difference between pro-rata and switching heuristics becomes less pronounced as timing becomes less critical. To generalize this notion, we do a similar analysis with vaccines becoming available eight weeks before, with the graphical results in Figure 9. Figures 10a, 10b, and 10c show the general trend of vaccine effectiveness at different shipping times.

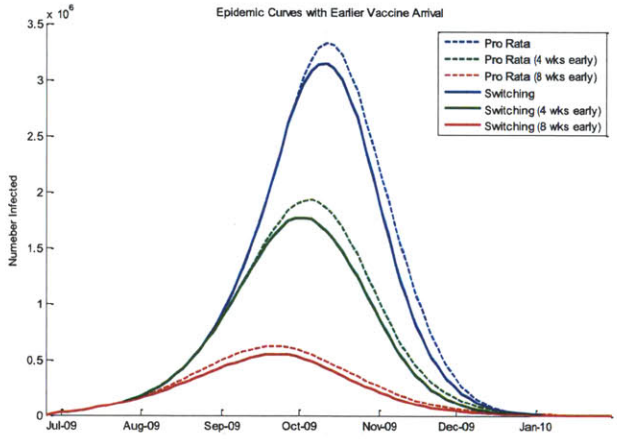


Figure 9: Cumulative epidemic curves of the pro-rata and switching heuristics with vaccines arriving at 0, 4, and 8 weeks earlier than in 2009.

As we see from Figure 10b, the absolute difference between the number of averted infections with the switching algorithm and the pro-rata heuristic is not monotonic, the relative difference given in percentages in Figure 10c is highest in the most realistic situation, that is when vaccines arrive relatively late as they had in 2009. It is at that point that exploiting the real-time dynamically evolving differences between regions becomes most important and effective, providing as much as a 25% improvement in lowering the toll of an outbreak.

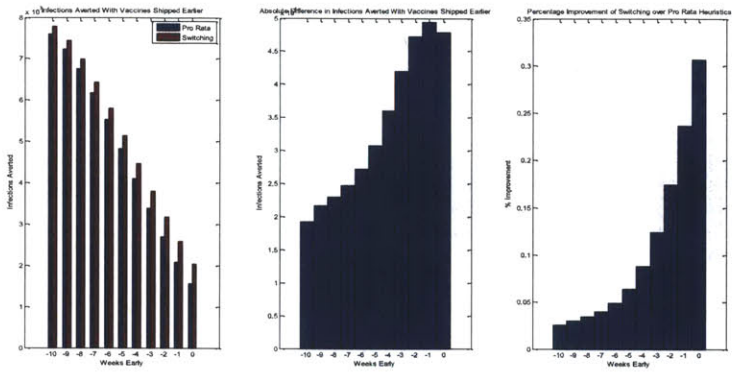


Figure 10: Figure 10a contains the total numbers of infections averted as a function how early the vaccines were shipped. Figure 10b contains the difference in the total number of infections averted by the pro rata and the switching heuristic. Figure 10c contains the percentage difference in the total number of infections averted by the two heuristics.

Finally, we would like to discuss a point that is unlikely to occur in a real-life scenario but needed to be addressed in the analysis needed for Figure 10. When vaccines arrive particularly early, that is the vaccines take on a more preventative role, we might have to make a decision before having reasonable information about the epidemic curve in some regions. In particular, in the hypothetical case, in which vaccines arrived 10 weeks earlier they had, none of the regions have experienced any significant epidemiological activity. In reality, it is unlikely that we will have to make decisions before at least some information is available and we can make some estimates of R_0. However, when faced with this situation, the switching algorithm tends to allocate very little vaccine to a region with little progression. When estimating parameters for a late-start region with minimal number of infections, we estimated an approximate value for R_0 from other regions and assigned it to the late-start region with the assumption that the first infection will occur 2 weeks after the current decision time. The resulting allocation favored the regions where the outbreak was more prevalent.

7.2.2 Finer region segmentation

The CDC currently uses 10 regions to separate the United States into manageable sections and to allocate vaccines track influenza spread. After allocating vaccines using any of our heuristics to one of such regions, the vaccines would be spread in some manner to the constituent states within each region. That can happen either by allocating the vaccines according to population (pro-rata), or to once again use another heuristic to try to maximize the effectiveness within the area. For a country with a population of hundreds of millions of people 10 regions make up a very general geographical split. A finer split would theoretically offer us more degrees of freedom and thus more opportunities to explore the difference in the epidemic's behavior in different states. We ran our simulation by splitting the United States into 20 regions (see Appendix E for the geographical view) according to data from the H1N1 outbreak. The resulting epidemic curves are recorded in Table 5.

Strategy	Total Infections	95% CI	Difference in Infections Averted (%)
No Vaccine	84,871,000	(84,865,000, 84,877,000)	---
Pro-Rata	69,011,000	(68,991,000, 69,031,000)	---
Pre-Peak	66,606,000	(66,576,000, 66,636,000)	15.16%
Critical Period	66,320,000	(66,267,000, 66,377,000)	16.97%
Switching Algorithm	63,528,000	(63,594,000, 55,560,000)	34.24%

Table 5: The cumulative number of infections using the pro-rata, pre-peak, critical period, and switching algorithm if the US were split into 20 regions.

The simulation results are slightly counter-intuitive. The effectiveness of the pro-rata heuristic essentially stayed the same, while the switching algorithm heuristic made minimal improvement in the base case. Due to this, the additional complications associated with splitting the region into 20 regions is not necessarily justified.

7.2.2. Higher vaccine uptake

Overall, the amount of vaccine actually administered to patients, especially toward the end of the H1N1 pandemic, was fairly low. The total percentage was never higher than 60% of available vaccine, even in the first few days of vaccine shipment. As a last example, we provide in Table 6 the simulation results of what would happen if the total percentage of vaccine used in each region were increased by a factor of 0.2.

Strategy	Total Infections	95% CI	Difference in Infections Averted (%)
No Vaccine	84,391,000	(84,388,000, 84,394,000)	---
Pro-Rata	66,428,000	(66,400,000, 66,456,000)	---
Pre-Peak	64,085,000	(64,051,000, 64,119,000)	13.04%
Critical Period	65,802,000	(65,589,000, 66,016,000)	3.48%
Switching Algorithm	61,793,000	(61,762,000, 61,824,000)	25.80%

Table 6: The cumulative number of infections using the pro-rata, pre-peak, critical period, and switching algorithm if vaccines if were administered 20% more often.

The results encourage the use of public awareness campaigns to increase the total number of vaccines actually accepted by the public. A 20% increase in the acceptance of vaccines can avert anywhere from 2.3 to 7 million extra infections.

8. Incorporating complexity and constraints

While the switching allocation algorithm produces better results than the pro-rata heuristic, allocation that is not based on the population of each region raises the possibility of favoring certain regions significantly over others. In particular, in our testing example from Figure 8, Region 4 received enough vaccine to cover 26% of the population -- proportionally less than half of the amount received by Region 1. We can visualize the differences in vaccine effectiveness on each region by analyzing the effects of the pro-rata and switching heuristic on a regional scale. In Appendix F, we include a more detailed version of Table 2,the total number of infections for each heuristic is included for each of the ten US regions. We would like to know which regions experienced a positive effect from the switching heuristic, and which regions became worse off as a result. From Table 7, we can tell that half of the regions would actually see a net increase in the total number of infections. The possible negative effect of the switching heuristic results in

at most an infection increase of 12%, as is the case in region 10. The highest possible benefit, as in region 1, is over 76%.

Region	Number Infections Averted	Percentage Infections Averted from Pro-Rata*
1	3,393,000	76.01
2	1,323,000	24.25
3	2,796,000	27.12
4	-1,515,000	-11.29
5	811,000	7.29
6	-798,000	-9.3
7	4,000	0.12
8	-279,000	-11.57
9	-650,000	-9.28
10	-302,000	-12.71

Table 7: The region by region summary of infections averted by using the switching heuristic over the pro-rata heuristic. A positive number implies an *increase* in the total number of infections when the switching heuristic is used. *This column is calculated by taking the difference between the number of infections using the two heuristics, and expressing that as a percentage of infections that would have occurred had the pro-rata heuristic been used.

8.1 Fairness constraints

While the switching constraint's allocation might be mathematically justified by a net decrease in the total number of infections, a policy-setting body like the CDC might feel that it is inappropriate to allocate with such a tangible preference of one region over another. We can easily include this consideration in the Switching Algorithm heuristic by turning our problem into a constrained one. We impose the constraint that we cannot take away more vaccines from a region than is regulated by certain fairness limitations.

8.1.1 Hard constraints

One simple way of incorporating fairness is by mandating that on each day, each region i must receive no less than $l_i(t)$ vaccines. $l_i(t)$ may be an absolute number or a percentage of total available vaccines on day t. In the extreme case of the pro-rata heuristic, $l_i(t) = \frac{n_i}{\sum_{j=1}^{R} n_i} * v(t)$. The switching algorithm is then modified as follows:

Initialize with $v_i(k) = v_i^*(k)$ for $k < t$ and $v_i(k) = l_i(t) + (v(k) - \sum_{j=1}^{R} l_j(k)) * \frac{n_i}{\sum_{i=1}^{R} n_i}$ for $k \geq t$ for all i;

Initialize a R x R x $(T - t + 1)$ matrix A

while $\left(\exists\, i,j,k \mid I(v_i) + I(v_j) > I(\overrightarrow{v_i} + \overrightarrow{e_k}) + I(\overrightarrow{v_j} - \overrightarrow{e_k})\right)$

for every triple (i,j,k) such that $\overrightarrow{v_j}(k) > l_j(k)$ and $k \geq t$ calculate $I(\overrightarrow{v_i} + \overrightarrow{e_k}) + I(\overrightarrow{v_j} - \overrightarrow{e_k})$

Let $(i^*, j^*, k^*) = argmin(A)$

$\overrightarrow{v_{i^*}} = \overrightarrow{v_{i^*}} + \overrightarrow{e_{k^*}}$

$\overrightarrow{v_{j^*}} = \overrightarrow{v_{j^*}} - \overrightarrow{e_{k^*}}$

end

end

8.1.2 Relative constraints

Instead of imposing hard constraints, we might simply state that one region may not have any significant preference over another. That is, we can impose a lower bound parameter $\rho < 1$, such that for any two regions i and j, $\frac{v_i(t)}{n_i} \geq \rho \frac{v_j(t)}{n_j}$. This results in the following algorithm:

Initialize with $v_i(k) = v_i^*(k)$ for $k < t$ and $v_i(k) = v(k) * \frac{n_i}{\sum_{i=1}^{R} n_i}$ for $k \geq t$ for all i;
Initialize a R x R x $(T - t + 1)$ matrix A
while $\left(\exists\, i,j,k \mid I(v_i) + I(v_j) > I(\overrightarrow{v_i} + \overrightarrow{e_k}) + I(\overrightarrow{v_j} - \overrightarrow{e_k})\right)$
    for every triple (i,j,k) such that $\frac{\overrightarrow{v_j}(k)}{n_j} > \min_l \rho \frac{v_l(t)}{n_l}$ and $k \geq t$ calculate
    $I(\overrightarrow{v_i} + \overrightarrow{e_k}) + I(\overrightarrow{v_j} - \overrightarrow{e_k})$

    Let $(i^*, j^*, k^*) = argmin(A)$
        $\overrightarrow{v_{i^*}} = \overrightarrow{v_{i^*}} + \overrightarrow{e_{k^*}}$
        $\overrightarrow{v_{j^*}} = \overrightarrow{v_{j^*}} - \overrightarrow{e_{k^*}}$
    end
end

The key to understanding how to impose these constraints is to recall that states where vaccines are inherently less effective, that is, states where the epidemic is in a late stage are also the states where the demand for vaccines is relatively low. In those states most of the vaccines allocated in 2009 have gone unused. Those states that received their vaccines especially late, more than 5 weeks after the peak of infection, used only about 15% of the vaccines allotted to them. Those states, coincidentally were the states to which the switching algorithm allocated the least numbers of vaccines. Constraints based on demand forecasts will be most effective in maintaining fairness while increasing the effect of limited available vaccines.

8.2 Risk parameter

Imposing some kind of fairness constraints on the vaccine allocation problem lands us somewhere on the spectrum between the unrestricted solution obtained by the unconstrained switching algorithm and the pro-rata heuristic obtained by the CDC's actions in 2009. The place on the spectrum can be determined by the policy-makers as they tailor their strategy to their constituency. We propose introducing a riskiness parameter σ that determines how risky (switching) or safe (pro-rata) to make our allocation strategy.

Given an allocation decision to be made on day t, if $v_i(t)$ is the decision made by the pro-rata heuristic in region i on day t, and $w_i(t)$ is the corresponding decision made using the switching algorithm we allocate $(1 - \sigma)v_i(t) + \sigma w_i(t)$ to region i. The corresponding allocation lies on the spectrum of risky and relatively safe distributions. See Figure 11 for the effectiveness of the vaccines as a function of σ.

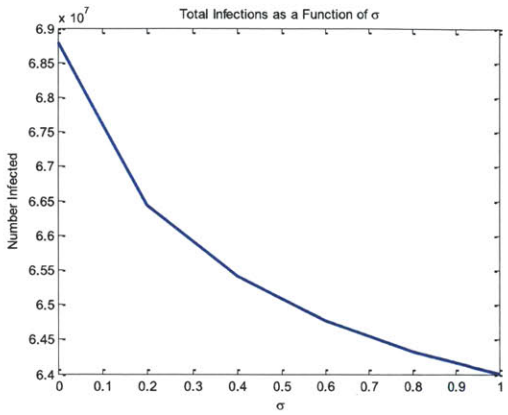


Figure 11: The total number of infections using the risk parameter combined with the pro-rata and switching heuristics.

Note that the number of infections occurring during an epidemic is convex in σ. That is, the benefits of a riskier switching algorithm heuristic can be reaped with a fairly low σ value. Even a little deviation from the safest Pro Rata heuristic provides success in averting infections overall.

The benefit of using a risk parameter is avoiding the pitfalls of incomplete information available in real time. A bad fit to data may cause a drastic decision to send a majority of the stockpiles to one region that might in reality not be justified. A low value of σ will safeguard against such pitfalls but will likely result in a lower number of averted infections. To make σ more adaptive, we can make it dependent on our confidence in the available data. One possible way of doing this is to create a parameter $\sigma(t)$ that is dependent on the quality of the least squares fit to our data. We ran the simulation using $\sigma(t) = K\frac{1}{\gamma(t)}$ where $\gamma(t)$ square norm of our fit to the existing data and $K$is a normalization constant.

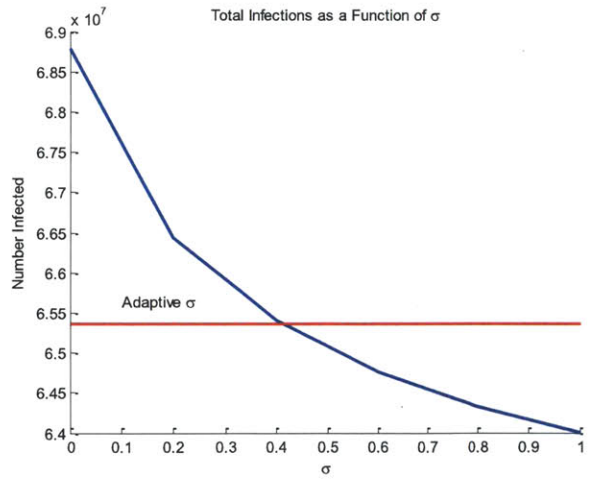


Figure 12: The total number of infections for various lines. The red line corresponds to the threshold number of infections attained by using the adaptive risk parameter.

The adaptive risk parameter performs quite well and the number of infections lies closer to those of the switching algorithm than the pro-rata heuristic. At the same time, we can see from Figure 13 that the

algorithm that uses the adaptive risk parameter allocates vaccines similarly to the Pro Rata heuristic at the beginning of the epidemic when data is still uncertain. This is changed as time goes on and the fit to data becomes more certain.

We can look also at the regional effect of the risk parameter on the total number of infections. As in Table 7, we consider the effects of the adaptive risk parameter algorithm compared to the pro-rata heuristic in Table 8. While 6 of the regions still experience a net loss from this algorithm, the extent of the largest possible loss is halved to 6% from 12%.

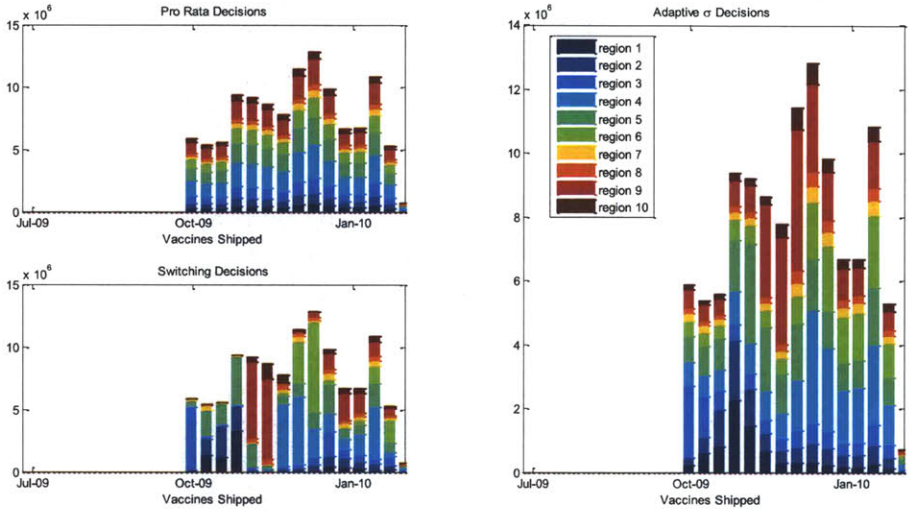


Figure 13: Decisions made by the pro-rata, switching, and adaptive risk parameters heuristics.

Region	Number Infections Averted	Percentage Infections Averted from Pro-Rata*
1	2,761,000	61.85
2	766,000	14.03
3	1,559,000	15.11
4	-623,000	-4.64
5	62,000	0.56
6	-314,000	-3.66
7	-66,000	-1.8
8	-128,000	-5.33
9	-439,000	-6.28
10	-143,000	-6.01

Table 8: The region by region summary of infections averted by using the adaptive risk parameter heuristic over the pro-rata heuristic. A positive number implies an *increase* in the total number of infections when the switching heuristic is used. *This column is calculated by taking the difference between the number of infections using the two heuristics, and expressing that as a percentage of infections that would have occurred had the pro-rata heuristic been used.

The risk parameter can be made more complicated still by adding our forecast of vaccine uptake percentages and other uncertainty parameters; it provides a balance to prevent drastic decisions from being made. We won't go into more detail here, but other adjustments to the algorithm may be made

when utilized in real time. For one, one can always use a different objective function to focus on lowering mortality or economic ramifications of a pandemic.

8.3 A few notes on parameter estimation

We would like to add a few final notes on parameter estimation used prior to making an allocation decision. Real time model fitting is very difficult to do with limited data, especially in the very beginning of an outbreak, when the data is still limited and uncertain. However, the most important factor in the switching algorithm heuristic is where on the epidemic curve the region is currently located. That is, whether the epidemic is just beginning so that exponential growth is still occurring, or whether it is close to the peak of infection when the incidence of new infections is slowing down. The switching heuristic compares this information between all regions and makes the allocation decision accordingly. Because of this, even though R_0 is important for estimating the effectiveness of vaccines, even rough estimates of R_0 and other parameters along with the aggregate data on the total number of infections produce very good results.

At the beginning of the outbreak, regions may estimate their modeling parameters such as R_0 from those values estimated in other regions of the world. For example, estimates of R_0 ranging from 1.4 to 1.6 in Mexico for the H1N1 pandemic surfaced as early as May 2009 [8], well before the US was experiencing major outbreaks. Even though the estimates were later adjusted, they were still within the same ballpark as the early findings. Even overall estimates gathered from modeling the United States as a single community and using aggregate data may yield good early estimates. Our estimates for the US regions ranged from 1.14 to 1.43, while the estimate that we would have achieved using the aggregate data for the US as a whole would have been 1.14. In the early stages, when data is not yet conclusive, such heuristics result in a less complicated parameter estimation process. Also, each region can provide their own estimation as the regional policy-makers are familiar with the interventions and actions that have been taken internally.

9. Conclusions

The problem of allocating limited numbers of vaccines to multiple regions is not only a mathematical one. This problem needs to be considered with the full weight of the requirements placed upon decision-makers in case of a real influenza emergency. Organizations that distribute vaccine, such as the CDC, are faced with uncertain data and epidemiological as well as political considerations that must be made every time there is vaccine to be distributed. Each decision lies on a spectrum, where riskier decisions –ones that disproportionally favor certain regions over others – result in lower infections than safer decisions such as the pro-rata heuristic currently used by the CDC in the eventuality of a worldwide influenza pandemic. We provide in this chapter the methodology that can help gain insights in the effectiveness of vaccines at different times in the outbreak, and to evaluate how well various decision-making strategies perform in reducing the total number of infections. This methodology is a tool in the hands of the decision-maker, who must choose just how risk-averse the allocation method should be. With that in mind, a combination of the switching heuristic, the pro-rata heuristic, and perhaps other allocation schemes can provide the most cost-effective and politically viable strategy for mitigating the toll of influenza.

References

1. Hopkins J. H1N1 After-action reports: lessons on vaccine distribution. *MIT ESD Working Paper*. 2011.
2. Centers for Disease Control and Prevention. Selecting the Viruses in the Seasonal Influenza (Flu) Vaccine. Available at: http://www.cdc.gov/flu/professionals/vaccination/virusqa.htm#manufacture. Accessed January 6, 2011.
3. Centers for Disease Control and Prevention. The 2009 H1N1 Pandemic: Summary Highlights, April 2009-April 2010. June 16, 2010. Available at: http://www.cdc.gov/h1n1flu/cdcresponse.htm. Accessed January 12, 2011.
4. Centers for Disease Control and Prevention. 2009 H1N1 Vaccine Doses Allocated, Ordered, and Shipped by Project Area. February 2, 2010. Available at: http://www.cdc.gov/h1n1flu/vaccination/vaccinesupply.htm. Accessed January 20, 2011.
5. Diamond J, Wolfe N. Where will the next pandemic emerge? *Discover*. 2008.
6. Rappuoli R. Cell-Culture-Based vaccine production: technological options. *The Bridge*. 2006;36(3):25-30.
7. Maugh T. Cell-culture influenza vaccine proves effective, could speed production. *Los Angeles Times*. February 2011.
8. Fraser C, Donnelly CCS, Hanage W, et al. Pandemic potential of a strain of influenza A (H1N1): early findings. *Science*. 2009;1557-1561:324.
9. Box G, Hunger J, Hunter W. *Statistics for Experimenters: Design, Innovation, and Discovery*. 2 ed: Wiley; 1978.

Appendix A: Efficient switching algorithm

Initialize with $v_i(k) = v_i^*(k)$ for $k < t$ and $v_i(k) = v(k) * \frac{n_i}{\sum_{i=1}^{R} n_i}$ for $k \geq t$ for all i;
Initialize a R x R x $(T - t + 1)$ matrix A
Let $\delta_v = 2^{\lfloor \log(\max_{i,k} v_i(k)) \rfloor}$
$\delta_c = 2^{\lfloor \log(I(v^*)) \rfloor}$
while $\delta_c \geq 1$
    $\delta_v = 2^{\lfloor \log(\max_{i,k} v_i(k)) \rfloor}$
    while $\delta_v \geq \delta_c$
        while $\left(\exists\, i,j,k \mid I(\overrightarrow{v_i}) + I(\overrightarrow{v_j}) > I(\overrightarrow{v_i} + \delta_v * \overrightarrow{e_k}) + I(\overrightarrow{v_j} - \delta_v * \overrightarrow{e_k})\right)$
            for every triple (i,j,k) such that $\overrightarrow{v_j}(k) > \delta_v$
                calculate $A(i,j,k) = I(\overrightarrow{v_i} + \delta * \overrightarrow{e_k}) + I(\overrightarrow{v_j} - \delta * \overrightarrow{e_k})$
                let $(i^*, j^*, k^*) = argmin(A)$
                if $A(i^*, j^*, k^*) > \delta_c$
                    $\overrightarrow{v_{i^*}} = \overrightarrow{v_{i^*}} + \delta_v * \overrightarrow{e_{k^*}}$
                    $\overrightarrow{v_{j^*}} = \overrightarrow{v_{j^*}} - \delta_v * \overrightarrow{e_{k^*}}$
                else
                    if $A(i^*, j^*, k^*) > 0$
                        $\overrightarrow{v_{i^*}} = \overrightarrow{v_{i^*}} + \delta_v * \overrightarrow{e_{k^*}}$
                        $\overrightarrow{v_{j^*}} = \overrightarrow{v_{j^*}} - \delta_v * \overrightarrow{e_{k^*}}$
                    end
                    $\delta_v = \frac{\delta_v}{2}$
                end
            end
        end
    end
    $\delta_c = \frac{\delta_c}{2}$
end

This algorithm is a modified version of the gradient descent algorithm due to the inherently discrete nature of the problem. We start with moving large numbers of vaccines but only when the cost improvement is significantly large. If moving large numbers of vaccines is not possible we lower the cost and vaccine thresholds until no vaccines can be moved to any improvement at all.

Appendix B: Stochastic spread model

We used the deterministic heterogeneous community influenza spread model as a basis for the stochastic model. The model had two levels of heterogeneity, λ and s. At the onset of the epidemic, λ was distributed as a Poisson distribution with mean μ over the entire population. μ was found by using a fit to the data from the 2009-2010 H1N1 epidemic. s corresponded to an individual's susceptibility to infection given contact with an infectious person, and was set to have two values, one for a non-vaccinated person and one for a person that has been vaccinated against the current influenza strain. We modify the deterministic model to incorporate the stochastic nature of a real outbreak during our Monte Carlo Simulations.

Recall from the deterministic heterogeneous community influenza spread model, that for each day i, we calculate the quantities

$$N_{i+1}^I = N_i^S \int_0^1 \int_0^\infty p(I|\lambda,s) * f_i^S(\lambda,s) d\lambda\, ds$$

$$N_{i+1}^S = N_i^S - N_{i+1}^I$$

$$f_{i+1}^I(\lambda,s) = \frac{p_i(I|\lambda,s) * f_i^S(\lambda,s)}{\int_0^1 \int_0^\infty p(I|\lambda,q,s) * f_i^S(\lambda,q,r,s) d\lambda\, ds}$$

$$f_{i+1}^S(\lambda,s) = \frac{\left(1 - p_i(I|\lambda,s)\right) * f_i^S(\lambda,s)}{\int_0^1 \int_0^\infty \left(1 - p(I|\lambda,s)\right) * f_i^S(\lambda,s) d\lambda\, ds}$$

When calculating the same quantities for a stochastic model, we follow similar steps. The probability that a given interaction is with an infectious person remains as before,

$$\beta_i = \frac{N_i^I E_i^I(\lambda)}{NE(\lambda)}$$

Similarly, the probability of becoming infected given an infectivity level λ,

$$p_i(I|\lambda,s) = 1 - e^{-\lambda s \beta_i}$$

remains the same.

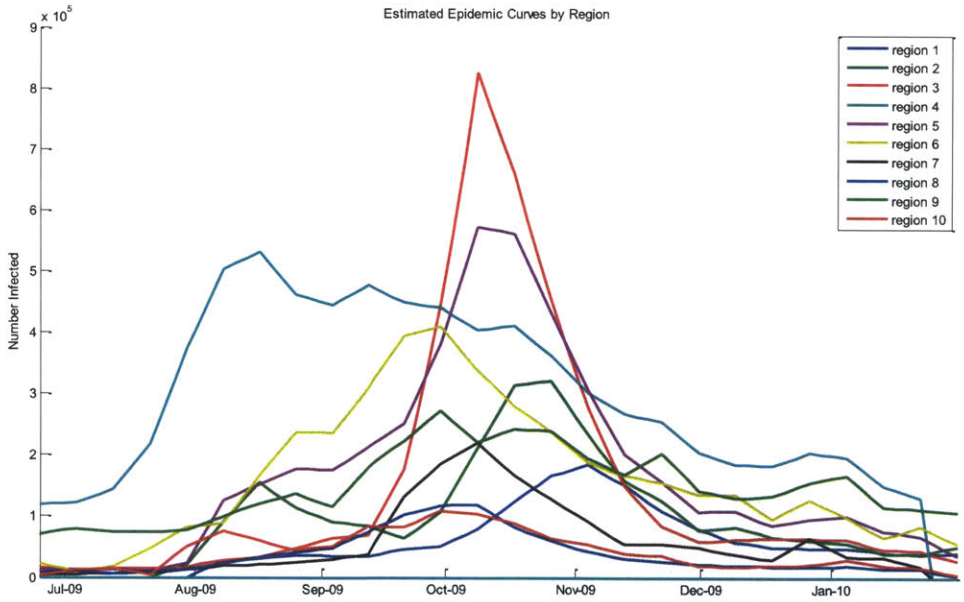
Thus, the number of infections of people of infectivity level λ and susceptibility level s is a random variable picked from a binomial distribution with n =number of susceptible Individuals with parameter λ and $p = p_i(I|\lambda,s) = 1 - e^{-\lambda s \beta_i}$.
After selecting a value from this binomial distribution for each pair (λ, s) we proceed as before to sum the total to number infections for the total N_{i+1}^I. N_{i+1}^S, $f_{i+1}^I(\lambda,s)$, $f_{i+1}^S(\lambda,s)$ were then calculated based on the properties of the individuals randomly selected to become infected.
For the $\left|\frac{1}{\sqrt{n}}\left(\sqrt{\frac{1-p}{p}} - \sqrt{\frac{1-p}{p}}\right)\right| < .3$ [**9**], the binomial random variables were approximated with the appropriate normal distribution.

Appendix C: Parameter values

We used the 10 regions specified by the CDC and fit our parameters to the data available from 2009-2010. For the breakdown of the regions by state see Appendix E.

Region	Population	% Covered by Administered Vaccine
1	14,429,720	32.6
2	19,949,192	20.7
3	28,891,734	23
4	60,580,377	20.1
5	51,766,882	24.2
6	37,860,549	17.8
7	13,610,802	23.5
8	10,787,806	25.4
9	47,495,705	22.1
10	12,734,126	25

Approximated Epidemic Curves by Region



Vaccine Schedule [4]

Date	Allocated	Ordered	Shipped
01/29/10	147,323,810	119,086,300	118,922,220
01/28/10	147,301,010	118,989,300	118,899,820
01/27/10	147,089,010	118,920,900	118,841,120
01/26/10	146,940,510	118,840,000	118,606,720
01/25/10	146,267,090	118,699,900	118,279,120
01/22/10	145,132,850	118,630,900	116,226,920
01/21/10	144,741,750	118,072,700	116,206,320
01/20/10	144,092,400	117,711,000	115,822,020
01/19/10	141,602,700	115,853,000	115,630,620
01/15/10	139,334,000	115,591,600	115,487,820
01/14/10	138,934,400	115,246,700	114,850,120
01/13/10	138,402,600	114,972,000	114,504,220
01/12/10	138,369,200	114,517,000	113,293,220
01/11/10	137,697,600	113,270,700	112,538,020
01/08/10	137,359,300	112,599,500	110,202,920
01/07/10	135,981,600	112,013,100	110,202,520
01/06/10	131,661,400	110,400,200	109,016,520
01/05/10	130,386,600	109,506,200	107,351,020
01/04/10	130,386,600	108,474,200	105,266,520
12/31/09	118,241,100	106,247,500	99,366,920
12/30/09	118,241,100	104,249,900	99,362,620
12/29/09	116,266,700	101,697,200	98,764,920
12/28/09	111,947,700	100,931,500	97,281,820
12/24/09	110,477,700	98,681,000	92,659,820
12/23/09	107,786,500	97,584,000	92,659,820
12/22/09	107,786,500	95,514,300	91,863,720
12/21/09	107,156,700	93,940,500	90,384,920
12/18/09	100,082,700	91,630,300	85,988,420
12/17/09	99,439,500	89,187,200	85,980,220
12/16/09	96,385,200	86,710,200	84,493,120
12/15/09	94,628,100	85,298,200	82,617,120
12/14/09	92,883,100	83,656,200	76,355,920

Fitted parameters

Region	μ	R_0	Epidemic Start Time	Initial Number Infected
1	13.32	1.43	Aug 13, 2009	508
2	11.66	1.26	July 21, 2009	905
3	12.29	1.32	July 28, 2009	2,921
4	10.81	1.18	June 30, 2009	10,000
5	11.06	1.21	July 11, 2009	4,607
6	10.80	1.18	July 2, 2009	8,940
7	11.35	1.24	July 14, 2009	1,710
8	10.88	1.18	July 11, 2009	3,386
9	10.44	1.14	June 30, 2009	6,128
10	10.58	1.16	July 11, 2009	5,656

Other fixed global parameters

Parameter	Value
$s_{vaccinated}$	.02
$s_{unvaccinated}$	.1
Duration of generation	2.3 calendar days
Number of iterations for each heuristic	500

Appendix D: Runtime and testing modifications

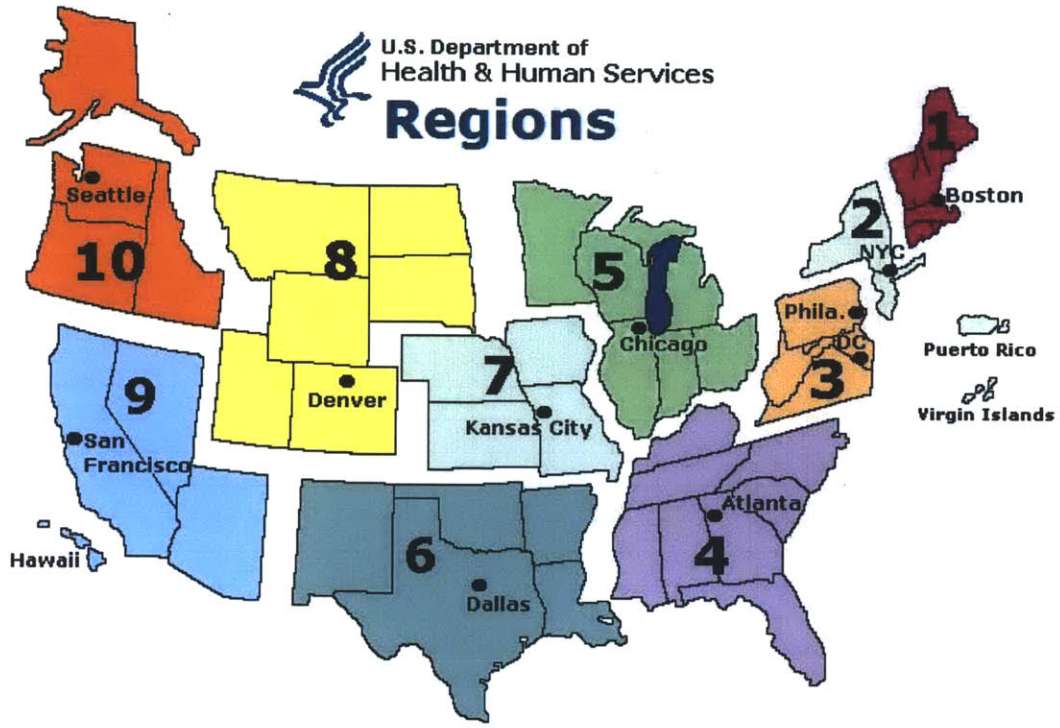
The runtime a single iteration of each of the given heuristics were on average:

Strategy	Runtime (seconds)
No Vaccine	1.1
Pro-Rata	1.3
Pre-Peak	1.3
Critical Period	25.9
Switching Algorithm	87.1

This code has been written in MATLAB and run for 500 on a 4GB dual-core processor.
To speed up runtime, some small modifications to the switching algorithm have been made. For example the minimal value for δ_c has been set to 1024. This did not strongly effect the optimal solution found but reduced the runtime of a single iteration significantly. In a real-life scenario such a modification need not be made as the algorithm only needs to be run once during each decision point. However, during the testing stage, such modifications were necessary.

Appendix E: U.S. regions

Currently, the CDC uses these 10 regions for ILI surveillance.



http://www.cdc.gov/flu/weekly/regions2009-2010/hhssenusmap.htm

Region	States
1	Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, Vermont
2	New Jersey, New York
3	Delaware, Maryland, Pennsylvania, Virginia, West Virginia
4	Alabama, Florida, Georgia, Kentucky, Mississippi, North Carolina, South Carolina, Tennessee
5	Illinois, Indiana, Michigan, Minnesota, Ohio, Wisconsin
6	Arkansas, Louisiana, New Mexico, Oklahoma, Texas
7	Iowa, Kansas, Missouri, Nebraska
8	Colorado, Montana, North Dakota, South Dakota, Utah, Wyoming
9	Arizona, California, Hawaii, Nevada
10	Alaska, Idaho, Oregon, Washington

To test the effect of splitting the region into smaller subregions we propose the following new division of the United States:

Region	States
1a	Maine, New Hampshire, Vermont
1b	Connecticut, Massachusetts, Rhode Island
2a	New York
2b	New Jersey
3a	Pennsylvania
3b	Delaware, Maryland, Virginia, West Virginia
4a	Florida, Georgia, Alabama
4b	Kentucky, Mississippi, North Carolina, South Carolina, Tennessee
5a	Ohio, Michigan, Indiana
5b	Illinois, Minnesota, Wisconsin
6a	Texas
6b	Arkansas, Louisiana, New Mexico, Oklahoma
7a	Iowa, Nebraska
7b	Missouri, Kansas
8a	Colorado
8b	Montana, North Dakota, South Dakota, Utah, Wyoming
9a	California
9b	Arizona, Hawaii, Nevada
10a	Washington
10b	Alaska, Idaho, Oregon

Appendix F: Region by region infection numbers for different heuristics

Region	No Vaccine	Pro Rata	Switching	Adaptive σ
1	6,990,000	4,464,000	1,071,000	1,703,000
2	6,825,000	5,454,000	4,131,000	4,688,000
3	11,610,000	10,311,000	7,515,000	8,752,000
4	15,231,000	13,414,000	14,929,000	14,037,000
5	14,460,000	11,116,000	10,305,000	11,054,000
6	9,484,000	8,577,000	9,375,000	8,891,000
7	4,236,000	3,659,000	3,655,000	3,725,000
8	2,811,000	2,414,000	2,693,000	2,542,000
9	9,861,000	7,002,000	7,652,000	7,441,000
10	2,884,000	2,385,000	2,687,000	2,528,000
Total	84,391,000	68,795,000	64,012,000	65,362,000

Chapter VI: Effective timing and allocation of different types of NPI resources

1. Introduction

In the final chapter of this thesis, we will discuss the effective use of non-pharmaceutical interventions (NPIs) to mitigate the spread of a pandemic in a single geographical region. On a large scale, the most useful pharmaceutical tools we have to curb the effects of influenza are vaccines and antiviral medication. As we saw in Chapter 5, allocating vaccines efficiently is important for high level decision-makers. Now, we would like to zoom in to a smaller population representative of a single US state, where the quantity and timing of vaccines received by the population is controlled from the outside. During the 2009-2010 H1N1 pandemic, some states had to wait up to 3 months [**1**] from the start of the outbreak until the time when vaccines started being administered to the general population. However, during that time, there were a number of measures available to state healthcare officials to help mitigate the outbreak as it was happening through non-pharmaceutical means. In this chapter we will focus on two specific types of non-pharmaceutical interventions and introduce methodology for determining the most effective ways of using NPI tools available to health policy decision-makers in a smaller region.

We focus here on two different types of interventions, though this methodology can be expanded to include other available interventions as well. The first type, we label as *Top Down interventions*. These interventions, like school and business closures, are mandated by the state's government and have a drastic effect on the way that a virus is spread within the community. Recall that if $R_0 = \lambda p$, a school closure reduces the value of λ. School children are a high-activity group, often making hundreds of close contacts in one day, so reducing their contact rates by keeping them home can be very effective in reducing the overall transmissibility of the disease. Of course, just because schools are closed, there are no guarantees that students stay home; they might easily socialize outside of school. However, recent research on school closures in Alberta, Canada in 2009 during the H1N1 pandemic suggests that the rate of transmission of the virus in the first wave was reduced by as much as 50% after schools were closed throughout the province [**2**]. Further literature on school closure effectiveness suggests similar decreases in incidence (42-50%) [**3, 4**]. A very detailed, historical review of school closure effectiveness, published in the middle of the 2009-2010 H1N1 outbreak, suggests that the total number of infections might only be reduced by 15%. However, the peak of the infection and the strain on the healthcare system would be reduced by as much as 40% [**5**].

Unfortunately despite the potential effectiveness of school closures, a community simply cannot keep schools closed for a long time. Similarly, even the most technologically advanced company can hardly keep its doors closed for longer than a few weeks. Furthermore, school closures are often made on a district level causing high levels of variation in the timing and duration of school closures even within a single US state [**6**]. Hammond and Epstein analyzed the potential economic effects of school closures, taking into account the worker absenteeism resulting from parents needing to care for their children. They estimated that closing all schools in the United States for 4 weeks could cost anywhere between $10 and $47 billion, costing somewhere between $35 and $157 per student per week of school closure [**7**]. With such costs, it would be impossible to keep schools closed for any period of time longer than those four weeks, and even four weeks would be difficult. The economic impact of such interventions is the reason we consider these to be "acute" interventions. They are drastic and short lived. Throughout this chapter we will refer to these types of interventions as Top Down or acute interventions interchangeably.

One might conjecture that Top Down interventions should be enacted as soon as possible after the discovery of a new, dangerous virus. Indeed, in 2009, schools in Mexico City were closed within a week of the pandemic alert. Business and other non-essential activities were shut down another week later. In total, the schools were closed for 3 weeks, reopening in the middle of May, 2009 [**8**]. Even though schools were reopened well after the peak of the epidemic curve, when new cases were at a minimum, once schools and businesses reopened, Mexico saw a resurgence of H1N1, with a second peak coming in June of 2009. While school closures curbed the initial transmission of the virus, once contact rates were restored to their previous levels, the transmissibility of the virus soared again. A recent paper by Hollingsworth et al., confirms that closing schools as soon as possible might not be the best course of action. Depending on the objective at hand, short term interventions such as school closures can be most effective later during the outbreak, closer to the peak of the epidemic curve [**9**].

The second type of intervention we will use in this chapter is termed *Bottom Up interventions*[1]. These are the non-pharmaceutical interventions that, unlike school closures, can be used over a prolonged period of time. Such interventions from a governing body include public awareness advertising campaigns and educational seminars to promote good hygienic behavior. Public awareness campaigns can take the form of local TV and radio advertising or ad campaigns in public spaces promoting hand washing and telecommuting. Diligent hand washing and sanitizing has been shown to reduce influenza infections by 20% to 95% [**10**]. Of course, such campaigns are not binding to the public. Even though they are sponsored by the government, the initiative to implement suggested behavior is left to the individuals. However, unlike the short-lived effect of Top Down interventions, the effect of behavioral changes in the population may last for a long time. A combination of Top Down and Bottom Up interventions can go a long way in holding off a major outbreak until vaccines and other pharmaceutical agents arrive to a community, and can even stop a minor outbreak altogether.

2. The multiple-intervention timing problem

Consider the problem faced by a single US state in August 2009. The state is preparing for an outbreak of the novel H1N1 influenza virus and has been promised vaccine shipments that will start in October. The outbreak is just beginning in this particular state, and we have various NPIs at our disposal. We would like to know how the state should focus its efforts to achieve the best effect in curbing the total number of infections.

2.1 Problem formalization

To formalize the problem we will once again use a discrete time heterogeneous model from Chapter 2 with two levels of heterogeneity, λ, and s. We include heterogeneity in the form of λ and s to show the effect of various interventions focused on decreasing people's daily contacts and decreasing the probability of spreading infection from one person to another. We choose to exclude infectivity heterogeneity to avoid unnecessary computations that would not help in drawing the insights from the methodology in this chapter. Finally, recall that a model "day" or "generation" lasts approximately 2.3 real days. We will consider each model day as a decision point, at which a government body can choose to enact various interventions while adhering to budget constraints.

We assume that the vaccines will start being delivered to this state in T_v weeks. After that, they will be shipped on a weekly basis just as they had been in 2009. For modeling purposes, we assume that T_v is a

[1] We will also recall to this class of interventions as *continual interventions* in this chapter.

known deterministic value. In reality, this is likely to only be an estimate of when vaccines will become available, and will have some variance. Given a known value of T_v we consider two NPI options available to state healthcare officials.

The state can use a Top Down NPI (school/workplace closures, etc) for the duration of L_a days. This will have the effect of reducing the value of λ for some fraction F_a of the population. Since the acute interventions in this chapter generally target schoolchildren, who are known to have higher than average activity levels, [**11**] in our modeling we will take into account the fact that the people affected by school closures will be relatively active. To model Top Down NPIs, we first need to quantify the effect of any NPI on reducing the spread of the infection in the population. We do this by introducing the term *NPI-resource*. NPI-resource is a unitless parameter that quantifies the amount of NPI used by the government on a given day. We will talk about fitting values to this parameter later on in the chapter, but for now we call the amount of total NPI-resource provided by school closures X.

X – The amount of daily NPI-resource provided by implementing the Top Down NPIs. That is, implementing a Top Down NPI at time t_a, provides X units of *NPI-resource* for days $t = t_a, t_a + 1, \ldots t_a + L_a - 1$. $X \geq 0$, since we assume there is no such thing as negative resource. This implies that no NPI would actually increase anybody's λ or s.

Due to the costs associated with a Top Down campaign, we assume that this intervention may only be used once during the epidemic period.

In addition to the Top Down interventions, we can also implement a public service announcement program that will encourage people to reduce their overall values of λ and s through telecommuting, hand-washing, and other behavioral changes. This NPI program affects the entire population. We enact it as follows: during each generation t, the government can reduce the public's λ, and s values, by using a certain amount of available Bottom Up NPI resource y_t, in addition to school closure reductions if they are in effect.

y_t – The amount of NPI-resource provided by implementing the Bottom Up intervention on day t. $y_t \geq 0$ for all t.

Note that while X is a fixed amount, provided by a one-time Top Down intervention, y_t is variable and can be changed on a daily basis by the authorities.

This effect is not long-lasting, however, and experiences geometric decay. That is, if a public service campaign provided y_t amount of resource on day t, then some of this effect lingers on day $t + 1$. That effect of the previous day's advertisement becomes βy_t for some $\beta < 1$. We call β the *knowledge decay parameter*. The next day, the leftover effect from day t is $\beta^2 y_t$ and so on.

We split our population into two groups. The first is the fraction $F_a$of the population that is affected by acute as well as continual interventions (e.g. schoolchildren). The second group includes the rest of the population, which is only affected by continual interventions. We can now combine the total amount of NPI-resource by representing the total amount of resource experienced by each of these groups. Let

z_t^{SA}– The amount of NPI-resource experienced by the group affected by Top Down intervention due to the acute, Top Down intervention, on day t.

This is

$$z_t{}^{SA} = \sum_{i=0}^{i=t} \beta^{t-i} y_i$$

Similarly,

z_t^{SC} – The amount of NPI-resource experienced by the group affected by Top Down intervention due to the continual, Bottom Up intervention, on day t.

This is

$$z_t{}^{SC} = \sum_{i=0}^{i=t} x_t$$

where x_t is X when acute interventions are in effect, and 0 otherwise.

Finally, we define

z_t^N – The total amount of NPI-resource experienced by the group not affected by Top Down intervention on day t.

$z^N{}_t$ is just:

$$z_t{}^{N} = \sum_{i=0}^{i=t} \beta^{t-i} y_i = z_t^{SA}$$

since this group is not affected by school closures.

Now we are left with calculating the effect of $z_t{}^N, z_t{}^{SA}$ or $z_t{}^{SC}$ worth of NPI-resource on λ and s. We assume that Bottom Up interventions will decrease a typical person's λ and s values by some factor, while Top Down interventions will decrease a typical schoolchild's λ value by a different factor. To quantify this we define the following constructs:

$\gamma_s^S(z_t^{SC})$ – the proportional decrease in s experienced by a person in the Top Down affected group (e.g. schoolchild) when the total amount of Bottom Up resource experienced by the group is z_t^{SC}.

$\gamma_\lambda^S(z_t^{SA}, z_t^{SC})$ – the proportional decrease in λ experienced by a person in the Top Down affected group (e.g. schoolchild) when the total amount of acute resource experienced by the affected group is z_t^{SA} and the total amount of continual resource experienced by the group is z_t^{SC}.

In other words, if we were to pick a person, from the affected group, whose original (pre-epidemic) heterogeneity parameters equaled λ and s, her new parameters to be used in the modeling would now be

$\tilde{s} = \left(1 - \gamma_s^S(z_t^{SC})\right) s$, and

$$\tilde{\lambda} = \left(1 - \gamma_{\lambda}^{S}(z_t^{SA}, z_t^{SC})\right)\lambda.$$

Similarly,

$\gamma_s^S(z_t^N)$ – the proportional decrease in s experienced by a person in the group not affected by Top Down interventions when the total amount of resource experienced by the group is z_t^N.

$\gamma_{\lambda}^{S}(z_t^N)$ – the proportional decrease in λ experienced by a person in the group not affected by Top Down interventions when the total amount of resource experienced by the group is z_t^N.

We are now left with calculating $\gamma_s^S(z_t^{SC}), \gamma_{\lambda}^{S}(z_t^{SA}, z_t^{SC}), \gamma_s^S(z_t^N)$ and $\gamma_{\lambda}^{S}(z_t^N)$.

We start with $\gamma_{\lambda}^{S}(z_t^{SC}, z_t^{SC})$:

$$\gamma_{\lambda}^{S}(z_t^{SA}, z_t^{SC}) = \left(1 - e^{-\sigma_{\lambda}^{S}(z_t^{SC} + z_t^{SA})}\right) M_{\lambda}^{S}, \text{ for some } 0 \leq \sigma_{\lambda}^{S} \leq \infty, 0 \leq M_{\lambda}^{S} \leq 1$$

Here σ_{λ}^{S} is an intensity parameter representing the responsiveness of people's λ values to implemented NPIs, and M_{λ}^{S} is the maximum attainable reduction in λ. Note that other monotonically decreasing functions bounded by 0 and 1 could be used here, but the one we chose is convenient. We need only to use a monotonically increasing function that follows the following three constraints:

- $\gamma_{\lambda}^{S}(z_t^{SA}, z_t^{SC}) = 0$ at $z_t^{SA} + z_t^{SC} = 0$,
- $\lim_{z_t^{SA} + z_t^{SC} = \infty} \gamma_{\lambda}^{S}(z_t^{SA}, z_t^{SC})$ approaches M_{λ}^{S}, for $M_{\lambda}^{S} \leq 1$, and
- $\frac{d^2\gamma}{dz_t^2} < 0$.

Concavity is necessary due to diminishing returns to scale. It is increasingly more difficult to get people to decrease their λ values those last few percent. A person might be willing to telecommute to work, but will still see her family during the day. Getting that person to quarantine herself entirely, and ensure that her $\lambda = 0$, is more difficult than simply limiting her contacts to a few important individuals.

$\gamma_s^S(z_t^{SC})$ is calculated similarly:

$$\gamma_s^S(z_t^{SC}) = \left(1 - e^{-\sigma_s^S(z_t^{SC})}\right) M_s^S, \text{ for some } 0 \leq \sigma_s^S \leq \infty, 0 \leq M_s^S \leq 1$$

Similar calculations apply to $\gamma_s^S(z_t^N)$ and $\gamma_{\lambda}^{S}(z_t^N)$.

We borrow some of the ideas behind knowledge decay and the total effect of multiple interventions from the study of advertising effects in marketing science [**12**], and we will fit the parameters to the data for a real US state in 2009. See Section 2.2 and Appendix B for fitted parameter values.

Given the ability to control these two types of interventions, we must now find the best time to implement the one-time acute intervention as well as the timing and intensity of the public awareness campaign. Naturally, we are under various constraints. As we mentioned before, acute interventions require significant costs and planning so they can only be implemented once on a community level. Depending on the government structure in the community, schools can close individually, but we consider only the

effect of organized school closures that affect a significant proportion of the population. The only decision variable related to these interventions is the variable T_a, representing the time when the intervention begins. After Top Down NPIs have been deployed, they last for the entire intervention length L_a (a few weeks for school and business closures, longer for non-essential enterprises).

Our second non-pharmaceutical intervention, the public service campaign is also constrained. Due to budget constraints, we can only purchase so much advertising time and space, so the total amount of Bottom Up NPI-resource must be less than some given budget constraint B. Recall that we call the amount of Bottom Up resource provided at time t (via some sort of advertisement) the decision variable y_t. The corresponding budget constraint is given by

$$\sum_{t=0}^{t=\infty} y_t \leq B.$$

Given these two constraints we would like to solve the problem of minimizing the total number of infections having control only of T_a and y_t for all t. In Appendix A, we provide a summary of all definitions in this section that will be used to model decision making in the rest of this chapter. Of course, the total number and types of interventions can be tailored to the specific situation in a given region. We present here the methodology for minimizing the total number of infections using these two specific interventions with these given constraints, but the same methodology may be used with any number of different interventions and corresponding constraints.

2.2 Model parameter values

To try to solve the problem of minimizing the total number of infections we use a heterogeneous model of a community modeled after the state of Indiana during the 2009-2010 H1N1 pandemic. This state was fairly typical of the US as a whole, with vaccines first being administered at the same time as the epidemic reached its peak within that state. 26.5% of Indiana's population had been vaccinated by the end of the epidemic and we have weekly vaccine administration data provided by the state. We used these data to determine the parameters of our model (Table 1). See Appendix B for a detailed description of how each parameter was obtained.

Model Parameter	Description	Value
μ	Parameter for contact-rate distribution $f(\lambda)$ in the population	11.06
$s_{vaccinated}$	Susceptibility of a vaccinated person	.02
$s_{unvaccinated}$	Initial expected susceptibility of an unvaccinated person	.1
L	Length (in generations) of the Top Down NPI implementation (e.g. school-closures)	5
T_v	Time (in generations) from the beginning of the epidemic until the arrival of first vaccines	24
F_a	Proportion of the population affected by acute interventions[2]	.248
X	Amount of available Top Down NPI resource	.5
B	Maximum total effect of Bottom Up interventions.	.20
N	Population of the community	6,483,802
β	Knowledge decay factor	.5
$\sigma_\lambda^S; \sigma_s^S; \sigma_\lambda^N; \sigma_s^N$	The artificial variable that describes the concavity of the function that determines the effect of NPI use on a given day	.2

Table 1: Model parameters for the minimization problem. See Appendix B for modeling details.

Decision Variable	Description	Constraint
T_a	Time of the Top Down NPI implementation	[0, 100], integer
y_t for $t = 1 \ldots 100$	Bottom Up NPIs implemented	$\sum_{t=1}^{\infty} y_t \leq B$

Table 2: Decision variables in the minimization problem.

As in Chapter 5, we assumed that vaccines were 98% effective because that number was representative of the H1N1 vaccine. However, we also assume, unrealistically, that vaccines become effective immediately after they are administered. This chapter focuses on describing the methodology to find a balance between different types of interventions and gain insights on timing the implementation of these interventions during an outbreak. In preparation for a known outbreak, similar analyses should be performed with parameters that reflect the effectiveness of the available vaccines for that particular outbreak.

Given these estimated parameter values we first approach a simpler problem by looking at continual and acute interventions independently.

3. Top Down/Bottom Up intervention tradeoff

We first look at a simple problem of implementing only Top Down interventions. While Mexican officials in 2009 closed schools and businesses as soon as the H1N1 outbreak was determined to be

[2] Since school closures tend to affect those individuals on the right-hand-tails of the activity and susceptibility spectra in our modeling framework, we select the people affected by acute interventions proportionally to the product of the activity level and susceptibility level of the individual. That is, the probability that a person with activity level λ and susceptibility level s is affected by an acute intervention is $\frac{\lambda s}{E(\lambda s)}$.

particularly dangerous, researchers now believe that with these short term interventions a delayed response might be more appropriate [**9**]. That is, when an intervention may only be held in place for some limited period of time, it is useful to delay implementing the intervention until the epidemic is at its worst. We use the model introduced in Chapter 2 and the parameters Table 1 to find the best time to use Top Down NPIs in a region modeled after the state of Indiana. For now, we ignore the availability of vaccines or Bottom Up interventions.

See Figure 1a for the percentage of total infections averted by using acute NPIs at different points in time. Judging from the figure, the most effective time to start acute NPIs is right before the peak of infections, so that the NPIs remain in effect through the peak of the outbreak and last until just after the peak is over.

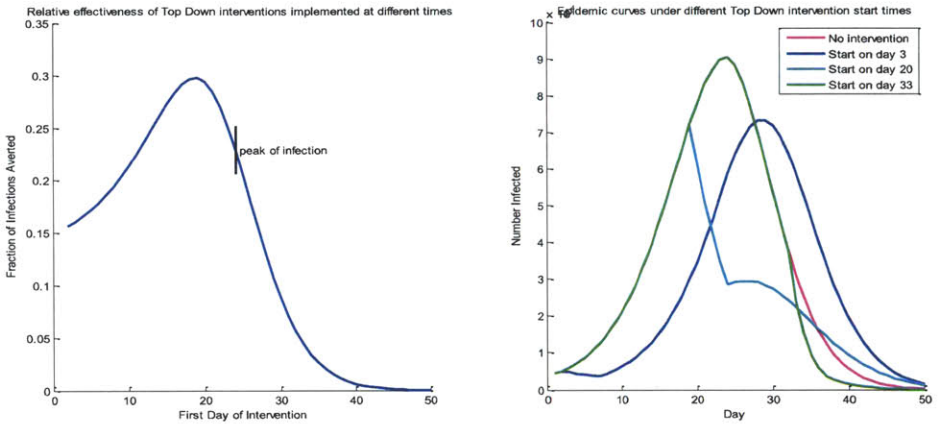


Figure 1: Effectiveness of acute interventions implemented at different times. The left graph shows the number of infections averted through acute interventions only, as a function of the first day of implementation. The right graph contains the epidemic curves for a few select intervention scenarios.

Note from Figure1b that the epidemic curve is only delayed when acute NPIs are implemented too early and its shape remains relatively the same as it would have been without interventions. On the other hand, late NPIs only lower the tail of the epidemic when it is already well on the decline. The middle ground is most effective. Top Down NPIs implemented close to the peak of the epidemic curve cause the pandemic to be both delayed have lower peak.

Now that we know the best time for acute NPI implementation, let us look separately at the effect of Bottom Up interventions. Unlike the case with acute NPIs, which must be implemented just once, continual Bottom Up interventions can be implemented at different times at any level of intensity, and we will discuss how to choose the times and intensity levels in Section 4. For now, however, we will try to get some intuition by restricting our decision-making freedom for Bottom Up interventions. Suppose that we may only use one Bottom Up intervention, such as a public safety campaign, throughout the entire duration of the outbreak. We have no control over this intervention; the amount of intervention used by the government at any time must stay constant throughout the entire epidemic period. We relax this restriction later in this chapter.

Suppose, for example, that the intervention strategy for this outbreak is defined by $y_t = .016$ for $t = 1...100$, and $\beta = .2$ as in Appendix B. In other words, for the entire 100-generation duration of the

epidemic, the government allocates $y_t = .016$ amount of Bottom Up NPI-resource to the population with a knowledge decay parameter of $\beta = .2$. No other interventions or vaccines are available.

The green line in Figure 2 represents the percentage of total infections averted (against a baseline of no interventions) if the only interventions available to us had been a constant level of continual intervention. That is, approximately 24% of infections that would have occurred with no intervention have been averted by using the Bottom Up NPIs described above.

If we had a hypothetical choice between our restricted continual interventions and acute interventions, acute interventions would only be preferable if they were implemented at a time when they are more effective than Bottom Up NPIs. Using the parameters fit from the outbreak in Indiana in 2009-2010, we see that the Top Down NPIs (blue curve) are only more effective than Bottom Up NPIs (green line) when they are implemented sometime within the period of 3 weeks, beginning just before the peak of the epidemic. If we could only choose one type of intervention, we would want to choose acute-interventions only if we knew with high confidence that we could estimate the timing of the peak of the epidemic curve within 3 weeks. Otherwise, we would be more effective going with the continual intervention choice.

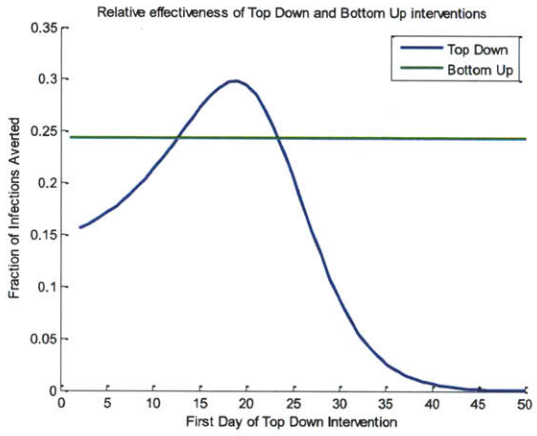


Figure 2: Comparison of the relative effectiveness of acute interventions implemented at different times and continual interventions implemented over the entire outbreak period. The green line represents the total number of infections averted if only Bottom Up interventions are used throughout the entire time of the epidemic.

In Figure 3, we test the sensitivity of our results to our model parameters. These graphs show the relative effectiveness of the interventions with different epidemic parameters Figure 3a contains the variation in effectiveness for several values of R_0 while Figure 3b includes the percentage of infections averted assuming different levels of effectiveness for the continual intervention.

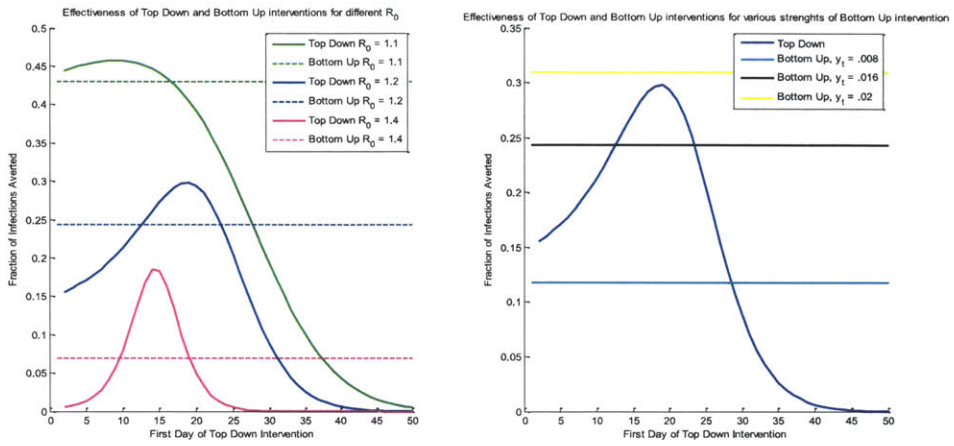


Figure 3: Effectiveness of acute and continual interventions with varying R_0 and effectiveness of the continual interventions.

The key insight from this quick study is that acute interventions are only as effective as a constant continual intervention if they are implemented during a certain period close to the peak of infection. This intuition can guide us in the next section.

In reality, continual interventions need not be spread out over the entire period of the infection. Policy-makers have control over how to allocate their resources for disease prevention campaigns. Bottom Up resources can be used in any amounts at any time subject to some reasonable budget, so from now on we relax the constraint on continual interventions. y_t can now take on any value as long as $y_t \geq 0$ and $\sum_{t=0}^{t=\infty} y_t \leq B$.

In addition, policy-makers are not faced with a tradeoff decision. Budgets are often allocated for both acute and continual interventions. When dealing with two types of interventions, especially when the Bottom Up interventions need not be used throughout the entire period, we need to resort to more intricate ways of finding the solution to the multi-intervention problem.

4. Decision algorithm

Now that we have defined the problem, we can utilize our discrete time steps in our model to formulate a dynamic program to solve it.

Let us define a state S_t on day t to be the tuple (Θ, Q, U, W). Here, Θ is the aggregate variable that contains all of the information that is needed to define a state of an outbreak. That is, the number of susceptibles in the population at time t, the number of infected people at time t, and the respective λ and s distributions in each of those subpopulations. In addition, Q, U, and W represent the resources that are still available to be used on day t. Q is the amount of Top Down resource that can still be used, that is, $Q = X$ if Top Down NPIs have not been implemented yet, and $Q = 0$ otherwise. Similarly U represents the amount of Bottom Up intervention that has still not been used, and equals $B - \sum_{i=0}^{i=t-1} y_t$. Finally, W is the aggregate state variable that represents the total amount of Top Down and Bottom Up interventions carried over from previous days. That is $W = \{\sum_{i=0}^{i=t-1} \beta^{t-i} y_i\,;\, x_i\}$ where

$$x_i = \begin{cases} X \text{ if Top Down NPIs have been implemented in the last } L_a \text{ days} \\ 0 \text{ otherwise} \end{cases}.$$

Our objective is to minimize the total number of infections resulting from the outbreak using the two available NPIs. This state space is too complex to be solved via dynamic programming directly due, once again, to the non-Markovian nature of an influenza spread model. Instead, we can use a heuristic algorithm to determine the approximate solution that would help choose the best timing for various interventions available to healthcare officials.

The first approximation we must use is splitting up the decision space into m segments so that the Bottom Up intervention that can be used at a given time t can only be used in increments of $\frac{B}{m}$.

Given that, we can still make m choices in each generation, so the total decision space for Bottom Up interventions is still large, $m^{number\ of decision\ days}$. We can still not solve this exactly. However, we can use a roll-out algorithm where we solve a small problem (k time periods, for example) at each step, and use the solution to move forward to the next generation. To use the algorithm, we need to use some sort of heuristic estimate for the objective value of given tuple (Θ, Q, U, W). That is, for any state at time t, represented by (Θ, Q, U, W), we estimate the approximate total number of infections that will result after the outbreak has passed if we used our available resources to the best of our abilities -- $V(\Theta, Q, U, W)$.

To make this estimation at time t we can only use the information available to us about the epidemic curve from the beginning of the outbreak until time t. The model parameters from the simulation are not available to the decision makers. Therefore, to make an estimate of the number of infections that will occur in the future we first estimate the parameters of the model (i.e. average initial contact rates and susceptibility levels in the population as well as the transmissibility of the flu) based on available data. After the parameters have been estimated, we can use the estimated model to calculate an estimate for $V(\Theta, Q, U, W)$.

We propose the following estimates: $V(S_t) = V(\Theta, Q, U, W)$:

> If the peak has already passed on day t, $V(S_t)$ is the model-generated total number of infections that would occur if we
>
> - use any available acute NPIs (Q) immediately, and
> - use all of the available continual NPIs (U) over the next 3 days [3].
>
> If the peak has not yet passed on day t, we first estimate the time when the peak will occur using a fit to the data observed up to time t. $V(S_t)$ is then the model-generated total number of infections that would occur
>
> - if any available acute NPIs (Q) are used exactly at the (model-estimated) peak, and

[3] We chose 3 days as a tradeoff because it produced the best results in estimating the value of a given state. Using all NPIs in one or two days is not very effective because it provides only temporary relief to the outbreak. On the other hand, spreading the resource out for has little effect during the peak of the epidemic, the time when this resource is most effective.

- the available continual NPI (U) are used in the 3 days around the peak, i.e. day before the peak, day of peak, and day after the peak.

The decision to use Top Down NPIs during the peak of the outbreak is based on evidence Figure 2, which implies that acute NPIs are most effective when they begin at a time close to the peak of the epidemic curve.

With this estimate we find the best decision by moving forward in time. For each generation t, we find the decisions y_t and determine if $T_a = t$ by using the approximation $V(S_{t+k})$ for the value of the objective function at time $t + k$.

The additional parameters used for this simulation are included in Table 3 below.

Algorithm Parameter	Description	Value
k	Time periods considered using roll-out algorithm	4
m	Continual NPI budget constraint increment	10

Table 3: Parameters used for the ADP problem with two types of intervention.

5. Results

We ran this simulation on the data from Indiana with the parameters from Table 1 in Section 2. In the 2009-2010 H1N1 outbreak Indiana started administering its vaccines on the same week as the peak of the epidemic curve occurred. Prior to this, some school closures were in effect locally, but no centralized decision has been made. We ran the algorithm to establish how effective various NPIs could be in a situation where vaccines are delayed for months as has happened in Indiana in the summer and fall of 2009, with the best decisions displayed in Figure 4. Note that the acute NPIs are once again used four generations (about 1.5 real life weeks) prior to the peak at $t = 24$, which coincides with the beginning of vaccine distribution. While the total numbers can only be used as an approximation to gain insight in this experiment, the total number of model-generated infections that would occur without NPI use is approximately 1.8 million while the number of infections with NPIs is cut down to 1.2 million -- a 33% reduction in the total number of infections.

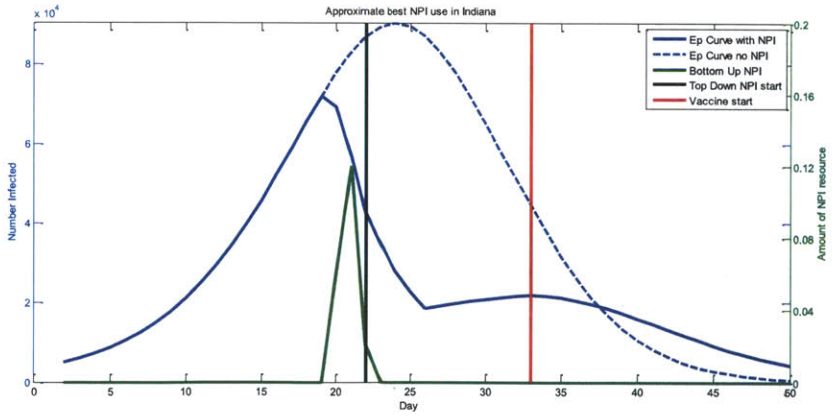


Figure 4: The continual and acute NPI decisions generated by our algorithm and the resulting epidemic curves for the hypothetical NPI decisions for a community with the demographics of Indiana during the 2009-2010 H1N1 epidemic.

Consider the effect of each type of intervention. From Figure 2b and Figure 4 we can discern that the effect of an acute intervention on the epidemic curve is temporary but easily noticeable. Once acute NPIs are enacted, the slope of the epidemic curve decreases significantly, often turning negative. In Figure 4 the epidemic curve slopes downward for 5 generations until acute interventions are no longer in effect. After the acute NPI is no longer in effect, the slope turns positive again, and the epidemic continues slowly with the help of vaccines until it dies out more than 50 generations after its start.

The effect of Bottom Up interventions is more subtle. Bottom Up interventions are more spread out than the acute interventions. Their effect is manifested through a lower, flatter peak of the epidemic curve. Due to their effect, the epidemic lasts for a longer time, but the peak infection rates are lowered causing two positive effects. The first is the decrease in the immediate stress on the health care system, as hospitals are less likely to experience a surge in the number of cases and admissions. The second is the delay in the most stressful period of the infection, giving the community more time until it starts receving the more effective intervention - vaccines - from the government. There is a secondary effect of lowering the area under the epidemic curve – the total number of infections . However this effect is lower than the other two, and does not occur at all if all of the Bottom Up resources are used too early.

These observations regarding individual interventions are consistent with the findings of Hollingsworth et al. Using these individual insights, we can combine two types of interventions to obtain the most benefit during the waiting period for vaccine arrival. In the case when vaccines start arriving fairly late but at a time when they can still make a difference (e.g. peak of epidemic curve), continual interventions are the first line of defense. They cause the epidemic to halt its growth and flatten the peak of the infection buying more time for vaccine arrival. The second line of defense, acute intervention, must be timed well to obtain the most effect a few generations before the projected peak of the disease. The combination of the two provides enough delay to maximize the combined effect of the non-pharmaceutical interventions and influenza vaccines.

Naturally, the way in which the two should be combined is not always the same as in Indiana. The optimal combination is strongly dependent on the time when vaccines finally start arriving. During the 2009-2010 H1N1 epidemic, the first vaccines arried as early as 11 weeks prior to the peak in some states and as late as 6 weeks after the peak in others [**1**]. We ran this algorithm for vaccines ariving from six weeks prior to the peak to 5 weeks after the peak, with the schedule of administered vaccine remaining the same as in 2009-2010. In Figure 5 we display the best algorithm-generated decisions for this community if vaccines arrived at different times.

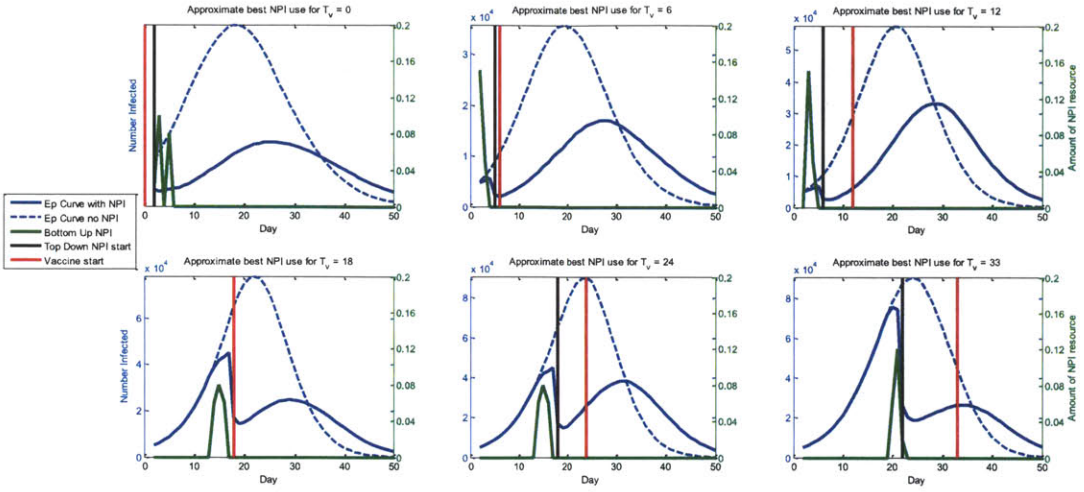


Figure 5: Best decisions for Top Down and Bottom Up interventions with vaccines arriving at different times.

We summarize the key observations from this experiment:

- **Top Down NPIs**
  - If vaccines start arriving before peak, acute NPIs are initiated some time (not more than five generations) before the initial vaccinations. This works to immediately decrease the transmission rate in preparation for vaccines. Recall that we assume that vaccines become effective immediately. In reality, vaccines are most likely to come into effect about 10 days after they have been administered causing, so acute NPIs are likely to be most effective a few generations later than in our results.
  - Otherwise, NPIs are most effective right before the peak.
- **Bottom Up NPIs**
  - These are used almost always before acute interventions. The only exception to this rule occurs when vaccines are administered at time $t = 0$.
  - When vaccines arrive particularly late, Bottom Up NPIs are combined with acute NPIs for maximal effect close to the peak of the infection.
  - When vaccines do not arrive too late, the Bottom UP and Top Down NPIs are spread out to flatten the epidemic curve and hold off the major outbreak for as long as possible until at least some vaccines become available.

In essence, while temporary non-pharmaceutical interventions do not have a substantial net effect when used early, properly timed interventions are noticeably helpful in decreasing the total number of infections and easing the strain on a community during a serious influenza epidemic event.

6. Discussion

The results in Section 5 should not be taken as the exact solution for Indiana in the event of a future pandemic. However, the methodology described above allows us to gain insights on how to time the different interventions that are available to healthcare workers prior to being able to distribute permanent immunity through vaccines. Indeed, in this chapter we are benefitting from hindsight in the knowledge of

the epidemiological characteristics of the virus such as R_0. This allows us to model the shape of the epidemic curve and the possible effects of different interventions on the number of infections caused by the outbreak with great accuracy. During the next pandemic event, however, we will no longer have this information, and we will need to use the approach described above with possibly weak estimations of the parameters.

Despite this drawback, the main benefits of this exercise are the insights that we draw from performing this analysis in preparation to an outbreak. The results in this chapter show consistent trends in the best use of continual and acute types of interventions. The key to minimizing infections seems to lie in estimating the timing of the peak of the epidemic curve and its relation to the time that vaccines become available. When vaccines are expected reasonably early, NPIs are most useful in flattening the epidemic curve and delaying the peak until vaccines are available. This can be done through attacking the outbreak by scheduling acute interventions such as school closures a few weeks prior to this peak. Furthermore, continual interventions are most useful in delaying the peak when scheduled a few weeks prior to start of the acute interventions. If very late vaccines are anticipated, on the other hand, we benefit most by using the NPIs at the time when they are most effective, close to the peak of the epidemic.

We have only used one objective function in this analysis. Many other objective functions may be used instead of simply minimizing the total number of infections in the community. For example, one potential objective for a smaller community than an entire country is to minimize the surge strain on the healthcare providers during the peak of the outbreak. The objective of flattening the epidemic curve is discussed along with others in Hollingsworth et al. The methodology in this chapter can include other objectives and other types of intervention; it should be used to gain insight on the best ways to mitigate the effects of an outbreak when pharmaceutical interventions are not easily available.

References

1. Finkelstein S, Hedberg K, Hopkins J, Hashmi S, Larson R. Vaccine availability in the United States during the 2009 H1N1 outbreak. *Am J Disaster Med.* Jan-Feb 2011;6(1):23-30.
2. Earn D, He D, Loeb M, Fonseca K, Lee B, Dushoff J. Effects of School Closure on Incidence of Pandemic Influenza in Alberta, Canada. *Annals of Internal Medicine.* 2012;156:173-181.
3. Heymann A, Chodick G, Reichman B, Kokia E, Laufer J. Influence of school closure on the incidence of viral respiratory diseases among children and on health care utilization. *The Pediatric Infectious Disease Journal.* July 2004;23(7):675-677.
4. Haber M, Shay D, Davis X, et al. Effectiveness of interventions to reduce contact rates during a simulated influenza pandemic. *Emerg Infect Dis.* April 2007;13(4):581-589.
5. Cauchemez S, Ferguson N, Wachtel C, Tegnell A, Saour G. Closure of schools during an influenza pandemic. *The Lancet Infectioius Diseases.* August 2009;9(8):473-481.
6. Klaiman T, Kraemer J, Stoto M. Variability in school closure decisions in response to 2009 H1N1: a qualitative systems improvement analysis. *BMC Public Health.* February 2011;11(73).
7. Lempel H, Hammond R, Epstein J. Economic cost and health care workforce effects of school closures in the US. *Center on Social and Economic Dynamics Working Paper.* September 2009.
8. Secretaria De Salud, Mexico. Situacion actual de la epidemia. October 2009, 2009. Available at: http://portal.salud.gob.mx/sites/salud/descargas/pdf/influenza/situacion_actual_epidemia_121009.pdf. Accessed February 13, 2012.
9. Hollingsworth T, Klinkenberg D, Heesterbeek H, Anderson R. Mitigation Strategies for Pandemic Influenza A: Balancing Conflicting Policy Objectives. *PLoS Computational Biology.* February 2011;7(2).
10. Finkelstein S, Prakash S, Nigmatulina K, McDevitt J, Larson R. A home toolkit for primary prevention of influenza by individuals and families. *Disaster Med Public Health Prep.* Dec 2011;5(4):266-271.
11. Mossong J, Hens N, Jit M, Beutels P, Auranen K. Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS Med.* 2008.
12. Little J. Aggregate advertising models: the state of the art. *Operations Research.* August 1979;27(4):629-667.
13. Corn J. Indiana Swine Flu Closing Hits Washington Township Elementary School. *Yahoo! News.* May 2009.
14. indystar.com. Schools have grappled with outbreaks before. *IndyStar.* Nov 2011.
15. US Census Bureau. State & Country QuickFacts. January 2012. Available at: http://quickfacts.census.gov/qfd/states/18000.html. Accessed February 29, 2012.

Appendix A: Summary of definitions from Section 2

Variable	Definition	Units	Type
T_v	Time from beginning of outbreak until the beginning of vaccine arrivals	generations	Fixed Parameter
L_a	Length of the acute-NPI implementation (e.g. school-closures)	generations	Fixed Parameter
F_a	Proportion of the population affected by acute interventions		Fixed Parameter
X	Amount of NPI-resource provided by Top Down interventions		Fitted Parameter
y_t	Amount of NPI-resource provided by Bottom Up interventions on day t		Decision Variable
β	Knowledge Decay parameter		Fitted Parameter
z_t^{SA}	Amount of NPI-resource experienced by the group affected by Top Down intervention due to the acute, Top Down intervention, on day t.		Calculated using y_t, β
z_t^{SC}	Amount of NPI-resource experienced by the group affected by Top Down intervention due to the continual, Bottom Up intervention, on day t.		Calculated using X, T_a
z_t^N	The total amount of NPI-resource experienced by the group not affected by Top Down intervention on day t.		Calculated using y_t, β
$\gamma_s^S(z_t^{SC})$	proportional decrease in s experienced by a person in the Top Down affected group.		Calculated using $z_t^{SC}, \sigma_\lambda^S, M_\lambda^S$
$\gamma_\lambda^S(z_t^{SA}, z_t^{SC})$	proportional decrease in λ experienced by a person in the Top Down affected group		Calculated using $z_t^{SC}, z_t^{SA}, \sigma_s^S, M_s^S$
$\gamma_s^N(z_t^N)$	proportional decrease in s experienced by a person not affected by Top Down interventions.		Calculated using $z_t^N, \sigma_\lambda^N, M_\lambda^N$
$\gamma_\lambda^N(z_t^N)$	proportional decrease in λ experienced by a person not affected by Top Down interventions.		Calculated using z_t^N, σ_s^N, M_s^N
$M_\lambda^S; M_s^S; M_\lambda^N; M_s^N$	Maximum intervention effectiveness		Fitted parameter
$\sigma_\lambda^S; \sigma_s^S; \sigma_\lambda^N; \sigma_s^N$	Responsiveness to NPI		Fitted parameter
T_a	Time when Top Down interventions are first implemented		Decision Variable
B	Budget for Bottom Up interventions		Fitted Parameter

Appendix B: Sources for parameter values

Model Parameter	Value	Source
μ	11.06	Chapter 5 Appendix C. Region 5 estimate.
$s_{vaccinated}$	.02	Chapter 5 Appendix C
$s_{unvaccinated}$	.1	Chapter 5 Appendix C
L	5	Estimated Indiana school closings in 2010 gathered from news reports. [**13, 14, 2**]
F_a	.248	Proportion of people under 18 years old in the Indiana population [**15**].
T_v	24	Information from the CDC, also mentioned in Chapter 5 and [**1**].
N	6,483,802	Indiana Population [**15**].
X	.5	Least Squares parameter fit to epidemic curve data.
B	.20	
β	.5	
σ	.2	

Chapter VII: Conclusions

Over the last century, pandemic influenza has proven to be a real threat with potentially catastrophic results. Deadly viruses affecting most of the world have been occurring since the first 1918 "Spanish Flu" pandemic. However, it is only recently with the 2009-2010 H1N1 outbreak, that the world experienced a pandemic event during which authorities had the resources and tools to track the disease as it spread throughout the world. For the first time, authorities have also been able to keep records of the kinds of interventions that were used to try to contain the outbreaks. The statistics and analysis from that outbreak will allow worldwide organizations such as the CDC, WHO and local healthcare authorities to use the lessons learned and improve their strategies in dealing with pandemic influenza. The operations research community has the opportunity to contribute to this goal through applying modeling and optimization framework to decision-making in the field of epidemiology.

1. Thesis summary

In this work we have focused on both theoretical and applied aspects of pandemic influenza modeling. In Chapter 2, we have introduced a generalized discrete time model of influenza spread in a single community. In essence, the model in Chapter 2 creates an infinite number of classes. The main advantage of the generalized model is its analytical tractability. Having a small number of equations allows us to manipulate these equations to draw analytical conclusions about the state of the epidemic at different points in time. We show that population characteristics change significantly as the outbreak progresses in a community. In particular, people with the highest λ and s values tend to get infected early and leave the susceptible population before others. This makes the people with high susceptibility and activity levels the "drivers" of the outbreak. They are potentially high-priority vaccination strategy targets.

In Chapter 3, we build on this model to discuss vaccination program planning in a single community. We use the heterogeneous spread model to determine what proportion of the population would need to be vaccinated in order to contain an outbreak through reaching herd immunity, thus bringing down the effective reproductive ratio, $R(t)$, to below 1. This proportion changes with time because, as we have shown in Chapter 2, the susceptibility, infectivity, and activity level characteristics of the population change with time. In addition, we show that the number of vaccines needed to reach herd immunity fall within a range, bounded by different targeting techniques. Allocating vaccines randomly provides a rough upper bound for the number of vaccines needed to contain an epidemic. However, targeting those individuals most likely to contract or spread the disease will make vaccines more effective in curbing spread, and thus lower the total number of vaccines needed. The most effective strategy we have found is to vaccinate those individuals with the highest values of the parameter $w = \lambda^2 rs$ where λ, r and s represent a person's activity, infectivity, and susceptibility levels respectively. We note, however, that such a targeting strategy would be difficult to implement in practice, as a person's susceptibility and infectivity levels are difficult to measure objectively.

In the next few chapters of this thesis we discuss the lessons learned from the 2009-2010 H1N1 pandemic. This pandemic was the first worldwide influenza event that occurred at time when the researchers had the resources to keep track of epidemic curves and intervention measures. Using data from the CDC and 11 single US states, we were able to evaluate the effectiveness of the vaccines shipped to different US states in 2009 using a "quick and dirty" homogeneous influenza spread model. The key conclusions from this analysis are mostly intuitive:

- High numbers of vaccines actually administered to patients imply high numbers of infections averted through vaccination.
- With respect to the peak of the infection in a region, early vaccines are more effective than late vaccines.
- Vaccines delivered to a community early are more likely to be administered to patients due to higher demand levels.

These observations motivate us to develop a dynamic vaccine allocation strategy in Chapter 5 using our heterogeneous influenza spread model. The current strategy by the CDC, allocating influenza vaccines proportionally to the population of each state, ignores the potential benefits from allocating more vaccines to those states that we predict will experience the highest marginal benefit from extra doses of vaccines shipped. The switching algorithm we developed takes into account information from all constituent regions to maximize the total effect of vaccines in decreasing virus spread. We hope that our methodology will be particularly helpful to governing bodies on all levels of society. Our analysis of the CDC's decision algorithm for US states may avert up to 30% more infections than the pro-rata heuristic in a scenario similar to the recent H1N1 pandemic. Local officials may also benefit from this methodology on a state or city level.

In the final chapter, we perform analysis of non-pharmaceutical interventions on a local level. While vaccines provide permanent immunity from a virus, they must be manufactured over a long period of time, during which local communities have to contain the epidemic using public awareness campaigns, along with school and business closures. We develop methodology for using limited non-pharmaceutical resources with highest effectiveness in preparation for vaccine arrivals. Together, Chapters 4-6 of this thesis are meant to cover local and national strategies for interventions meant to mitigate the effects of a large-scale influenza event.

2. Contributions and policy implications

One of the main contributions of our work is the introduction of the generalized model for influenza spread in a heterogeneous population.

2.1 Generalized model for heterogeneous populations

The model we introduced in Chapter 2 eliminates the need for arbitrary class distinctions used in standard heterogeneous SIR models. Instead of segregating the population into a finite number of groups in which all people behave identically, our model essentially uses an infinite number of classes by introducing a joint probability distribution over the activity, infectivity, and susceptibility levels of individuals in the population. This probability distribution more realistically approximates the true characteristics of a population at any given time than class-based modeling. In addition to being more realistic, this fact that this model is based on discrete Bayesian updating steps improves the analytical tractability over standard SIR models due to the small number of equations needed to describe the state of the epidemic in the population at any point.

The standard modeling techniques for describe a heterogeneous population using a compartmental SIR model involves a growing number of coupled linear differential equations. For comparison, we give examples of the two modeling approaches in Table 1. The set of equations on the left shows a typical SIR influenza transmission model involving just two compartments [1] while the set of equations on the right

reproduces equations (3)-(6) from Chapter 2. The second set of equations includes the entire spectrum of activity, infectivity and susceptibility levels.

$$\frac{dS_1}{dt} = B_1(N_1(t))N_1(t) - (\mu_1 + d_1)S_1(t) - \lambda_1(t)S_1(t) + d_2S_2(t),$$
$$\frac{dS_2}{dt} = B_2(N_2(t))N_2(t) - (\mu_2 + d_2)S_2(t) - \lambda_2(t)S_2(t) + d_1S_1(t),$$
$$\frac{\partial I_1}{\partial t} + \frac{\partial I_1}{\partial a} = -(\mu_1 + r_1 + b_1)I_1(a,t) + b_2I_2(a,t), \quad 0 < a \leq \tau,$$
$$\frac{\partial I_2}{\partial t} + \frac{\partial I_2}{\partial a} = -(\mu_2 + r_2 + b_2)I_2(a,t) + b_1I_1(a,t), \quad 0 < a \leq \tau,$$
$$\frac{dR_1}{dt} = r_1 \int_0^\tau I_1(a,t)da + I_1(\tau,t) - (\mu_1 + c_1)R_1(t) + c_2R_2(t),$$
$$\frac{dR_2}{dt} = r_2 \int_0^\tau I_2(a,t)da + I_2(\tau,t) - (\mu_2 + c_2)R_2(t) + c_1R_1(t),$$
$$\lambda_i(t) = k_i \int_0^\tau I_i(a,t)da,$$
$$N_i(t) = S_i(t) + R_i(t) + \int_0^\tau I_i(a,t)da,$$
$$I_i(0,t) = \lambda_i(t)S_i(t), \quad i = 1,2,$$

$$N_{t+1}^I = N_t^S \int_0^1 \int_0^1 \int_0^\infty p(I|\lambda,s) * f_t^S(\lambda,r,s)d\lambda\, dr\, ds$$
$$N_{t+1}^S = N_t^S - N_{t+1}^I$$
$$f_{t+1}^I(\lambda,r,s) = \frac{p_t(I|\lambda,s) * f_t^S(\lambda,r,s)}{\int_0^1 \int_0^1 \int_0^\infty p_t(I|\lambda,s) * f_t^S(\lambda,r,s)d\lambda\, dr\, ds}$$
$$f_{t+1}^I(\lambda,r,s) = \frac{(1 - p_t(I|\lambda,s)) * f_t^S(\lambda,r,s)}{\int_0^1 \int_0^1 \int_0^\infty (1 - p_t(I|\lambda,s)) * f_t^S(\lambda,r,s)d\lambda\, dr\, ds}$$

Table 1: Two compartment SIR model [1] compared with the continuous attribute heterogeneous spread model form Chapter 2.

The relative simplicity of these equations allows us to gain insights about the population during the epidemic and to prove some general results rigorously. We encourage this kind of approach in a field where analytical proofs are not often attempted due to the fact that epidemic curves cannot be described through closed-form equations.

We believe there are three key policy implications to be drawn from this kind of modeling: planning and ordering vaccines before they are available in bulk; non-pharmaceutical interventions in preparation for vaccines before they become available; and dynamic vaccine allocation during an outbreak.

2.2 Non-pharmaceutical interventions

The biggest problem with vaccines in a pandemic is that they cannot be prepared well ahead of time. Once the virus is identified, the world is forced to wait up to six months (in the US) for bulk quantities of vaccines to be shipped and additional quantities follow after that. By that time, the virus is likely to start propagating in many places. A single region, such as a US state, will likely be faced with the problem of stopping the spread of the infection by using non-pharmaceutical interventions. Our research shows that just closing schools is unlikely to stop the outbreak altogether. However, school and business closures in conjunction with encouraging the public to use good hygiene practices and telecommute to work should tide the region over until vaccines arrive. In some cases it might even stop the outbreak altogether. Slowing the growth of the epidemic curve goes a long way in making sure that once vaccines arrive the virus is eradicated from the population quickly, without any extra burden on the healthcare infrastructure. Our methodology should be tailored to the budget constraints and types of resources available in specific regions. It will provide guidelines for how to use different resources most efficiently in order to lower and delay the peak of the epidemic curve in preparation for vaccine arrival.

2.3 Vaccine order planning

Once the threat of an outbreak has been established, healthcare officials on national levels must place orders for vaccines to be developed and shipped to their constituents. The important decision is to choose

how many vaccines should be ordered so as to avoid the economic strain of ordering too much vaccine and the healthcare strain of ordering too little. The first step in deciding how many vaccines will be necessary for some population is to choose the value of the threshold parameter $R(t)$ that the healthcare officials would like to achieve. $R(t) = 1$ corresponds to herd immunity, the commonly accepted threshold, but lower values of $R(t)$ would ensure a faster decline of the outbreak in the population. The methodology in Chapter 3 provides a way to find a range for the proportion of the population that would need to be vaccinated to achieve this threshold. In our examples, even the upper bound for the amount of vaccine needed amounted to 33% of the population in a pandemic with $R_0 = 1.5$. While the R_0 for the H1N1 virus was lower [**2, 3**], some countries ordered enough vaccine to cover as much as 100% of the population [**4, 5, 6**]. Careful analysis of threshold parameters and vaccine strategies should prevent such gross overestimations of necessary vaccines. We also advocate vaccinating those members of the population who contribute the most to spreading the epidemic so that limited available doses have the most impact when administered to patients.

2.4 Real time vaccine allocation

A governing body that faces the decision of vaccine allocation over multiple regions must take into account a variety of factors such as vaccine effectiveness, fairness, and efficiency when allocating limited resources. In the US, the CDC's current strategy of allocating vaccines on a pro-rata basis achieves the goal of fairness. Unfortunately, equal distribution of vaccines does not necessarily use these vaccines in an efficient manner. In Chapter 5 we analyze in detail other, potentially more effective heuristics for vaccine allocation. The key to being able to use this methodology is to maintain a degree of equity between states so that no region gets significantly more vaccine than any other, especially when the demand for vaccines is high in both regions. This can be accomplished through accurate data gathering accompanied by precise parameter estimation and feasible equity constraints on the solution. In our testing framework introduction of the adaptive risk parameter ensured that vaccines were allocated effectively while all regions received reasonable amounts of vaccine. In the next section we discuss the future work that should enable our methodology to be implemented on a national level.

3. Future work

3.1 Modeling research

In the first three chapters of this work we focused on developing heterogeneous model of influenza spread in a community. The model in Chapter 2 provides the analytic tractability to draw insights about the dynamics of an epidemic spread. The theorems we establish in this chapter are just one important application of this model. We hope that further work will be done in discovering the role of heterogeneity on the dynamics of influenza epidemics. We note, at the end of Chapter 2, that work has been done on estimating the distribution of expected contact rates, $E(\lambda)$, in some European populations [**7, 8**]. This work makes modeling epidemic spread in a region with accurate estimates of contact rate heterogeneity possible. The estimation of infectivity and susceptibility parameters r and s, however, is more difficult. A person's relative susceptibility is difficult to quantify. Doing so will most likely include expert opinions as well as hospital studies that have not been explored so far.

Some recent literature does study the effect of interventions like hand-washing on decreasing the probability of virus spread [**9**]. However, the absolute values of r and s for individuals have still not been

determined. We hope that further research, such as the recent work done by Everitt et al [**10**] on the effect of genes on susceptibility and mortality will offer some methods for estimating these parameters.

3.2 Dynamic vaccine distribution tools

Another tool that has the potential to significantly decrease the effect of a pandemic is the methodology developed in Chapter 5 of this work. If the CDC takes into account the state of the epidemic in each of its constituent regions when allocating available vaccines, those vaccines may be used more effectively than when allocated using the pro-rata heuristic. There are still some obstacles to overcome before this methodology may be used in a real life scenario. The most important issue that needs to be explored is data collection. In our work, we have mostly used clear data provided to us by the CDC and individual states with the benefit of hindsight. Unfortunately in real time, cases are often underreported [**11, 12, 13**]. It is difficult to make accurate estimations of the necessary transmissibility parameters such as R_0 which are crucial to dynamic vaccine allocation described in Chapter 5 in real time. Future research should focus on gaining access to statistically reliable data that could be used to estimate parameters as the outbreak is developing. We are optimistic, however, that such estimates are possible. Estimates of R_0 for the 2009-2010 H1N1 pandemic surfaced as early as May 2009 [**12**]. Once methodology for early parameter estimation is established, software packages should be developed to deliver this tool to decision makers on national and regional levels.

The contributions of this work span decision-making on local as well as national levels. When applying this thesis to real situations, we must acknowledge that the purpose of this work is not to give exact solutions to decision-makers' problems. Instead, we hope that this thesis will guide healthcare officials to insights into strategies that will help decrease the devastating effects of pandemic influenza on the population and the economy.

References

1. White M, Zhao X. Threshold dynamics in a time-delayed epidemic model with dispersal. *Mathematical Biosciences*. January 2009;218:121-129.
2. Towers S, Feng Z. Pandemic H1N1 influenza: predicting the course of a pandemic and assessing the efficacy of the planned vaccination programme in the United States. *Euro Surveill*. 2009;13(41).
3. White L, Wallinga J, Finelli L, Reed C. USA, Estimation of the reproductive number and the serial interval in early phase of the 2009 influenza A/H1N1 pandemic in the. *Influenza and Other Respiratory Viruses*. September 2009;3(6):267-276.
4. Bone A, Guthmann J, Nicolau J, Levy-Bruhl D. Population and risk group uptake of H1N1 influenza vaccine in mainland France 2009-2010: Results of a national vaccination campaign. *Vaccine*. November 2010;28(51):8157-8161.
5. Balmer C. France wants to sell millions of surplus flu shots. *Reuters*. January 2010.
6. Dutchnews.nl. Netherlands tries to sell excess flu vaccine. *Dutchnews.nl*. January 2010:2010.
7. Fu Y. Measuring personal networks with daily contacts: A single-item survey question and the contact diary. *Soc. Networks*. 2005;27:169-186.
8. Mossong J, Hens N, Jit M, Beutels P, Auranen K. Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS Med*. 2008.
9. Finkelstein S, Prakash S, Nigmatulina K, McDevitt J, Larson R. A home toolkit for primary prevention of influenza by individuals and families. *Disaster Med Public Health Prep*. Dec 2011;5(4):266-271.
10. Everitt A, Clare S, Pertel T, John S, et al e. IFITM3 restricts the morbidity and mortality associated with influenza. *Nature*. March 2012;484:519-523.
11. Lipsitch M, Lajous M, O'Hagan J, T C, al e. Use of cumulative incidence of novel influenza A/H1N1 in foreign travelers to estimate lower bounds on cumulative incidence in Mexico. *PLoS One*. July 2009;4(9).
12. Fraser C, Donnelly CCS, al e. Pandemic potential of a strain of influenza A (H1N1): early findings. *Science*. 2009;1557-1561:324.
13. Centers for Disease Control and Prevention. Underreporting of 2009 H1N1 influenza cases. Available at: http://www2c.cdc.gov/podcasts/media/pdf/EID_12-09_FluEstimates.pdf. Accessed June 10, 2011.