Computational and Analytical Mathematical Techniques for Modeling Heterogeneity

by

Benjamin Morin

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved February 2012 by the
Graduate Supervisory Committee:

Carlos Castillo-Chavez, Co-Chair
David Hiebeler, Co-Chair
Daniel Hruschka
Sergei Suslov

ARIZONA STATE UNIVERSITY

May 2012

ABSTRACT

This dissertation is intended to tie together a body of work which utilizes a variety of methods to study applied mathematical models involving heterogeneity often omitted with classical modeling techniques. I posit three cogent classifications of heterogeneity: physiological, behavioral, and local (specifically connectivity in this work). I consider physiological heterogeneity using the method of transport equations to study heterogeneous susceptibility to diseases in open populations (those with births and deaths). I then present three separate models of behavioral heterogeneity. An SIS/SAS model of gonorrhea transmission in a population of highly active men-who-have-sex-with-men (MSM) is presented to study the impact of safe behavior (prevention and self-awareness) on the prevalence of this endemic disease. Behavior is modeled in this examples via static parameters describing consistent condom use and frequency of STD testing.

In an example of behavioral heterogeneity, in the absence of underlying dynamics, I present a generalization to "test theory without an answer key" (also known as cultural consensus modeling or CCM). CCM is commonly used to study the distribution of cultural knowledge within a population. The generalized framework presented allows for selecting the best model among various extensions of CCM: multiple subcultures, estimating the degree to which individuals guess yes, and making competence homogenous in the population. This permits model selection based on the principle of information criteria. The third behaviorally heterogeneous model studies adaptive behavioral response based on epidemiological-economic theory within an *SIR* epidemic setting. Theorems used to analyze the stability of such models with a generalized, nonlinear incidence structure are adapted and applied to the case of standard incidence and adaptive incidence.

As an example of study in spatial heterogeneity I provide an explicit solution to a generalization of the continuous time approximation of the Albert-Barabási scale-free network algorithm. The solution is found by recursively solving the differential equations via integrating factors, identifying a pattern for the coefficients and then proving this observed pattern is consistent using induction. An application to disease dynamics on such evolving structures is then studied.

DEDICATION

For Lydia & Ari: with patience and love I've been allowed to keep you near.

My parents Richard and Susan Morin: you taught me what it means to work hard for what you want.

My grandfather, Pépé Bob: never has there been someone so willing to sing my praises so loudly, often, and to so many people.

ACKNOWLEDGEMENTS

There are many people without whom this work would have never been completed, and perhaps without their influence I would have never pursued a career in mathematics. My first experiences with Dr. Sharon Crook at the University of Maine, shaped an interest in dynamical systems that was continued while writing my undergraduate thesis with Dr. Robert Franzosa. Under their tutelage I learned an appreciation for understanding what the mathematics meant in the context of an application, not just the elegance and formality. It was with Dr. David Hiebeler where I learned that mathematics could be fun. The time I spent with him focused on asking the right questions prepared me for a future in applied mathematical research. And finally, Dr. Carlos Castillo-Chavez was the one who really taught me what it means to go from being a good "technician" to a good researcher.

Dr. Karl Hadeler has been nothing if not the perfect sounding board for my investigations bordering on the purer side of our discipline. His critical mind and willingness to "work it out", as well as Dr. Sergei Suslov's love of discovery, was a boon in finding the solutions in Chapter 6. Dr. Fred Brauer serves as the inspiration for much of the work with calculations of thresholds in epidemic models (as well as motivating the comparison of final epidemic sizes to answer his "worst case scenario" question). The way Dr. Brauer seems to make a game out of making innumerable changes and generalizations to a model is a challenge I embrace. My fr[ei]nds Joseph Abbruscato and Challie Facemire~~, without whom I'd drink coffee alone,~~ are due a specific thanks for editing a Few chapters of this thesis for proper use of the written word. ~~using words good.~~ I've always been one to want a patient ear and I wish to thank them here: doctors Faina Berezovskaya, Georgy Karev, Christopher Kribs, Sergei Suslov, Daniel Hruschka, Eli Fenichel, David Murillo, Fabio Sanchez and Nicolas Lanchier as well as my colleagues Oscar Petterson, Anarina Murillo, Ilyssa Summer, Maytee Cruz, Emmanuel Morales, and Kehinde "Kenny" Salau. Finally, for their faith in my

abilities as an applied, computational scientist I again acknowledge the two men without whom I would not be doing something I love: doctors David E. Hiebeler and Carlos Castillo-Chavez .

TABLE OF CONTENTS

Chapter 1

Introduction

One of the most commonly used tools in the mathematical modeling of natural phenomena is the differential equation system. Differential equations (i.e., ordinary, partial, delay, integro-, etc...) have been applied to many applied problems: ecological persistence, birth/death processes, and disease transmission to name a few. Critical to the success of these models is the intuitive construction and proliferation of quantitative and qualitative techniques for their analysis.

The applicability of differential models to the natural sciences may be summarized with two criteria: estimability of model parameters and correctness of underlying mechanism. The former will not be discussed at any length in this work, but proper treatments for various models may be found [23, 45, 62, 79, 80, 118, 121, 122]. The underlying mechanism of a process is often where particular differential models are weakest (e.g., assumptions of large numbers, non-explicit treatments of space, and, most relevantly to this study, homogeneity). Heterogeneity within differential models may be included a number of ways: increasing the number of states/compartments, stochastic modeling, or distributed parameterization.

This work does not aim at being a comprehensive overview of heterogeneity in models. Rather, I pose 3 meaningful classifications of heterogeneity and then outline examples of how such heterogeneity may be included in a model. I posit that there are two classifications of heterogeneity concerned with properties specific to an individual unit: physiological and behavioral. The third heterogeneity is a *local* measure which involves the individual's placement within its environment. Any of these heterogeneities may also be static (unchanging) or dynamic.

Not all individuals respond to biological stimuli in the same way. Individuals who differ physiologically may have very different, or surprisingly similar, reactions to phenomena (such as exposure to a disease). In the biological sense this heterogeneity is akin to the physiological characteristics of the individual in question which give rise to differential response to stimuli (age, gender, etc...). With all else equal, a model characterized by *physiological heterogeneity* draws a distinction in how individual responses are influenced by factors omitting behavior.

Similarly, all individuals do not proceed through their day, or respond to social stimuli, in the same way. An activity regime for an individual may be broken up into two parts: the type of activity and the frequency of the activity. For instance an individual may be at work and participate in a few contacts a day with a large number of people and then return home to make a large number of contacts with very few people (e.g., spending on average 20 minutes a day with each of 10 coworkers but spending 3 hours a day with 2 friends).

In compartmental epidemic models, *behavioral heterogeneity* is often indistinguishable from physiological heterogeneity. For example, if there is a population with two different forces of infections it could be due to 1) one of the groups being naturally more susceptible to the disease, 2) a group comprising a population where a co-infection decreases, or decreases, the body's ability to resist similar infection, or 3) one group participating in many more contacts than the others. When modeled simply via parameterization (e.g., $c_1\beta_1 + c_2\beta_2$ where c_i is a contact number and β_i is the probability of infection for individuals in group i) the effect of the heterogeneity is obfuscated in an estimability sense. An increase in the number of compartments may increase the ability to observe the effect of the heterogeneity, but may lead to less tractable models (i.e., multiple sex, pair forming, core-group models).

Individuals who make the same number of contacts per unit-time may make these contacts over very different structures (sex workers versus a highly sexually active,

monogamous couple). In this formulation of *local heterogeneity* it is important to study the "neighborhood" structure of individuals and to investigate the effects this has on the course of dynamics. The heterogeneity here may often be a result of the demographics of individuals and activity type modeled, as opposed to a fundamental property of the individuals. Additionally, the spatial location of an individual on their landscape can have a large impact on the possible dynamics that individuals may participate in (i.e., Are you standing in or near the fire?).

For a particular model it is important to determine what type(s) of heterogeneity is essentially considered[1]. It may be the case that more than one type of heterogeneity is present in a model, or that a single type manifests itself through more than one mechanism. Sociophysiology is a study on the concomitant relationship between physiology and social behavior; i.e. group dynamics influenced by physiological metrics. In these situations it may be most straightforward to consider Physiological-Behavioral Heterogeneity. Should activity type also dictate the structure of contact networks (e.g. the connectivity heterogeneity is a secondary effect) then a Behavioral-Local Heterogeneous model using techniques such as "meta-networks", with differential node type and connectivity, may be applicable.

This dissertation is divided into three Parts. In the first, Physiological Heterogeneity, I review some results on the application of distributed parameters to disease models while utilizing a representation theorem to reduce the effective dimensionality of the problem to that of a *transport system*[2]. The result is not an approximation and thus the dimensionality is not *truly* reduced, but a significant ability to perform computation/analysis of the model is gained. The qualitative equivalence between the

[1]Even in the event that the effects of physiological and behavioral heterogeneity are indistinguishable, it is important to make the type heterogeneity explicitly known so that the mechanisms are clearly understood.

[2]The work presented in this chapter follows from a previous collaboration with Georgy Karev and Irina Kareva (on avoiding the tragedy of the commons through taxing overconsumers and subsidizing underconsumers) and discussions with Artem Novozhilov and Carlos Castillo-Chavez.

transport system and the undistributed formulation of the model is demonstrated for the basic *SIR* model and an *SIR* model with distributed susceptibility. Using a monotonicity argument I make a worst case scenario assertion with regards to the final epidemic sizes for the models. The weakness of the representation theorem in working with more complicated models (models with demographics, recovery without immunity, etc...) is highlighted with respect to "blue-sky" births. In these examples the nature of the non-autonomous parameter specification cannot be tractably investigated and thus 1) the nature of the transient behavior cannot be confidently studied and 2) the asymptotic equivalence cannot be guaranteed. Two situations of open population *SIR* models (with inheritance of disease state and sterility brought on by infection) where the theory works are discussed and analyzed.

Part 2, Behavioral Heterogeneity, contains three separate models. The selection of these models is done to cogently distinguish between static and adaptive (dynamic) behavioral heterogeneity. Chapter 3 is a paper[3] analyzing the impacts of static behavioral effects on the dynamics of gonorrhea transmission in a single-sex population. Chapter 4 is another example of static behavioral heterogeneity where there is no underlying dynamic process. The question of model selection in test theory without an answer key is discussed in relation to its application to a population with distinct cultures[4]. Chapter 5 presents a prototype of incorporating positive, adaptive behavior into a standard epidemiological model using concepts from the economics of risk aversion[5]. Adaptive behavior is a subset of more general dynamic behavior.

Local heterogeneity is covered in Part 3. Explicit spatial models (e.g., PDE's) are omitted in favor for connectivity (network) models. Chapter 6 presents a new explicit solution to a generalization of the scale-free paradigm popularized by Albert and

[3] Authorship of this paper, published in the Journal of Theoretical Biology, is B.R. Morin, L. Medina-Rios (an undergraduate student at the time of research), E.T. Camacho, and C. Castillo-Chavez.

[4] Authorship of this manuscript is Benjamin R. Morin and Daniel Hruschka.

[5] Authorship of this submitted manuscript is Benjamin R. Morin, Eli P. Fenichel and Carlos Castillo-Chavez.

Barabási, and uses this model to calculate the epidemic threshold for an *SIR* model. This calculation demonstrates that with reasonable biological parameters the basic reproduction number of such a disease on a scale-free-like network is always greater than one.

Part I

Physiological Heterogeneity

Chapter 2

A Transport Equation Approach to Variable Susceptibility

Epidemic/compartmental models typically have the population broken up into a number of compartments that describe the individual disease state of its members. Along with the compartments describing the states of individuals, there are a set of biologically motivated parameters describing the transitionary flow between compartments. It is commonly the case that these parameters are described as the average quantity for a given population (e.g., the average duration of infection). The *SIR* model is one such example that has had a great deal of intellectual effort put forth to its understanding
([5, 6, 8, 9, 11, 19, 21, 29, 31, 35, 36, 38, 40, 41, 50, 58, 68, 71, 75, 78, 86, 98, 113, 134] to name but a few). Individuals in this model are either susceptible to a disease, S, infected/infectious, I, or have recovered and are now immune to reinfection, R. This formulation involves the rate that a contact between a susceptible and an infectious individual results in a new infection, β, and a recovery rate, γ. Furthermore, a standard assumption to the contact structure is that individuals make contacts at random with the entire population. This results in the system

$$\begin{aligned} \dot{S(t)} &= -\beta S(t)\frac{I(t)}{N}, \\ \dot{I(t)} &= \beta S(t)\frac{I(t)}{N} - \gamma I(t), \\ \dot{R(t)} &= \gamma I(t), \end{aligned} \tag{2.1}$$

where $\dot{X(t)} = \frac{dX(t)}{dt}$. While the explicit solution is not available, there are two quantities often calculated with respect to the behavior of this model: the basic reproduction number $\mathfrak{R}_0$ and the final epidemic size relation. For this simple model we have $\mathfrak{R}_0 = \frac{\beta}{\gamma}$ and the final epidemic size relation of $S_\infty = S(0)e^{-\frac{\beta}{\gamma}\left(1-\frac{S_\infty}{N}\right)}$.

The goal of this chapter is to invoke the supposition that the population is not homogenous in its parameters (i.e., individuals may be further stratified by differential susceptibility, infectivity, and/or recovery). General theory on the use of transport equations/variables for models of the form

$$\begin{aligned} \dot{X}(t,w) &= X(t,w)F(X(t,w;\vec{\theta}),\vec{Y(t)},t;\vec{\theta}), \\ \dot{\vec{Y(t)}} &= G(X(t,w),\vec{Y(t)},t;\vec{\theta}), \end{aligned} \tag{2.2}$$

has been done by Karev [81, 82], and was used by Novozhilov, [109], on the dynamics within a closed population *SIR* setting. Novozhilov's results are somewhat replicated here in Section 2.1 as a means to warm up to the method. While the equivalence of asymptotic behavior for distributed models of the form given in System (2.2) and their undistributed counter parts was proven by Karev, Subsection 2.1 demonstrates the technique for proving it for this specific case[1]. The representation theory explained in [81, 82] excludes models that exhibit *blue-sky births* (i.e., entries into the distributed class at a rate not proportional to the class itself). I demonstrate that the representation may be carried through in some open population models, and the final outcome is a finite dimensional, non-autonomous system. The difficulty in applying the equivalence theory arrises in studying the non-autonomous parameter's evolution over time.

2.1 Differential Susceptibility *SIR* With a Closed Population

I start here with an introduction of variability in susceptibility[2] modeled via a parameter w and a resultant value of $\beta(w)$. Assume that $\beta(w) \geq 0$ (for biological feasibility) and is finite (in order to ensure all populations are finite in finite time), and denote the susceptible individuals with a particular susceptibility of w via $S(t,w)$. The resulting system takes on the form

$$\begin{aligned} \dot{S(t,w)} &= -\beta(w)S(t,w)\frac{I(t)}{N}, \\ \dot{I(t)} &= \int \beta(w)S(t,w)dw\frac{I(t)}{N} - \gamma I(t), \end{aligned} \tag{2.3}$$

[1] This dynamic equivalence is similar to that of the age structured model, see [29].

[2] The full derivation of this model may be found in [109].

with $R(t)$ omitted due to the constant population size. Based on the representation theory of Karev [81–83], one introduces a transport variable $q(\dot{t}) = -\frac{I(t)}{N}$ and arrives at the solution

$$S(t,w) = S(0,w)e^{\beta(w)q(t)}$$

through separation of variables for the susceptible class. Note that $S(t) = \int S(t,w)dw$ is the total susceptible population which satisfies

$$S(\dot{t}) = -\overline{\beta(t)}S(t)\frac{I(t)}{N},$$

where

$$\begin{aligned}
\overline{\beta(t)} &= \frac{\int \beta(w)S(t,w)dw}{\int S(t,w)dw}, \\
&= \frac{\int \beta(w)S(0,w)e^{\beta(w)q(t)}dw}{\int S(0,w)e^{\beta(w)q(t)}dw} \\
&= \frac{d}{d\lambda}\left[\ln\left(M_\beta(0,\lambda)\right)\big|_{\lambda=q(t)}\right.\!.
\end{aligned}$$

$M_\beta(t,\lambda)$ is the moment-generating function of the time t density of property w within the susceptible population. The system in (2.3) may be of arbitrarily large dimension, since w may take on values along a continuum, and is now reduced to two non-autonomous differential equations and an integral expression[3], or a *transport system* given by:

$$\begin{aligned}
S(\dot{t}) &= -\overline{\beta(t)}S(t)\frac{I(t)}{N}, \\
I(\dot{t}) &= \overline{\beta(t)}S(t)\frac{I(t)}{N} - \gamma I(t), \\
q(\dot{t}) &= -\frac{I(t)}{N}, \\
\overline{\beta(t)} &= \frac{\int \beta(w)S(0,w)e^{\beta(w)q(t)}dw}{\int S(0,w)e^{\beta(w)q(t)}dw}.
\end{aligned} \tag{2.4}$$

[3]What is interesting about this method is that it constructs non-autonomous differential equations which implies that one is still searching an extremely large solution space. This has shifted the continuum of state variables onto an autonomous transport variable ODE and a time dependent parameter. However, with the introduction of this transport variable, coupled with the ability to solve $S(t,w)$ in terms of it and initial data, the integral expressions are "solvable" numerically and may be represented via moment-generating functions of the initial data.

It may be assumed that the initial condition for the distribution of $S(0,w)$ and the values for $\beta(w)$ are known. The system (2.4) may be recast where the incidence is a nonlinear function. Denote the moment-generating function for the distribution describing the selection of an arbitrary susceptible individual with susceptibility $\beta(w)$ via

$$M_\beta(0,q(t)) = \int \frac{S(0,w)}{S(0)} e^{\beta(w)q(t)} dw.$$

We may use this to further alter the representation of (2.3) by finding

$$\begin{aligned}
\frac{1}{S(t)}\frac{d}{dt}S(t) &= \overline{\beta(t)}\frac{d}{dt}q(t), \\
&= \frac{d}{d\lambda}\left[\ln\left(M_\beta(0,\lambda)\right)\big|_{\lambda=q(t)}\right]\frac{d}{dt}q(t), \\
\frac{d\ln(S(t))}{dt} &= \frac{d}{dt}\ln(M_\beta(0,q(t))), \\
\frac{S(t)}{S(0)} &= M_\beta(0,q(t)), \\
q(t) &= M_\beta^{-1}(0,S(t)/S(0)).
\end{aligned}$$

This allows one to write

$$\dot{S(t)} = -\frac{d}{d\lambda}M_\beta(0,\lambda)|_{\lambda=M_\beta^{-1}(0,S(t)/S(0))}S(0)\frac{I(t)}{N}, \tag{2.5}$$

and by the inverse function theorem (i.e., $(f^{-1})'(b) = \frac{1}{f'(a)}$ where $b = f(a)$) results in

$$\dot{S(t)} = -\left[\frac{d}{d\lambda}M_\beta^{-1}(0,\lambda)|_{\lambda=S(t)/S(0)}\right]^{-1}S(0)\frac{I(t)}{N} = -h(S(t))\frac{I(t)}{N}, \tag{2.6}$$

where $h(S(t))$ is a non-linear function of $S(t)$.

This formulation is particularly useful for the calculation of final epidemic size relation. Calculating $\frac{dS(t)}{dR(t)}$ results in

$$\begin{aligned}
dS(t) &= -\frac{S(0)}{N\gamma}\left[\frac{d}{d\lambda}M_\beta^{-1}(0,\lambda)|_{\lambda=S(t)/S(0)}\right]^{-1} dR(t),\\
-\frac{1}{N\gamma}dR(t) &= \frac{d}{d\lambda}M_\beta^{-1}(0,\lambda)|_{\lambda=S(t)/S(0)}\frac{dS(t)}{S(0)},\\
-\frac{R_\infty - R(0)}{N\gamma} &= \int_1^{S_\infty/S(0)} dM_\beta^{-1}(0,\lambda),\\
-\frac{N-S_\infty}{N\gamma} &= M_\beta^{-1}(0,S_\infty/S(0)),\\
S_\infty &= S(0)M_\beta\left(0,\frac{S_\infty - N}{N\gamma}\right).
\end{aligned}$$

Furthermore, as a straightforward application of the results in [25], *SIR* models with this nonlinear form have a basic reproduction number[4] $\mathfrak{R}_0 = \overline{\frac{\beta(0)}{\gamma}}$. With these threshold quantities (final epidemic size and basic reproduction number) one may address several questions: when is the beginning of the epidemic in the distributed case identical to the non distributed case, when is the final size relationships between the two models the same, and when is the qualitative behavior between the two models identical.

Equating the two basic reproduction numbers results in $\beta = \overline{\beta(0)}$. Thus, if the traditional β is chosen to be the initial mean of the distribution of $S(t,w)$ then the initial behavior of the two models is identical. Supposing that the solution to $S_\infty = S(0)e^{-\frac{\beta}{\gamma}\left(1-\frac{S_\infty}{N}\right)}$ is identical to that of the distributed problem implies this solution must satisfy

$$e^{\frac{\beta}{\gamma}\frac{R_\infty}{N}} = M_\beta\left(0,\frac{1}{\gamma}\frac{R_\infty}{N}\right),$$

where R_∞ is the limiting recovered population, supposed identical on either side of the expression. Note that the β on the left hand side is the particular value (from the classical model) while that on the right is a distributed variable. Since distributions are

[4]The limitations of such a quantity should be apparent in such a case where the infectivity is a function of time. Nevertheless, it is presented as a standard threshold computation.

uniquely identified by their Moment-generating function, we may conclude the initial distribution of w in the susceptible population must be delta, i.e.,

$$\begin{aligned} M_\beta\left(0, \frac{1}{\gamma}R_\infty\right) &= \int \frac{S(0,w)}{S(0)} e^{\frac{\beta(w)}{\gamma}\frac{R_\infty}{N}} dw, \\ &= \int \delta_{\beta(w)-\beta} e^{\frac{\beta(w)}{\gamma}\frac{R_\infty}{N}} dw, \\ &= e^{\frac{\beta}{\gamma}\frac{R_\infty}{N}}. \end{aligned}$$

We may then conclude that if the final size is identical between the two models (distributed and undistributed) then the initial epidemic behavior is identical; however, the converse is not true. One may choose any number of initial distributions such that the mean at time zero is equivalent to β. This is particularly important when estimating parameters from initial epidemic data (the initial phase of exponential growth). These estimations typically assume a delta distribution of infectivity and therefore may be used to incorrectly project final epidemic size (whether the estimation based on homogeneity considerations is over or under that of a particular distributed case is discussed later). Furthermore, Novozhilov demonstrated that if the same distribution were supposed with equilvalent mean and different variance, then there are some situations (e.g., Gamma) where the more variable population may be proven to result in a greater final epidemic size.

Dynamic Equivalence

For the undistributed model, all points of the form $(S^*, 0)$ are equilibria. Qualitatively, all values $S^* > \frac{\gamma}{\beta}N$ are unstable fixed points and all $S^* < \frac{\gamma}{\beta}N$ are stable. The fixed points for the transport system (2.4) may pose a particular challenge because the system is now non-autonomous. However, assuming a non-degenerate situation (i.e., $\overline{\beta(t)} \neq 0$), the equilibria are still of the form $(S^*, 0)$. The linearization of the distributed system gives the condition for stability as

$$S(t) < \frac{\gamma}{\overline{\beta(t)}}N.$$

For general distributions, the stability threshold may create a complicated phase space where $\overline{\beta(t)}$ forms an implicit (in time) boundary which may induce oscillations (necessarily damped) in the phase space (e.g., there may exist some times τ_1, τ_2, and τ_3 such that for $t \in [0, \tau_1)$ and $t \in (\tau_2, \tau_3)$ the stability condition is not satisfied and for $t \in (\tau_1, \tau_2)$ and $t \in (\tau_3, \infty)$ the stability condition is satisfied, causing a second peak), but one may prove that this will not occur. Define the threshold $T(t) = \frac{\gamma N}{\overline{\beta(t)}}$ and consider

$$\begin{aligned} \frac{dT(t)}{dt} &= -\frac{\gamma N}{\left(\overline{\beta(t)}\right)^2} \frac{d\overline{\beta(t)}}{dt}, \\ &= \frac{\gamma Var(\beta(t)) I(t)}{\left(\overline{\beta(t)}\right)^2}. \end{aligned} \tag{2.7}$$

By Equation (2.7) it is clear that $T(t)$ is monotonically increasing (furthermore, its slope approaches 0 as $Var(\beta(t))$ approaches 0, i.e., as $\frac{S(t,w)}{S(t)}$ approaches a delta distribution). Since $T(t)$ is monotonically increasing, the amount of the $S(t)$-axis in the phase space for which the points are stable is also increasing (non-decreasing in the event that $\frac{\gamma}{\overline{\beta(t)}} > 1$ for some $t < \infty$). Since $I(t)$ begins to decrease once it crosses $T(t)$, and because $T(t)$ is monotonically increasing, there is no way to induce an oscillation on $I(t)$ regardless of the distribution on w (i.e., once $I(t)$ decreases it may not increase). Similarly, there will be no oscillations in either $S(t)$ or $S(t,w)$.

Lemma 2.1.1 (Closed SIR Equivalence/Worst Case Distribution)**.** *Due to the monotonicity of $\overline{\beta(t)}$,the transient and asymptotic qualitative behaviors of the distributed and non-distributed SIR models are identical. Additionally, the initial behavior and final epidemic size of the two models are identical if $S(0,w) = \delta_{\beta(w)-\beta} S(0)$. Finally, the initial behavior of the two models is identical if and only if $\overline{\beta(0)} = \beta$. Also due to the monotonic decrease in $\overline{\beta(t)}$, over all distributions chosen with equivalent initial mean, the one that produces the most infection is the delta (undistributed) distribution.*

Proof. The equivalence claims are all proven in the text preceding this Lemma. To prove the worst case scenario claim observe that since $q(0) = 0$, and $I(t) \to 0$, we have that $q(t)$ monotonically decreases to some value $\eta \in (-\infty, 0)$. The derivative of $\overline{\beta(t)}$ with respect to $q(t)$ is $Var(\beta(t)) > 0$, implying $\overline{\beta(t)}$ decreases monotonically to $\varepsilon \in (0, \overline{\beta(0)})$. Note the final epidemic size calculation

$$S_\infty = S(0)e^{-\int_0^\infty \overline{\beta(u)} I(u)du}, \tag{2.8}$$

and the inequality

$$\int_0^\infty \overline{\beta(u)} I(u)du \leq \overline{\beta(0)} \int_0^\infty I(u)du = -\frac{\overline{\beta(0)}}{\gamma} \frac{(S_\infty - N)}{N}. \tag{2.9}$$

This implies that the final epidemic size for the undistributed case is minimal, given that the distribution has an equivalent initial mean susceptibility, i.e.,

$$S(0)e^{-\int_0^\infty \overline{\beta(u)} I(u)du} \geq S(0)e^{-\frac{\overline{\beta(0)}}{N\gamma}(N - S_\infty)}.$$

□

2.2 "Blue-Sky" Births & Open Populations

The most straightforward manner to "open" the population of the aforementioned *SIR* model is to suppose newborns are susceptible and each class experiences proportionate removal from the system. This results in the distributed system

$$\begin{aligned} \dot{S(t,w)} &= \Lambda N(t,w) - \beta(w)S(t,w) \int \frac{I(t,w)}{N(t)} dw - \mu S(t,w), \\ \dot{I(t,w)} &= \beta(w)S(t,w) \int \frac{I(t,w)}{N(t)} dw - (\gamma + \mu) I(t,w), \\ N(t,w) &= N(0,w)e^{(\Lambda - \mu)t}. \end{aligned} \tag{2.10}$$

However, both $I(t,w)$ and $R(t,w)$ produce members of $S(t,w)$ (and thus the ODE for $S(t,w)$ cannot be solved via separation of variables). Nevertheless, continuing as

before, let $\dot{q}(t) = -\frac{I(t)}{N(t)}$ to get

$$\begin{aligned} \dot{S(t,w)} + \left(\beta(w)\frac{I(t)}{N(t)} + \mu\right) S(t,w) &= \Lambda N(t,w), \\ \frac{dS(t,w)e^{-\beta(w)q(t)+\mu t}}{dt} &= \Lambda N(0,w)e^{\Lambda t - \beta(w)q(t)}, \\ S(t,w)e^{-\beta(w)q(t)+\mu t} - S(0,w) &= \Lambda N(0,w)\int_0^t e^{\Lambda r - \beta(w)q(r)}dr, \end{aligned}$$

and thus

$$S(t,w) = \left(\Lambda N(0,w)\int_0^t e^{\Lambda r - \beta(w)q(r)}dr + S(0,w)\right) e^{\beta(w)q(t)-\mu t}. \tag{2.11}$$

The method for solving this equation was via the integrating factor $e^{-\beta(w)q(t)+\mu t}$ as opposed to separation of variables, as in the closed population case. We may now consider the transport system

$$\begin{aligned} \dot{S(t)} &= \Lambda N(t) - \overline{\beta(t)}S(t)\frac{I(t)}{N(t)} - \mu S(t), \\ \dot{I(t)} &= \overline{\beta(t)}S(t)\frac{I(t)}{N(t)} - (\gamma+\mu)I(t), \\ N(t) &= N(0)e^{(\Lambda-\mu)t}, \\ \overline{\beta(t)} &= \frac{\Lambda\int\beta(w)P(w)e^{\beta(w)q(t)}\int_0^t e^{\Lambda r-\beta(w)q(r)}drdw + \int\beta(w)P_S(w)e^{\beta(w)q(t)}dw}{\int\left(\Lambda P(w)\int_0^t e^{\Lambda r-\beta(w)q(r)}dr + P_S(w)\right)e^{\beta(w)q(t)}dw}, \\ &= \frac{\Lambda\int_0^t e^{\Lambda r}\frac{d}{d\lambda}M_\beta(0,\lambda)|_{\lambda=q(t)-q(r)}dr + \frac{d}{d\lambda}M_{\beta|S}(0,\lambda)|_{\lambda=q(t)}}{\Lambda\int_0^t e^{\Lambda r}M_\beta(0,q(t)-q(r))dr + M_{\beta|S}(0,q(t))}, \\ \dot{q(t)} &= -\frac{I(t)}{N(t)}, \end{aligned}$$

where $P(w) = \frac{N(0,w)}{N(0)}, P_S(w) = \frac{S(0,w)}{N(0)} \approx P(w)$, and $M_{\beta|S}$ is the conditional moment generating function based on $P_S(w)$. It should be clear that the set of possible qualitative behaviors from the undistributed case[5] are the only options for the evolution of this transport system. However, the nature of $\overline{\beta(t)}$ (whether it is

[5]The undistributed system is a homogeneous system and thus, by rescaling to proportionate variables we may equate stability analysis of fixed points of the rescaled system to stability analysis of the exponential trajectories of the original, undistributed system. The disease free equilibrium (trajectory) is attracting if and only if $\beta < \gamma + \Lambda$. When the disease free state is not attracting there is an endemic equilibrium (trajectory) in the relevant phase space which is stable.

increasing, decreasing, or both) is intractable. The derivative of $\overline{\beta(t)}$ with respect to t reduces to

$$-\frac{I(t)}{N(t)}Var\left(\overline{\beta(t)}\right)+\frac{\Lambda\left(\overline{\beta(t)}-\overline{\beta_N(t)}\right)}{S(t)}\left(\mu N(t)-\overline{\beta_N(t)}\right), \tag{2.12}$$

where $\overline{\beta_N(t)}=\frac{\int\beta(w)N(0,w)dw}{N(t)}$. Note that $\overline{\beta(0)}\approx\overline{\beta_N(0)}=\overline{\beta_N(t)}$. Thus at time $t=0$ it is true that $\overline{\beta(t)}$ is decreasing. For $t>0$ the sign of $\frac{\Lambda\left(\overline{\beta(t)}-\overline{\beta_N(t)}\right)}{S(t)}\left(\mu N(t)-\overline{\beta_N(t)}\right)$ is equivalent to that of $\left(\overline{\beta(t)}-\overline{\beta_N(t)}\right)\left(\mu e^{(\Lambda-\mu)t}\left(\int N(0,w)dw\right)^2-\int\beta(w)N(0,w)dw\right)$. Given this information it is feasible that the exponential trajectory for the transport system oscillates between being attracted to the disease free trajectory and the endemic trajectory. Furthermore, if $\overline{\beta(w)}\geq 1$ for all w then the derivative of $\overline{\beta(t)}$ with respect to $q(t)$ is always positive:

$$\frac{d\overline{\beta(t)}}{dq(t)}=Var(\overline{\beta(t)})+\frac{\Lambda}{S(t)}\int N(t,w)\left(\overline{\beta(t)}-1\right)dw>0. \tag{2.13}$$

Since $q(t)$ is monotonically decreasing we may infer in this case that there exists a time $\tau<\infty$ such that for all $t\geq\tau$, $\overline{\beta(t)}<\gamma+\Lambda$. This implies that the disease will eventually "burn itself out" and the disease free trajectory will be stable.

This does not seem to have opened many analytical pathways as in the closed case, however this should be seen as a boon for numerical computation. The original system involved (in general) an infinite number of ordinary differential equations to integrate numerically, a task not easily undertaken by any computer. However, in the above system there are three ODE's to solve numerically and a quantity involving numerical integration based on initial conditions ($S(0,w),N(0,w)$, and $\beta(w)$) and the solution trajectory of $\dot{q(t)}$ up to and including the current time. A careful coding of any forward-backward numerical solver can handle this system.

Pure Inheritance

A method to circumvent the "blue-sky" births into $S(t,w)$ is to assume the malady, and immunity to it, is transferred to new borns. This inheritance mechanism is weak at

best because 1) the additions and removals to the system are not solely births and deaths in general but could be immigration and emmigration from the area in question and 2) we have to further assume the father's status confers nothing onto new-borns. With these caveats in mind, one may formulate:

$$\begin{aligned}
\dot{S(t,w)} &= \Lambda S(t,w) - \beta(w)S(t,w)\frac{I(t)}{N(t)} - \mu S(t,w),\\
\dot{I(t,w)} &= \Lambda I(t,w) + \beta(w)S(t,w)\frac{I(t)}{N(t)} - (\gamma+\mu)I(t,w),\\
\dot{R(t,w)} &= \Lambda R(t,w) + \gamma I(t,w) - \mu R(t,w),\\
N(t,w) &= N(0,w)e^{(\Lambda-\mu)t}.
\end{aligned} \tag{2.14}$$

The solution to $S(t,w)$ may then be found via separation of variables as

$$S(t,w) = S(0,w)e^{(\Lambda-\mu)t+\beta(w)q(t)},$$

with $\dot{q(t)} = -\frac{I(t)}{N(t)}$. Integrating each ODE over w gives the transport system

$$\begin{aligned}
\dot{S(t)} &= \Lambda S(t) - \overline{\beta(t)}S(t)\frac{I(t)}{N(t)} - \mu S(t),\\
\dot{I(t)} &= \Lambda I(t) + \overline{\beta(t)}S(t)\frac{I(t)}{N(t)} - (\gamma+\mu)I(t),\\
\dot{R(t)} &= \Lambda R(t) + \gamma I(t) - \mu R(t),\\
N(t) &= N(0)e^{(\Lambda-\mu)t},\\
\overline{\beta(t)} &= \frac{\int \beta(w)S(0,w)e^{\beta(w)q(t)}dw}{\int S(0,w)e^{\beta(w)q(t)}dw},\\
\dot{q(t)} &= -\frac{I(t)}{N(t)}.
\end{aligned} \tag{2.15}$$

We may recast System (2.15) into a system with proportionate variables $s(t) = \frac{S(t)}{N(t)}$, $i(t) = \frac{I(t)}{N(t)}$ and $r(t) = \frac{R(t)}{N(t)}$, each trapped within the interval $[0,1]$. The resulting autonomous system is given by

$$\begin{aligned}
\dot{s(t)} &= -\overline{\beta(t)}s(t)i(t),\\
\dot{i(t)} &= \overline{\beta(t)}s(t)i(t) - \gamma i(t),\\
\dot{r(t)} &= \gamma i(t),\\
\dot{q(t)} &= -i(t),
\end{aligned} \tag{2.16}$$

with the definition of $\overline{\beta(t)}$ left unchanged. This system exhibits the same dynamics as the closed *SIR* population transport equations in System (2.4). This implies that opening the population as in System (2.15) may not induce oscillations, where in the original open system given by (2.10) we were not able to definitively rule out oscillatory behavior (it could not be shown that $\overline{\beta(t)}$ was monotonic).

Sterilization

It is conceivable that particular infections can infer sterility on the individual. The zoonoses *Trichomoniasis*, *Salmonellosis*, and *Leptospirosis* are such infections in cows [132]. Once a heifer has been infected with these diseases the next pregnancy will result in abortion. With *Salmonellosis* and *Leptospirosis* it is unclear if future pregnancies result in abortions even if the cow shows no signs of infectiousness, but it is true that upon true recovery, after a short time spent immune to the disease, the heifer is again susceptible to infection but may conceive and calf normally until reinfected. This dynamic, similar to an *SIS* model (the immunity is so short that the rate from *R* to *S* will be disproportionately large), can be shown to be completely incompatible with the transport equation technique[6].

Papillomaviruses in sheep have both an acute and chronic stage. During the acute stage the sheep is infectious and any pregnancy during which the sheep is in the acute phase will result in abortion [111]. The passing to the chronic phase causes scarification of the fallopian tubes, as it does in humans. This scarring causes infertility in addition to making the sheep more susceptible to other STDs and STIs. While in the chronic phase the sheep is still infectious, but at a much lower level than when in the acute phase [111]. I simplify this by supposing the infections caused by

[6]I've omitted showing the calculations for *SIS* and *SIRS* models but the reentry into the susceptible class causes the distributed equations to be completely unsolvable in any meaningful way. The solution for $I(t)$ in the distributed susceptibility *SIS* model looks very similar to the solution of the non-autonomous *SIS* model [94], but it can be easily shown that the solution is both implicit (the parameters "depend" on $I(t)$) and incomplete (the parameters require that $I(t,w)$ be solved, which cannot be done).

sheep in the chronic phase is negligible and cast the dynamics into an *SIR* setting with variable susceptibility. The following model will suppose a population whose growth is naturally limited, modeled via logistic growth, and is single sex (females only). I will introduce papillomavirus into the population noting that it 1) causes no death due to infection and 2)causes permanent infertility in infectious (acute) and recovered/immune individuals (chronic). The variable susceptibility serves an amalgamation of effects that contribute to susceptibility: nutrition, infection history, cleanliness of environment, etc.... Suppose a logistic growth for the population given by

$$\dot{N(t)} = \lambda N(t) - \frac{\lambda}{K} N^2(t), \tag{2.17}$$

and rationalize the terms mechanistically as a birth process $\lambda N(t)$ and a density dependent death process $\lambda N(t)\frac{N(t)}{K}$. By introducing a sterilizing disease, and imparting differential susceptibility, one arrives at

$$\begin{aligned}
\dot{S(t,w)} &= \lambda S(t,w)\left(1-\frac{N(t)}{K}\right) - \beta(w)S(t,w)\frac{I(t)}{N(t)},\\
\dot{I(t)} &= \int \beta(w)S(t,w)dw\frac{I(t)}{N(t)} - \left(\gamma+\lambda\frac{N(t)}{K}\right)I(t),\\
\dot{R(t)} &= \gamma I(t) - \lambda R(t)\frac{N(t)}{K},\\
\dot{N(t)} &= \lambda\int S(t,w)dw - \lambda N(t)\frac{N(t)}{K}.
\end{aligned} \tag{2.18}$$

Introduce the transport variables $\dot{u(t)} = -\frac{N(t)}{K}$ and $\dot{v(t)} = -\frac{I(t)}{N(t)}$ to arrive at

$$S(t,w) = S(0,w)e^{\lambda t+u(t)+\beta(w)v(t)},$$

and thus

$$S(t) = e^{\lambda t+u(t)}\int S(0,w)e^{\beta(w)v(t)}dw.$$

Defining

$$\overline{\beta(t)} = \frac{\int \beta(w)S(0,w)e^{\beta(w)v(t)}dw}{\int S(0,w)e^{\beta(w)v(t)}dw},$$

gives the transport system

$$\begin{aligned}
\dot{S(t)} &= \lambda S(t)\left(1-\frac{N(t)}{K}\right) - \overline{\beta(t)}S(t)\frac{I(t)}{N(t)},\\
\dot{I(t)} &= \overline{\beta(t)}S(t)\frac{I(t)}{N(t)} - \left(\gamma+\lambda\frac{N(t)}{K}\right)I(t),\\
\dot{u(t)} &= -\frac{N(t)}{K},\\
\dot{v(t)} &= -\frac{I(t)}{N(t)},\\
\dot{N(t)} &= \lambda S(t) - \lambda N(t)\frac{N(t)}{K}.
\end{aligned} \tag{2.19}$$

Supposing the total population is less than K, one may rescale to state variables in $[0,1]$ and define the biologically valid domain via
$T = \{(s,i,n)|s \geq 0, i \geq 0, n \in [0,1], s+i \leq 1\}$:

$$\begin{aligned}
\dot{s}(t) &= \lambda s(t)(1-n(t)) - \overline{\beta(t)}s(t)\frac{i(t)}{n(t)},\\
\dot{i}(t) &= \overline{\beta(t)}s(t)\frac{i(t)}{n(t)} - (\gamma+\lambda n(t))i(t),\\
\dot{u}(t) &= -n(t),\\
\dot{v}(t) &= -\frac{i(t)}{n(t)},\\
\dot{n}(t) &= \lambda s(t) - \lambda n^2(t),\\
\overline{\beta(t)} &= \frac{\int \beta(w)s(0,w)e^{\beta(w)v(t)}dw}{\int s(0,w)e^{\beta(w)v(t)}dw}.
\end{aligned} \tag{2.20}$$

The qualitative behavior of the undistributed case forms the bifurcation diagram that this non-autonomous transport system now moves through (the bifurcation parameter is a function of time). I will discuss a fixed point whose coordinates involve $\overline{\beta(t)}$. This "fixed curve" is a trajectory in $\mathbb{R}^3$ and should the trajectory of the state variables, s, i, and r, come in contact with it, in the space-time sense, then their dynamics will cease for a moment. However, if $i(t) \neq 0$ then $\dot{v(t)} \neq 0$ and $\overline{\beta(t)}$ may change. The state-dynamics will then not be at equilibrium and continue. Effectively this trajectory corresponds to turning points (local minimums, maximums or inflection points) that

occur for all three states simultaneously. There are two simple fixed points for the undistributed system:

$$(s^*, i^*, r^*) = (0,0,0), (1,0,0). \tag{2.21}$$

The trivial fixed point, all states 0, is a saddle-type node (i.e., attracting down the $i(t)$ and $r(t)$ axes and repelling down the $s(t)$-axis). The disease free equilibrium, DFE, $(1,0,0)$ has eigenvalues $-\lambda, -\lambda$, and $\overline{\beta(t)} - \lambda - \gamma$. Thus *when* $\overline{\beta(t)} < \lambda + \gamma$ then the disease free equilibrium is a stable node, at other times it is a saddle-type node. Looking at the nullclines of the $s(t)$ and $i(t)$ variables reveals an intersection at $\left(\frac{(\gamma+\lambda n(t))n(t)}{\overline{\beta(t)}}, \frac{\lambda n(t)(1-n(t))}{\overline{\beta(t)}}\right)$. Treating this as a point one may find the r-coordinate via $\dot{r(t)} = \gamma i(t) - \lambda r(t) n(t) = 0$ which implies

$$r(t) = \frac{\gamma(1-n(t))}{\overline{\beta(t)}}. \tag{2.22}$$

These expressions for $s(t), i(t)$, and $r(t)$, together with $s(t) + i(t) + r(t) = n(t)$, gives that $n(t) = \frac{\gamma}{\overline{\beta(t)} - \lambda}$ along this fixed curve. This gives a final form for the time t coordinates of the "fixed point" in terms of $\overline{\beta(t)}$ as

$$(s(t), i(t), r(t)) = \left(\frac{\gamma^2}{\left(\overline{\beta(t)} - \lambda\right)^2}, \frac{\lambda\gamma\left(\overline{\beta(t)} - \lambda - \gamma\right)}{\left[\overline{\beta(t)}(\overline{\beta(t)} - \lambda)\right]^2}, \frac{\gamma\left(\overline{\beta(t)} - \lambda - \gamma\right)}{\overline{\beta(t)}\left(\overline{\beta(t)} - \lambda\right)} \right). \tag{2.23}$$

This "point", call it the endemic equilibrium trajectory, is valid biologically only when $\overline{\beta(t)} - \lambda - \gamma > 0$ (when both the trivial and disease free equilibria are saddle nodes). Eigenvalues about this "point" are $-\frac{\gamma\lambda}{\overline{\beta(t)} - \lambda}$ and

$$\frac{1}{2\overline{\beta(t)}(\overline{\beta(t)} - \lambda)^2} \left[b \pm \sqrt{b^2 - 4\overline{\beta(t)}^2 \gamma\lambda(\overline{\beta(t)} - \lambda)^2(\overline{\beta(t)} - \gamma - \lambda)} \right],$$

where $b = \lambda(\overline{\beta(t)} - \lambda)(\overline{\beta(t)} - \gamma - \lambda)(\overline{\beta(t)} - 1)$. When the endemic equilibrium trajectory is inside T, the sign of the real part of the two complicated eigenvalues is determined by the sign of $\overline{\beta(t)} - 1$ (the simple eigenvalue is negative). The real part of these eigenvalues are positive if and only if $\overline{\beta(t)} > 1$. If the endemic equilibrium

trajectory is not in T, $\overline{\beta(t)} < \gamma + \lambda$, then there is always a positive eigenvalue for the endemic equilibrium trajectory, causing it to be unstable.

I will demonstrate that $\lim_{t\to\infty} \overline{\beta(t)} = \varepsilon \in [0, \overline{\beta(0)}$ and that T is a proper bounding set for the dynamics. These two facts will give a complete understanding of the transient and asymptotic dynamics of this system.

Supposing $i_\infty = \lim_{t\to\infty} i(t) > 0$ immediately gives that $n_\infty > 0$ and for all t, $\dot{v(t)} < 0$ (implying that $v(t) \to -\infty$). To study the limiting behavior of $\overline{\beta(t)}$, suppose $s_\infty > 0$. Rewriting $\overline{\beta(t)}$ as

$$\overline{\beta(t)} = \frac{\int \beta(w) s(0,w) e^{\beta(w)v(t)+u(t)+\lambda t} dw}{s(t)}, \tag{2.24}$$

one would need to consider $\lim_{t\to\infty} [\beta(w)v(t) + u(t) + \lambda t] \asymp -\infty + \infty$, an improper form. From this we may not determine what $\overline{\beta(t)}$ limits to other than (with the knowledge that it is monotonically decreasing) some value $\varepsilon \in [0, \overline{\beta(0)})$.

To show that T is a proper bounding region we must show that on the boundary of T all flow is inwards. This can be seen by studying the case $s(t) + i(t) = 1$ (the positivity conditions are straightforward). $\dot{s(t)}$ is always negative along this line because

$$\dot{s(t)} = -\overline{\beta(t)} s(t)(1 - s(t)) \leq 0.$$

The $\dot{i(t)}$ equation simplifies to $(1 - s(t))(\overline{\beta(t)} - \gamma)$, and thus $\dot{i(t)} > 0$ for $s(t) > \frac{\gamma}{\overline{\beta(t)}}$. This could be problematic; if the magnitude of flow in the $i(t)$ direction is greater than that in the $s(t)$ direction then the flow would escape T. Assuming that $n(t) = 1$ and $s(t) \leq n(t)$, we have that $n(t)$ is decreasing. Thus $s(t)$ is decreasing more than $i(t)$ increases and the flow remains within the region T! Therefore T is a proper bounding region for the dynamics should $(s(0), i(0), n(0)) \in T$.

We are left with several possible transient/asymptotic dynamic situations for the fixed points $(s^*, i^*, r^*) = (0,0,0), (1,0,0), \left(\frac{\gamma^2}{(\varepsilon-\lambda)^2}, \frac{\lambda\gamma(\varepsilon-\lambda-\gamma)}{[\varepsilon(\varepsilon-\lambda)]^2}, \frac{\gamma(\varepsilon-\lambda-\gamma)}{\varepsilon(\varepsilon-\lambda)}\right)$, the trivial, disease free, and the limit of the endemic equilibrium trajectory:

1. The limit ε is both greater than $\gamma+\lambda$ and 1: The trivial equilibrium and the DFE are saddles. The trivial equilibrium attracts along the $i(t)$ and $r(t)$ axes and repels along the $s(t)$ axis. The DFE attracts along the $s(t)$ axis and the line $s(t)+r(t)=n(t)$ and repels in the direction of $s(t)+i(t)\leq n(t)$. The endemic equilibrium (EE) is also a saddle with two eigenvalues with positive real part and one negative eigenvalue. Solutions will oscillate about the current value of (2.23) and asymptotically approach oscillation about the limit of (2.23).

2. The limit ε is in $(\gamma+\lambda,1)$: The trivial equilibrium and the DFE are saddles. The EE is now an attractor for all trajectories in the interior of T.

3. The limit ε is less than $\gamma+\lambda$: The DFE is the attractor for all trajectories in the interior of T, and the EE is outside of T repelling trajectories into T.

4. $\overline{\beta(t)}>\gamma+\lambda$ for $t\in[0,\tau)$ and $\overline{\beta(t)}<\gamma+\lambda$ for $t\in(\tau,\infty)$: Until time τ the system will appear as in Case 1, Case 2, or Case 1 and then Case 2 (depending on the sign of $\overline{\beta(t)}-1$). After time τ the behavior will be as in Case 3.

2.3 Discussion

Albeit in a narrow context, this section has focused on one way to handle biological heterogeneity, via transport equations akin to the reduction theory of Karev. Lemma 2.1.1 follows from the monotonicity of $\overline{\beta(t)}$. This property proves qualitative equivalence (not a new result, but a new way to show it) as well as demonstrating that the undistributed case infects the most individuals. This latter result in closed populations is intuitive but was shown to hold for the pure inheritance model as well.

The simple, open *SIR* model was reduced to a non-autonomous, finite dimensional system of ODEs, but due to the resulting nature of $\overline{\beta(t)}$ (involving the solution trajectory of the transport variable) it challenging to analyze. Without the ability to demonstrate monotonicity (at least after some time τ) we are not afforded with the ability to rule out, or construct conditions for, sustained oscillatory behavior.

The pure inheritance model, System (2.15), is perhaps a biologically unrealistic work-around for the "blue-sky" births found in the simple, open *SIR* case, but it was in a form receptive to the transport system representation. The reduction of the homogeneous system to one with states in $[0,1]$ allowed for the fixed point analysis of the closed *SIR* to be applicable to stability analysis of exponential trajectories. Interestingly, despite being an *SIR* model with demographic dynamics, the births being split into the three classes prevents the presence of oscillatory solutions. This splitting of births possibly is perhaps what confers monotonicity on $\overline{\beta(t)}$.

The sterilization model, a simplification of papillomavirus dynamics in sheep given by System (2.18), exhibits a wide range of transient dynamics due to the stability of the DFE being dependent on the relationship between $\overline{\beta(t)}$ and the other vital rates λ and γ as well as its own magnitude related to unit. Motivation for splitting up the population into variable susceptibility is in the spirit of black boxing several cofactors: nutrition, cleanliness of environment, genetic variability, and epidemiological history of the individual sheep. Furthermore, the model does not incorporate the probable culling/removal of infected sheep, the partial infectivity of those sheep in the chronic phase, and I assumed that chronic infected (recovery) necessarily led to sterility (early detection and treatment can prevent the scarification from occurring although the sheep would probably be removed from the breeding population to prevent more infection).

I demonstrated that for particular values of the mean distribution of infectivity that the disease free equilibrium is asymptotically stable, Cases 3 and 4, and that in the latter situation there is an interesting transient behavior of the solution curve "chasing" an equilibrium which vanishes from the biologically meaningful space. I've also given conditions for oscillatory behavior should the mean susceptibility *not* limit to zero, Case 1. In this situation the dynamics are quite complex; because there is not a fixed limit cycle for finite time, there is a "moving" oscillation through the phase space (the trajectory is oscillating and where the oscillation occurs is moving). Finally in Case 2, I was able to show that the endemic state is a "global" attractor within T.

Extensions to the work of this chapter may include a more general theoretic version of the transport system theory of Karev as applied to models requiring separation of variables to be "solved", as in the case of the simple, open *SIR*. Additionally, heterogeneity may be introduced to infectivity and recovery rate as Novozhilov did for closed populations [109]. These heterogeneities were not introduced here because they induce further transport equations, and was beyond the scope of an explanation of the method through novel examples. The question of parameter estimation was raised in Section 2.1. While the robustness of estimators of parameters in epidemiological compartmental models has been demonstrated, [62], for all but behavioral effects, it is of interest to investigate if the distribution, or the parameters of an assumed distribution, of an epidemiological parameter may be estimated from collected data.

2.4 Acknowledgements

I would like to thank Georgy Karev and Artem Novozhilov for helping to answer my technical questions about the method. Additionally I acknowledge Juan Aparicio for putting in plain language some of the finer points.

Part II

Behavioral Heterogeneity

Chapter 3

Static Behavior Effects on Gonorrhea Transmission Dynamics in a MSM Population [102].

The task of disease eradication and prevention, while of interest to a society as a whole, is in the hands of "active" individuals. Active can be taken to mean any behavior or circumstance that puts that individual in the process of the disease spreading. Of interest here is the spread of sexually transmitted diseases (STDs), more specifically gonorrhea, and thus active is taken to mean highly-sexually active. The study of so-called *core groups* has been used, to some success, in modeling the prevalence of an STD [35, 49, 51, 72, 73].

In classical models controls are viewed as amalgamations of several effects that are treated via a single parameter (i.e. β as a "force of infection") [6, 8, 10, 66, 72]. The behavior of individuals in a population can limit the spread of disease and we place them into two categories–prevention and self awareness. Prevention is specific to the characteristics of each disease. With STDs such as gonorrhea, prevention education advocates safe sex practices, condom use for a larger proportion of the time and reduction in the number of sexual partners. Self awareness is simply possessing a knowledge of symptoms, the presence of asymptomatic dynamics, and getting tested and treated. Self awareness leads to frequent STD screening as a disease control method since individuals may be infected but have no knowledge of their own status. In particular, infected individuals may have asymptomatic gonorrhea but still as infectious as those that show symptoms [1, 73]. If an infected individual is asymptomatic, but relatively uneducated about the disease, they will not seek medical attention and their average infectious period is potentially much greater than that of a symptomatically infected individual. Typically, symptomatically infected individuals

recover 14 days after the start of treatment which, due to the pain associated with the symptoms, begins once symptoms show a few days after gonorrhea is contracted [1]. Seminal mathematical work was done by Hethcote and York on heterosexual gonorrhea transmission [73]. There, a two-sex model of symptomatically infected, asymptomatically infected, and susceptible populations with different activity levels was used. The focus of this study was the effect that contact tracing and increased STD testing would have on the dynamics of the disease. It was found that contact tracing of infectees was less effective than that of infectors. Here the infectors were identified as a core group, or highly sexually active subpopulation. Hethcote concluded that focusing on the core group's activities was key to controlling the spread of gonorrhea.

Li *et al.* specifically modeled STDs like gonorrhea using an *SIS* model with multiple strains and varied reaction to infection [92]. It was found that sufficient heterogeneity of the female in the form of contact structure, immune response or activity level was necessary in order to have coexistence of the multiple strains. The conditions for the existence and stability of an endemic equilibrium with two strains were found. Coexistence required that one strain be better at infecting one subpopulation while the other be better at infecting another. It was concluded that each strain creates reservoirs in the population that it is less able to infect.

Although not specifically looking at gonorrhea, previous mathematical epidemiological studies by Kemper *et al.* have developed a general model for curable diseases with symptomatic or asymptomatic infection [85]. There an SIS/SAS model was developed to consider the impact of asymptomatic attacks. However, this model does not treat the different recovery time of asymptomatic infected individuals, nor does it account for the different proportion of contacts with symptomatically infected individuals that lead to symptomatic versus asymptomatic infection and vice-versa.

As a case study in the effects of behavior change in the spread of a curable disease that confers no permanent immunity, we present an SIS/SAS model of gonorrhea transmission in a men-who-have-sex-with-men (MSM) population that incorporates the effects of education previously mentioned. The choice of the MSM population was because it is a population that is typically high risk and very active [73] and can make occasional contacts with the highly susceptible female population. Thus focusing on the eradication of STD in this high risk and active population can shed light on STD control measures in the population as a whole. In this work we assume that the occasional male-female contacts in the MSM population is negligible from the point of view of spread within the MSM population but has the potential to cause gonorrhea to spread in the larger population not modeled here. The influence of safe sex will be modeled exogenously with a weighted average on the effective contact rate that accounts for changing behavior with respect to condom use, both of which are fixed parameter values. Self awareness will be modeled via an increase in infectious period for asymptomatic versus symptomatic individuals with the asymptomatic individuals exhibiting either high or low levels of awareness (frequent or rare testing). Analysis is done describing the disease free and endemic dynamics as well as quantifying the changes necessary to eradicate gonorrhea in this population.

3.1 Mathematical Homosexual Gonorrhea Transmission Model

The population modeled is single sex with three homogeneous compartmental states available: susceptible individuals, S, symptomatically infected, I, and asymptomatically infected individuals, A. The model equations are:

$$\begin{aligned}
\frac{dS}{dt} &= \mu(N-S) - (\lambda_1+\varepsilon\lambda_2)\beta S\frac{I+A}{N} + \alpha I + (\rho\lambda_3+\sigma\lambda_4)\alpha A, \\
\frac{dI}{dt} &= (\lambda_1+\varepsilon\lambda_2)\beta S\left(q_1\frac{I}{N}+(1-q_2)\frac{A}{N}\right) - (\mu+\alpha)I, \\
\frac{dA}{dt} &= (\lambda_1+\varepsilon\lambda_2)\beta S\left((1-q_1)\frac{I}{N}+q_2\frac{A}{N}\right) - (\mu+(\rho\lambda_3+\sigma\lambda_4)\,\alpha)A.
\end{aligned} \tag{3.1}$$

The rate at which the susceptible population is lost to infection is $\beta S\frac{I+A}{N}$, where β is the effective contact rate. We introduce control into the system via λ_1, λ_2, and ε. The proportion of the population participating in non-safe sex, or equivalently the proportion of time an individual participates in non-safe sex, is λ_1 and the proportion of the population participating in safe sex is λ_2. The proportion of time safe sex prevents the transmission of gonorrhea is $(1-\varepsilon)$. Thus for the proportion of the time an individual participates in safe sex, λ_2 is multiplied by the reduction factor, ε. These parameters combine to become a total reduction factor on the force of infection, i.e. $(\lambda_1+\varepsilon\lambda_2)\beta$. Individuals are recruited into the I class from the S class at a rate $(\lambda_1+\varepsilon\lambda_2)\beta S\left(q_1\frac{I}{N}+(1-q_2)\frac{A}{N}\right)$, where q_1 and $(1-q_2)$ are the proportion of individuals that become symptomatically infected from contact with symptomatically infected individuals and asymptomatically infected individuals, respectively. Similarly individuals are recruited into the A class from the S class at a rate $(\lambda_1+\varepsilon\lambda_2)\beta S\left((1-q_1)\frac{I}{N}+q_2\frac{A}{N}\right)$. Individuals from the I class reenter the susceptible class due to treatment at a rate α. Since individuals in the A class do not know they have gonorrhea the average duration of infection of an asymptomatic person is assumed to be longer than that of a symptomatic individual. We represent this increase in infectious period via $\rho\lambda_3+\sigma\lambda_4\in[0,1]$. The proportion of people who engage in low levels of self awareness, do not get tested often, is denoted by λ_3 while the population of those who have a high level of self awareness are represented by λ_4. The parameters ρ and σ are extensions to the symptomatic treatment rate α caused by low and high levels of self awareness respectively, and by the nature of their meanings $\rho<\sigma$. Thus reentry into the susceptible population from the asymptomatic class occurs at the rate $(\rho\lambda_3+\sigma\lambda_4)\alpha A$.

The model assumes a constant population size, N, with constant recruitment and removal, μ. This limits our ability to extend the time scale of this work but allows for analysis to be carried out that would not be possible. For this reason we are choosing

college bound (a time of high-sexual activity and educational influence) as a further narrowing of the population. The reasonable time scale for the model is thus over 4 years, where it is assumed that model parameters cannot change and population will hold relatively constant. We also do not have memory in our population. Therefore, once diagnosed with asymptomatic gonorrhea an individual is not more likely to get frequent testing [73]. Here, new individuals will enter the system as old ones leave making the absence of memory an acceptable assumption. We present control here as a meaningful separation of the more classical parameters, but these are still averaged measures of behavior. Thus, we cannot make any tracking of how often a specific individual has been diagnosed asymptomatic or has become infected at all. We also do not look at dynamic changes in sexual mixing through the use of evolving contact structures. A few proposed generalizations will be made in the end of the paper as veins of future work but the thought governing this work is to start with as little detail as seems relevant.

Parameter Estimation

The system we model is a college attending (18-22 years old), MSM population thus μ is taken to equal $\frac{1}{4yr}$. This group was chosen due to the prevalence of both highly sexually active individuals and the ability for policy changes to potentially reach the entire population. The behavioral parameters desired for the model were not found from a single source or over a single demographic across many sources and thus we have put together several sources' worth of parameter estimates as an approximation to the potential behavior of the study group. The effective contact rate is the product of the number of risky contacts per year and the proportion these that lead to infection. Based on the information given by [110], roughly 50% of MSM contacts with an infected individual leads to a new infection in the susceptible partner. Making an estimate of 100 risky contacts per year we arrive at $\beta = 50$ infectious contacts per

year. To determine the amount of time spent practicing safe versus risky behavior, we consider a study of an MSM population by Shlay *et al.* [125] where 25.6% of participants report consistent condom usage, thus $\lambda_2 = .256$ and $\lambda_1 = 1 - \lambda_2 = .744$. To determine ε we take into account that although condoms are 97% effective at preventing gonorrhea infection with perfect use, many uses are imperfect due to slippage, breaking, etc. According to Shlay *et al.* and Stone *et al.* [129], there is a 16.6 - 17.3% usage failure of condoms in a MSM population. We chose to let $\varepsilon \approx .173$. According to [110] symptoms typically appear for men within 4-6 days, $\frac{4}{365} - \frac{6}{365}$ years, following infection, and treatment duration is 7 days, $\frac{7}{365}$ years. Thus we take $\alpha \approx \frac{365}{12yr}$. A study by Fortenberry *et al.* showed that roughly 50% of men got tested for STD's, including gonorrhea, annually [60]. This study was done without differentiating between hetero- or homo- sexual males and the age range was larger than that of your typical college bound of 18 to 22, (about 50% of their sample was in our age range). This therefore serves as a poor estimate, an estimate none the less, and is probably an overestimate of safe behavior. Using this information we let $\lambda_3 = \lambda_4 = .5$. For the values of ρ and σ we will use testing once every year and once every 3 months, the CDC recommendation for highly-active MSM [1]. Thus we have $\rho * \alpha = \frac{365}{372yr}$ and $\sigma * \alpha = \frac{365}{99yr}$, or $\rho = \frac{12}{372} = \frac{1}{31}$ and $\sigma = \frac{12}{99} = \frac{4}{33}$. Making ρ represent a year as asymptomatic is probably an underestimate for this group, but it is our hope that this models some form of "college" pressure: student health efforts, peer pressure, partner desire. The probabilities q_1 and q_2 are two difficult parameters to estimate. If data were collected that found the probability of exhibiting a certain type of infection given a successfully infectious contact with an infectious individual (not necessarily of the same type) then we'd immediately have the values. However, what we generally have is some presence proportion (i.e., a percentage of the infected population who is of a particular type). Thus, the probabilities would solve the necessary relationship in assumed endemic populations. According to [28] 29.9% of infected individuals are

asymptomatic. Thus we would have to solve $.299I_\infty = .701A_\infty$ for q_1 and q_2. This is daunting since the expression for the endemic equilibrium is very large and the result would not necessarily be an invertible function (i.e., more than one pair of the probabilities would give the same relation). A numerical investigation, using the nominal values for every other parameter and varying q_1 and q_2, provides a few values for the respective probabilities that are near the 29.9% estimate. Choosing $q_1 = .965$ and $q_2 = .034$ we get an endemic percentage of asymptomatic individuals of 29.85%.

3.2 Global Stability Analysis

With the model constructed we wish to do a full analysis on the qualitative behavior of the system. To that end the main focus of this section is to prove Theorem 3.2.1. The discussion proving this will involve a treatment of the stability of the disease free equilibrium, conditions on the number of endemic equilibria that may exist, and a preclusion of closed orbits which will make all stability arguments global.

Theorem 3.2.1 (Global Stability)**.** *System 3.1 has 2 fixed points: a disease free and an endemic equilibrium. The disease free equilibrium is globally stable when the control reproductive number is less than one and unstable when the control reproductive number is greater than one. The endemic equilibrium does not exist biologically when the control reproductive number is less than one and is globally stable when the control reproductive number is greater than one.*

To start we wish to formulate System 3.1 into a slightly easier form. Since $S(t)+I(t)+A(t) = N$ we may eliminate one state variable for the purpose of analysis. We may also rescale in both state and time to reduce the overall system. Using $x(\tau) = \frac{I}{N}$, $y(\tau) = \frac{A}{N}$, and $\tau = t\omega\beta(1-x-y)$, with $\omega = \lambda_1 + \varepsilon\lambda_2$ and $p = \rho\lambda_3 + \sigma\lambda_4$,

we arrive at:

$$\begin{aligned}\frac{dx}{d\tau} &= q_1 x + (1-q_2)y - \frac{q_1 x}{R_{II}(1-x-y)},\\ \frac{dy}{d\tau} &= (1-q_1)x + q_2 y - \frac{q_2 y}{R_{AA}(1-x-y)},\end{aligned} \tag{3.2}$$

where $R_{II} = \frac{\beta\omega q_1}{\mu+\alpha}$ and $R_{AA} = \frac{\beta\omega q_2}{\mu+p\alpha}$. Since each state variable S, I and A are positive we have that $x+y \leq 1$. Thus the rescaling is positive invariant. There is a disease free equilibrium that always exists, $DFE := (x^*, y^*) = (0,0)$, implying $S^* = N, I^* = 0$ and $A^* = 0$. To determine the local stability of the DFE the system is linearized about (x^*, y^*) resulting in the following:

$$J_{(x^*,y^*)} = \begin{pmatrix} q_1 - \frac{q_1}{R_{II}} & (1-q_2) \\ (1-q_1) & q_2 - \frac{q_2}{R_{AA}} \end{pmatrix}. \tag{3.3}$$

The characteristic polynomial of the above Jacobian is:

$$\lambda^2 - \left[q_1 + q_2 - \frac{q_1}{R_{II}} - \frac{q_2}{R_{AA}}\right]\lambda + \left(q_1 - \frac{q_1}{R_{II}}\right)\left(q_2 - \frac{q_2}{R_{AA}}\right) - (1-q_1)(1-q_2),$$

which is in the form $\lambda^2 - b\lambda + c$. It may be shown that conditions for the determinant of the jacobian to be positive, $c > 0$, are identical to R_E, the basic control number, being less than one. First, we use the next generation operator to compute the number of secondary infections a typical infectious individual creates in a completely susceptible population [134]. This method essentially creates two vectors F and V. The vector F contains any rates in the infected classes' ODEs that constitute recruitment from a non-infected class. The vector V contains the negative of all other rates except for the rates in non-infected classes that represent recruitment into the infected classes. This is more easily done with Equations 3.1 than the rescaled system

in Equations 3.2. This gives us

$$F = \begin{pmatrix} 0 \\ \omega\beta S\left(q_1\frac{I}{N} + (1-q_2)\frac{A}{N}\right) \\ \omega\beta S\left((1-q_1)\frac{I}{N} + q_2\frac{A}{N}\right) \end{pmatrix}, \tag{3.4}$$

$$V = \begin{pmatrix} -\mu(N-S) - \alpha I - p\alpha A \\ (\mu+\alpha)I \\ (\mu+p\alpha)A \end{pmatrix}. \tag{3.5}$$

We then define $\mathscr{F}$ and $\mathscr{V}$ as the Jacobians of F and V respectively. Specifically these are matrices where the $(i,j)^{th}$ element is the partial derivative of the i^{th} term of the vector with respect to the j^{th} state variable. The spectral radius, or dominant eigenvalue, of $\mathscr{F}\mathscr{V}^{-1}$ is then our basic reproductive number. The numerator of each term in Equation 3.6 is from $\mathscr{F}$ and each denominator is from $\mathscr{V}^{-1}$.

$$\mathscr{F}\mathscr{V}^{-1} = \begin{bmatrix} \frac{\beta\omega q_1}{\alpha+\mu} & \frac{\beta\omega(1-q_2)}{\mu+p\alpha} \\ \frac{\beta\omega(1-q_1)}{\alpha+\mu} & \frac{\beta\omega q_2}{\mu+p\alpha} \end{bmatrix} = \begin{bmatrix} R_{II} & R_{AI} \\ R_{IA} & R_{AA} \end{bmatrix}. \tag{3.6}$$

Interestingly, the terms in Equation 3.6 are each control reproductive numbers in and of themselves. The term R_{ij} is the average number of individuals recruited into class j from a typical member in class i per unit time. Thus the control reproductive number of the entire system is a function of each of these individual recruitment thresholds. The spectral radius of Equation 3.6 is

$$R_E = \frac{R_{II} + R_{AA} + \sqrt{(R_{II}-R_{AA})^2 + 4R_{IA}R_{AI}}}{2}. \tag{3.7}$$

Corollary 3.2.1. *The condition $R_E < 1$ is identical to $c > 0$ for the characteristic polynomial $\lambda^2 - b\lambda + c$ for system 3.2.*

Proof. We begin with the condition for $R_E < 1$:

$$\begin{aligned}
\frac{R_{II}+R_{AA}+\sqrt{(R_{II}-R_{AA})^2+4R_{IA}R_{AI}}}{2} &< 1,\\
(R_{II}-R_{AA})^2+4R_{IA}R_{AI} &< (2-(R_{II}+R_{AA}))^2,\\
R_{II}^2-2R_{II}R_{AA}+R_{AA}^2+4R_{IA}R_{AI} &< 4-4(R_{II}+R_{AA})+(R_{II}+R_{AA})^2,\\
-4R_{II}R_{AA}+4R_{IA}R_{AI}+4(R_{II}+R_{AA}) &< 4,\\
R_{II}+R_{AA}+R_{II}R_{AA}\left(\frac{1-q_1-q_2}{q_1q_2}\right) &< 1.
\end{aligned}$$

If we then consider the condition for $c > 0$:

$$\begin{aligned}
\left(q_1-\frac{q_1}{R_{II}}\right)\left(q_2-\frac{q_2}{R_{AA}}\right)-(1-q_1)(1-q_2) &> 0,\\
q_1q_2(1-R_{II}-R_{AA})-R_{II}R_{AA}(1-q_1-q_2) &> 0,\\
1-R_{II}-R_{AA} &> R_{II}R_{AA}\frac{1-q_1-q_2}{q_1q_2},\\
R_{II}+R_{AA}+R_{II}R_{AA}\left(\frac{1-q_1-q_2}{q_1q_2}\right) &< 1, \qquad (3.8)
\end{aligned}$$

we see Equation 3.8 is identical to the condition for $R_E < 1$. □

Thus we have a new condition for stability and may define

$$\hat{R}_E = R_{II}+R_{AA}+R_{II}R_{AA}\left(\frac{1-q_1-q_2}{q_1q_2}\right).$$

This quantity is somewhat easier to understand and its relationship to 1 is identical to R_E's.

If $R_E > 1$ then the DFE is unstable, but we haven't discussed how many equilibria may exist. Consider $z = x+y$ and the fact that if $\frac{dz}{d\tau} = 0$ and $\frac{dy}{d\tau} = 0$ then we would be at a fixed point for System 3.2. Solving $\frac{dz}{d\tau} = 0$ we get an expression for y in terms of z by noting $x = z-y$. Plugging this expression into $\frac{dy}{d\tau}$ and solving the new expression for zero we get $zf(z) = 0$, where $f(z) = z^2 - Bz + C$,

$$B = 2+\frac{q_1q_2\left(\frac{1}{R_{AA}}+\frac{1}{R_{II}}\right)}{1-q_1-q_2},$$

and

$$C = \frac{(1-q_1)(1-q_2) - q_1 q_2 \left(1 - \frac{1}{R_{II}}\right)\left(1 - \frac{1}{R_{AA}}\right)}{1-q_1-q_2} = \frac{c}{-(1-q_1-q_1)},$$

where c is from the characteristic polynomial. If $z = 0$ then $x = y = 0$, the DFE. The interest thus lies in where $f(z) = 0$. A relationship between C and R_E can be made using the existing relationship found in Corollary 3.2.1.

Corollary 3.2.2. *If $q_1 + q_2 < 1$ then $R_E < 1 \Leftrightarrow C < 0$ and $R_E > 1 \Leftrightarrow C > 0$. If $q_1 + q_2 > 1$ then $R_E < 1 \Leftrightarrow C > 0$ and $R_E > 1 \Leftrightarrow C < 0$.*

Proof. If $q_1 + q_2 < 1$ then C and c have differing sign and thus if $c > 0$ then $C < 0$ and their relationships to R_E are the opposite of one another. If $q_1 + q_2 > 1$ then C and c have the same sign and their relationships to R_E are identical. □

The equation $f(z)$ is a quadratic and thus may have 0, 1, or 2 roots in $(0,1)$. The relative signs of $f(0)$ and $f(1)$ will allow us to determine under what conditions $f(z)$ has a particular number of roots in the unit interval. Consider $f(0) = C$ and $f(1) = 1 + C - B$. We already have that the sign of C may be viewed as being dependent on the magnitude of R_E. We also have that

$$1 + C - B = -\frac{q_1 q_2}{1-q_1-q_2}\frac{1}{R_{II}R_{AA}},$$

whose sign depends on $1 - q_1 - q_2$. Thus in order to study the zeros of $f(z) = z^2 - Bz + C$ we must consider all 4 combinations of the sum of q_1 and q_2 with the magnitude of the control reproductive number.

1. $R_E < 1$ & $q_1 + q_2 < 1$

   In this situation $f(0) = C < 0$ and $f(1) = 1 + C - B < 0$. Thus there are no zeros of $f(z) \in (0,1)$.

2. $R_E > 1$ & $q_1 + q_2 < 1$ Here we have that $f(0) > 0$ and $f(1) < 0$. Thus there is a single root for $f(z) \in (0,1)$.

3. $R_E < 1$ & $q_1 + q_2 > 1$ In the most difficult situation we have that $f(0) > 0$ and $f(1) > 0$. The quadratic having exactly two roots occurs when

   a) $\frac{B}{2} \in (0,1)$, and

   b) $B^2 - 4C > 0$.

   In order for $B > 0$ we require that

$$\frac{R_{II} + R_{AA}}{R_{II}R_{AA}} < \frac{2(q_1 + q_2 - 1)}{q_1 q_2}.$$

   The second condition yields

$$\frac{(R_{II} + R_{AA})^2}{R_{II}R_{AA}} > \frac{4(q_1 + q_2 - 1)}{q_1 q_2}. \tag{3.9}$$

   These two conditions are contradictory. To illustrate the contradiction we invoke Equation 3.7 to get that $R_E < 1 \rightarrow R_{II} + R_{AA} < 2$. Rearranging Equation 3.9 results in

$$\frac{R_{II} + R_{AA}}{R_{II}R_{AA}} > \frac{4}{R_{II} + R_{AA}} \frac{(q_1 + q_2 - 1)}{q_1 q_2}, \tag{3.10}$$

   but $\frac{4}{R_{II}+R_{AA}} > 2$ which gives the contradiction. Thus there are no zeros of $f(z) \in (0,1)$.

4. $R_E > 1$ & $q_1 + q_2 > 1$ In this situation $f(0) > 0$ and $f(1) < 0$. Thus there is a single root for $f(z) \in (0,1)$.

Combining the arguments gives us that if $R_E < 1$ then the only solution that is biologically relevant for our system is the DFE. When $R_E > 1$ the two solutions, which are biologically relevant, the unstable DFE and a single endemic equilibrium, EE. The entire above argument is valid only if $1 - q_1 - q_2 \neq 0$. When $1 - q_1 - q_2 = 0$, the system exhibits a single EE,

$$(x, y) = \left(\frac{R_{II}(R_{II} + R_{AA} - 1)}{(R_{II} + R_{AA})^2}, \frac{R_{AA}(R_{II} + R_{AA} - 1)}{(R_{II} + R_{AA})^2} \right),$$

which is only valid if $R_{II} + R_{AA} - 1 \geq 0$. We wish to make assertions about the stability of the EE without having to do a linearization around the fixed point which is very term intensive. We may now disprove the existence of closed orbits in the plane and thus assert that when the EE exists, $R_E > 1$, it is globally stable.

Corollary 3.2.3. *System 3.1 has no closed orbits.*

Proof. By Dulac's criterion, if there exists a function $\varphi \in C^1$ such that $\frac{\partial(\varphi\dot{x})}{\partial x} + \frac{\partial(\varphi\dot{y})}{\partial y} \neq 0$, then the planar system $\dot{x}, \dot{y}$ has no closed orbits. Let $\varphi = \frac{1}{xy}$. Then:

$$\begin{aligned}
\frac{\partial(\varphi\dot{x})}{\partial x} &= \frac{\partial}{\partial x}\left[\frac{q_1}{y} + \frac{(1-q_2)}{x} - \frac{q_1}{R_{II}(1-x-y)y}\right] \\
&= -\frac{(1-q_2)}{x^2} - \frac{q_1}{R_{II}(1-x-y)^2 y} \\
\frac{\partial(\varphi\dot{y})}{\partial y} &= \frac{\partial}{\partial y}\left[\frac{(1-q_1)}{y} + \frac{q_2}{x} - \frac{q_2}{R_{AA}(1-x-y)x}\right] \\
&= \frac{-(1-q_1)}{y^2} - \frac{q_2}{R_{AA}(1-x-y)^2 x}.
\end{aligned}$$

Since $q_1, q_2 \in [0,1]$ and $R_{AA}, R_{II} > 0$ the sum $\frac{\partial(\varphi\dot{x})}{\partial x} + \frac{\partial(\varphi\dot{y})}{\partial y}$ is always negative. Thus there are no closed orbits for system 3.2. Since the dynamics are identical for system 3.1 we have precluded limit cycles in it and have shown what was intended. □

3.3 Sensitivity Calculations for the Gonorrhea Transmission Model

In a perfect world, public health initiatives would be simple, multifaceted, and have great effects on the dynamics of a disease. However, this is not always the case and moreover, economic costs have to be considered in determining which interventions to support. Another important question is whether the size of the population is important to the dynamics of the disease and intervention. In order to address these concerns, in this section we examine the sensitivity of the system to changes in the parameter values. We take two approaches, first the time-dependent sensitivity of the original system to changes in parameter values is calculated via elasticity and secondly we compute what changes to the parameters would be necessary to bring the control reproductive number to a value less than 1.

In order to discuss the importance of individual parameters one must investigate how their value affects the state variables over time. We will consider the concept of elasticity. Formally, one may define elasticity of a function $X(t;\theta)$ with respect to a parameter $\theta_i \in \theta$ as

$$E_{\theta_i}(t) = \frac{\theta_i}{X(t;\theta)} \frac{\partial X(t;\theta)}{\partial \theta_i}.$$

This measures the ratio of a percent change in a parameter to that of the function. The elasticity is a unit-less and scaled method with which to compare each parameter's affect on the solution $J(t) := I(t) + A(t)$. However, we do not have a closed form for $J(t)$ and thus we must make an approximation to both its value and its partial derivative.

In general, consider $\frac{dX}{dt} = f(t, X; \theta)$ where θ is a parameter vector. Now consider the vector of nominal parameter values, $\hat{\theta}$, and a small perturbation, Δ_i, of the i^{th} element, $\hat{\theta}_i$, and call this new parameter vector $\hat{\theta}^i$. If what we are interested in is $\frac{\partial X}{\partial \theta^i}$ near the nominal value then we could do the following. Numerically find the solutions $X(t;\hat{\theta})$

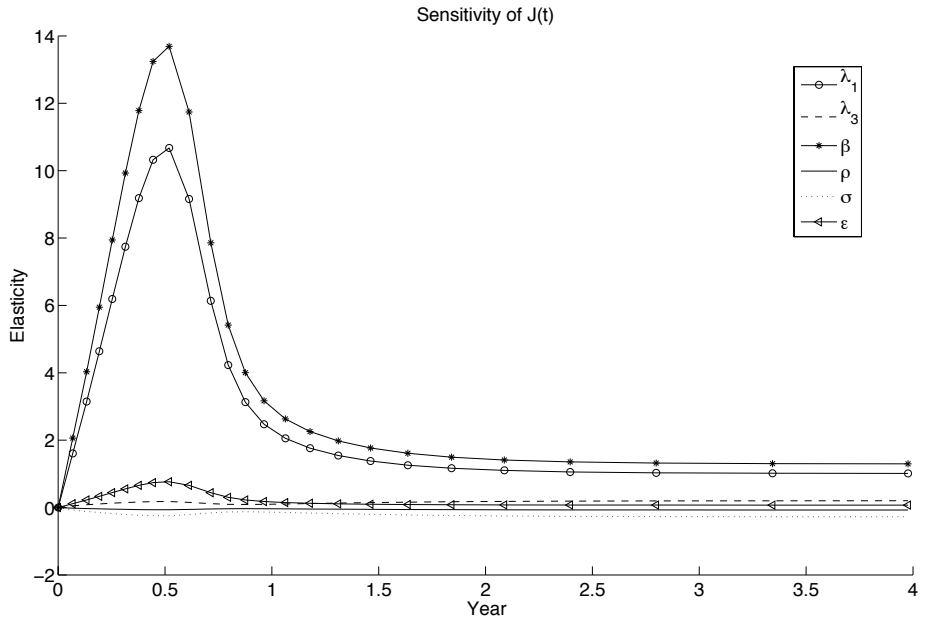

Figure 3.1: Sensitivity of the susceptible population with respect to each model parameter using the nominal values $(\mu,\lambda_1,\lambda_3,\beta,\alpha,\rho,\sigma,q_1,q_2,\varepsilon) = (1/4,.744,.5,50,365/12,1/31,4/33,.965,.034,.173)$ and $N = 10000$ with $S(0) = 99999, I(0) = 0$, and $A(0) = 1$.

and $X(t;\hat{\theta}^i)$ then take the difference quotient to arrive at

$$\frac{\partial X}{\partial \theta_i}(t) \approx \frac{X(t;\hat{\theta}^i) - X(t;\hat{\theta})}{\Delta_i}. \tag{3.11}$$

We may use this approximation in our computation of the elasticity of $J(t)$ with respect to each parameter.

The sensitivity of the total infected population with respect to the nominal parameters is presented in Figure 3.1 with $\Delta_i = .001\theta_i$ used in each of the approximations. Not surprisingly β acts as a large control on the infected population. Small increases in β will result in the largest changes to $J(t)$ and increase the number of individuals infected. The parameter modeling non-condom use, λ_1, intuitively is also very large. The remaining four control parameters, λ_3,ρ,σ, and ε are all near zero and thus each will require large changes to impact the infected population's size.

3.4 Reducing R_E

Holding all but one parameter constant we varied a free parameter until $R_E < 1$. Each parameter was varied by 0.1% in the correct direction to reduce the infected population, analogous to reducing R_E. Of these only two single effort campaigns resulted in a control reproductive number less than one, λ_1 and β. Reducing λ_1 to roughly .392, condom use increases to 61% of the time, or β to 28.239, 56.5 risky contacts per year, will result in $R_E < 1$. Each represents a fairly significant change in behavior. Even with the entire population following the CDC recommended testing frequency of once every 3 months the disease will remain prevalent.

Instead if one considers a multifaceted campaign to disease reduction we may vary each parameter an once, by 1/10%, until $R_E < 1$. This will represent smaller changes to each parameter. However, we fixed σ at the CDC recommendation to measure its effectiveness. If the control parameter vector reads

$$(\lambda_1, \lambda_3, \beta, \rho, \varepsilon) \approx (.5605, .3767, 37.67, .0428, .1303)$$

then the disease free equilibrium becomes stable. This translates to 43.95 % of the population using a condom, up from 25.6 %, 62.3 % of the population using the CDC's testing frequency recommendation, reduce the number of risky a contacts per year from 100 to 75.34, even if you do not follow the CDC recommendation for testing get tested once every 9 months and finally reduce condom failure rate from 17.3 % to 13 %.

3.5 Discussion

There are a few shortcomings in this model that may be able to be addressed. The removal of a constant population size could potentially remove the monotonicity of the system and thus would remove the hope for any kind of qualitative global stability analysis via criteria as supplied herein. It would fall upon the modeler to do a more

complex stability analysis where cycles may develop. It is unclear to the authors here if gonorrhea prevalence in this type of population merits the need for cyclic dynamics. In this chapter we considered the *average*, safe and risky behavior of the population. While this has given insights into how education affects disease transmission, it does not explicitly treat the heterogeneity present in such populations. This is a step forward from more classical models where the individual effects were not treated explicitly, but more work may be done for, if not analytical then numerical, results to suggest policy. Several treatments could simply increase the number of compartments while still not dealing with individual level concerns. Specifically, we may take each of the S, I, and A classes and divide them into those who are "safe" and those that are "risky" with respect to each behavior resulting in a system of 12 equations (i.e. susceptible population who are risky with respect to protection and self awareness, risky with only one behavior or risky with neither) with flow between like-state groups of different "risk" types. Work done in [22, 24, 35, 101] has also shown the importance of affinity-based mixing while others have turned to simulation to look at prevention strategies on networks (see [85]). Here we have proportionate mixing which should not be accurate given the pain of symptoms and the mixing assumed between risky and not risky individuals. Furthermore, there is this intuitive desire to model stubbornness in a system such as this. Two difficulties with educational measures is how broad of outreach one should invest in, how many people get to see the information, and how intense the education should be. The trade offs between should we remind everyone a little or remind a few constantly can be addressed with more game theoretic or individual based approaches [117, 123].

In this chapter, a simple model of homosexual gonorrhea transmission between men is analyzed and the parameter sensitivities are determined. In section 3.3, it is shown that disease education is useful in disease control, however education should be multifaceted in order to reduce the need of changing any one parameter by too much.

In detail, the system is most sensitive to β (the number of risky contacts per year), and this parameter may be changed all on its own enough to eradicate gonorrhea. As shown in Section 3.3, the sensitivities are time dependent. Although gonorrhea is best viewed as an endemic disease there are four-year residence systems, like a high school, where the introduction of a single infected individual is feasible.

We have shown that some single methods of education, within realistic bounds, are not effective at reducing disease prevalence. Testing may be too costly and inconvenient for some individuals while from a public health perspective frequent testing is best.

Chapter 4

Model Selection in Test Theory Without an Answer Key with Multiple, Fixed *a priori* Cultures [100].

Since their development in the 1980s, Cultural Consensus Models (CCM) have been widely used by social scientists to model the distribution of cultural knowledge in populations [17, 120, 136]. Derived from Test Theory Without an Answer Key, CCM is based on the assumption that people independently draw their responses from a common cultural answer key, with some individuals having better knowledge of the answer key than others. Researchers have used these models for a number of purposes: (1) to decide if there is a single cultural answer key, (2) to estimate such an answer key, (3) and to estimate individuals' relative knowledge of the answer key [136].

The standard implementation of the model involves estimating two sets of free parameters: the 'culturally correct' response to each of M dichotomous questions and a 'competence' for each of N individuals indicating the probability that he or she knows the correct answer. To provide some approximation to the maximum likelihood estimate for these parameters, the standard implementation uses a factor analysis of a person-by-person matrix of response agreement. A researcher then assesses the fit of the estimated model using two decision criteria derived from the factor analysis. First, the ratio of the first-to-second eigenvalues should be greater than three. Second, the elements of the first factor vector should be positive. If these criteria are met, then an investigator usually infers a single cultural answer key, reports the answer key, and may analyze individual differences in competence [17, 119, 120, 136].

In this chapter, we argue that this standard approach to fitting and assessing the cultural consensus model can lead to a number of inferential errors. For example, it can lead researchers to incorrectly infer that estimated competences reflect true

individual differences, a possibility first raised and examined using bootstrapping techniques by Weller [135], or that there is a single answer key, a possibility raised recently by Hruschka et al. [76]. We also describe a solution to these problems by applying a model selection framework grounded in the use of information criteria. The framework provides a statistically principled way to infer the existence of between-individual differences in competence and provides a foundation for assessing the existence of a single answer key. It also provides a flexible way to compare competing models for the distribution of cultural knowledge in populations. For example, applying the framework to one set of data, we show that between-individual differences in guessing yes (something not normally considered in the standard approach) is clearly supported, whereas between-individual differences in competences have only equivocal support support from the data depending on the specific model selection criteria used. We also show that a model in which people can draw from one of three related answer keys fits better than a single answer key. Publicly available code written for MATLAB provides tools for fitting and selecting among these models.

4.1 The Cultural Consensus Model and its Implementation

In Test Theory Without an Answer Key, Batchelder and Romney proposed a model for how people respond to questions about a domain of knowledge based on several assumptions [17, 120]. First, there is a single answer key from which people draw when formulating a response (Z_k). Second, people answer test questions based on conditional hit or miss probabilities (*high threshold model* (HTM)). These are defined by the probability of knowing the answer (e.g., competency, D) and a bias to guessing yes, g, if the answer is not known. These parameters can be specific to each individual and to each question, and define a probability that person i responds yes to the k^{th} question.

More formally, let $X = (X_{ik})_{N\times M}$ be the dichotomous response profile for N individuals to M questions and let $Z = (Z_k)_{1\times M}$ be the proposed answer key. Following Batchelder and Romney [17], we write the probability of the responses profile given an answer key as

$$P[(X_{ik}) = (x_{ik})|(Z_k) = (z_k)] = \prod_{k=1}^{M}\prod_{i=1}^{N} P(X_{ik} = x_{ik}|Z_k = z_k). \tag{4.1}$$

Based on the High Threshold Model, HTM, (also known as the General Concordant Model, GCM) we can specify the likelihood in terms of competences, $D = (D_i)_{N\times 1}$ (in [0,1]), and a bias toward guessing "yes", g (in [0,1]). This formulation assumes, among other things, a local independence condition. That is, students' answers to questions are independent from one another conditional on the answer key. Each individual conditional probability in Equation (4.1) is a Bernoulli random variable which can be expressed as

$$P(X_{ik} = x_{ik}|Z_k = z_k) = \begin{cases} (D_i z_k + (1-D_i)g_i), & x_{ik} = 1 \\ (D_i(1-z_k) + (1-D_i)(1-g_i)), & x_{ik} = 0 \end{cases}. \tag{4.2}$$

This term may be understood by assuming values of x_{ik} and z_k and reducing the expression to either a single term or a sum of two simple terms. D_i is the probability that the individual knew the answer, and g_i is the probability that the individual guessed "yes" if she did not know the answer.

This requires specification and estimation of a parameter vector in the form of $\Theta = [(D_i), (g_i), (Z_k)^T]$. To make approximation viable, the standard factor analytic implementation of the CCM was limited in several ways. First, answer key bias, the probability that a given question on the answer key is "yes", $\pi_k = \pi$, is assumed to be homogenous across questions. Second, bias to guessing "yes" is assumed to be fixed and homogeneous, $g_i = \frac{1}{2}$, across individuals. Third, a single answer key from which

all respondents draw is assumed. The estimation method then involved an unconstrained least squares that had some undesirable properties, including estimations of probabilities lying outside of $[0,1]$ and an unknown degree of approximation to the maximum likelihood estimate [17].

In the following work we examine several modifications of this standard implementation. First, in lieu of the standard factor analytic approach, we search for model estimates that maximize the likelihood of the data given the model. Second, we extend and restrict the model so that g can vary as a population-level or individual-level parameter and D can be fixed as uniform across individuals. Thus person i's parameter set could take on the forms $(D,g),(D_i,g),(D,g_i)$, or (D_i,g_i), where the second is a more general case of the classical approach of $(D_i,.5)$. Should (D_i,g) prove to be the best model given selection criteria then the classical approach would be strongly validated. For simplicity, we do not permit item difficulty to vary, but this is a straightforward extension implemented in Karabotsas and Batchelder [80].

4.2 Models of Subcultural Variation

We also explore the possibility that people draw from "subcultural" models, which in turn are drawn from a single "supercultural" model. The original idea presented by Batchelder and Romney for a multicultural approach was to estimate each culture's answer key [18]. This led to a large number of parameters having to be estimated and only permitted post-hoc comparison of answer keys. We present an alternative of a *filtered higher truth* model. Here, we assume a single supercultural answer key $Z^H = (Z^H_k)_{1\times M}$, and C subcultural groups each with their own answer keys $Z^c = (Z^c_k)$ for $c \in \{1,...,C\}$. Then it is assumed that each subculture has some probability of drawing an answer identically from the supercultural answer key

$$P(Z^c_k = Z^H_k) = \phi^c, \tag{4.3}$$

where the superscript c is not a power but an index. To reduce the number of parameters, we do not estimate each subculture's answer key. Rather, we only estimate a single measure of agreement with the higher truth, ϕ^c, for each subcultural answer key. This gives a likelihood function in the form

$$L(\Theta|X) = \prod_{c=1}^{C} \prod_{k=1}^{M} \sum_{A \in \{0,1\}} \left[P(Z_k^c = A | Z_k^H = z_k) \prod_{i \in G_c} P(X_{ik} = x_{ik} | Z_k^c = A) \right], \qquad (4.4)$$

where G_c is the index set of all individuals in culture c with the constraint that no individual may be in more than one group, $G_{c_i} \cap G_{c_j} = \emptyset$ for $i \neq j$. Additionally, the conditional probability of a subculture's answer for item k given the supercultural answer key is given by

$$P(Z_k^c = A | Z_k^H = z_k) = (\phi^c)^{y_k} (1 - \phi^c)^{1 - y_k}$$

where $y_k = 1$ if $A = z_k$ and 0 otherwise. To get the classical construction of a single culture one may let $C = 1$ and $\phi^c = 1$.

With the introduction of multiple cultures we may introduce a new generalization to the response profile parameters, that of subcultural homogeneity. A parameter is subculturally homogeneous if for everyone within a subcultural group the parameter is the same, but that parameter can vary across cultures. Thus an individual now may have parameter tuples in the forms (D, g), (D_c, g), (D_i, g), (D, g_c), (D, g_i), (D_c, g_c), (D_i, g_c), (D_c, g_i), or (D_i, g_i). Thus we redefine our parameter vector as $\Theta = [D, g, \phi, Z^T]$ to include the multi-cultural parameterization (ϕ).

Model Selection using Information Criteria

Model selection using maximum likelihood-based information criteria provides a statistical framework for selecting among these alternative models. Information criteria, such as the Akaike Information Criteria and Bayesian Information Criteria

used here, are measures that balance a model's goodness-of-fit with its complexity. This avoids the pitfalls of overfitting, i.e. adding more parameters will almost surely result in a better fit. The Akaike Information Criteria, or AIC, is defined as

$$AIC = 2k - 2\ln(L) \tag{4.5}$$

where k is the number of parameters and L is the maximized value of the likelihood function. Since $L \in [0,1]$ we can see that $\ln(L) \leq 0$, and thus while each parameter increases the AIC this can be balanced by increases in the likelihood, L. The Bayesian Information Criteria, or BIC, is defined as

$$BIC = k\ln(n) - 2\ln(L) \tag{4.6}$$

where n is the number of data points. In most cases, BIC penalizes model complexity more than the AIC.

Estimation Procedures

To find the MLE estimate, $\hat{\Theta}$, we start with 1000 randomly chosen parameter vectors $\tilde{\Theta}$ for a given response matrix (model). For each of the parameter vectors, the algorithm makes a small additive perturbation to a single parameter at a time ($\varepsilon = \pm 0.001$) or reverses the response for a single question in the answer key. If the perturbation increases the likelihood, then it is kept, otherwise the value pre-perturbation is kept. This is done iteratively first for Z, then ϕ, D and finally g. The algorithm proceeds until the increase in likelihood over 100 iterations is less than 0.001, and is effectively a simple "hill-climber"/ "greedy" algorithm.

4.3 Description of Data

As an example, we use data presented in Sibley *et. al.* on cultural theories of postpartum hemorrhage in Matlab, Bangladesh [76, 126]. This data covered 235 yes/no questions about the signs, causes, treatment and conditions of postpartum

hemorrhage. The study population may be broken up into three groups based on training and experience in assisting childbirth: lay women (LAY), traditional birth attendants (TBAs) and biomedically skilled birth attendants (SBAs). Hruschka *et. al.* showed that criteria traditionally used to infer the existence of a single answer key (i.e., the eigenvalue ratio criteria) indicated the existence of a single answer key. However, further investigation indicated distinct cultural models related to these meaningful subgroups in the population [76].

4.4 Comparison of Models

We fit the above models using the postpartum hemorrhage data. For the subcultural model, we specified three *a priori subcultural* groupings based on training and experience in assisting childbirth (TBAs, SBAs, and laywomen). For each of the nine models described in Sections 4.1 and 4.2, we chose the parameter estimates which provided the highest likelihood. The distribution of likelihoods across the 1000 starting points shows a clear compression at the maximum, suggesting the maximum of this distribution is at least close to the true maximum, as seen in Figure 4.1. One can see at least two pronounced peaks, possibly indicative of more than one supercultural answer key. The coloration indicates whether the estimated answer key is close to one of the two supercultural answer keys. Answer keys were clustered as follows. First, answer keys, $\tilde{Z}$, that have a reflexive symmetry (flipping all the answers on a given key, $\left\{1-\tilde{Z}_k\right\}_k$) are considered to be isomorphic. Thus, the distance between answer keys (A1 and A2) was judged as the minimum of the hamming distance between A1 and A2 and between A1 and "flipped" A2. Then, we considered that two answer keys that varied from one another by no more than 7 (this was found for this model to be the minimal hamming difference that created two mutually exclusive groups) should be grouped together as deviations from the same answer key. Answer keys from opposite groups always differed by more than 40 responses, indicating two clearly different answer key clusters. Note the marked drop offs to the

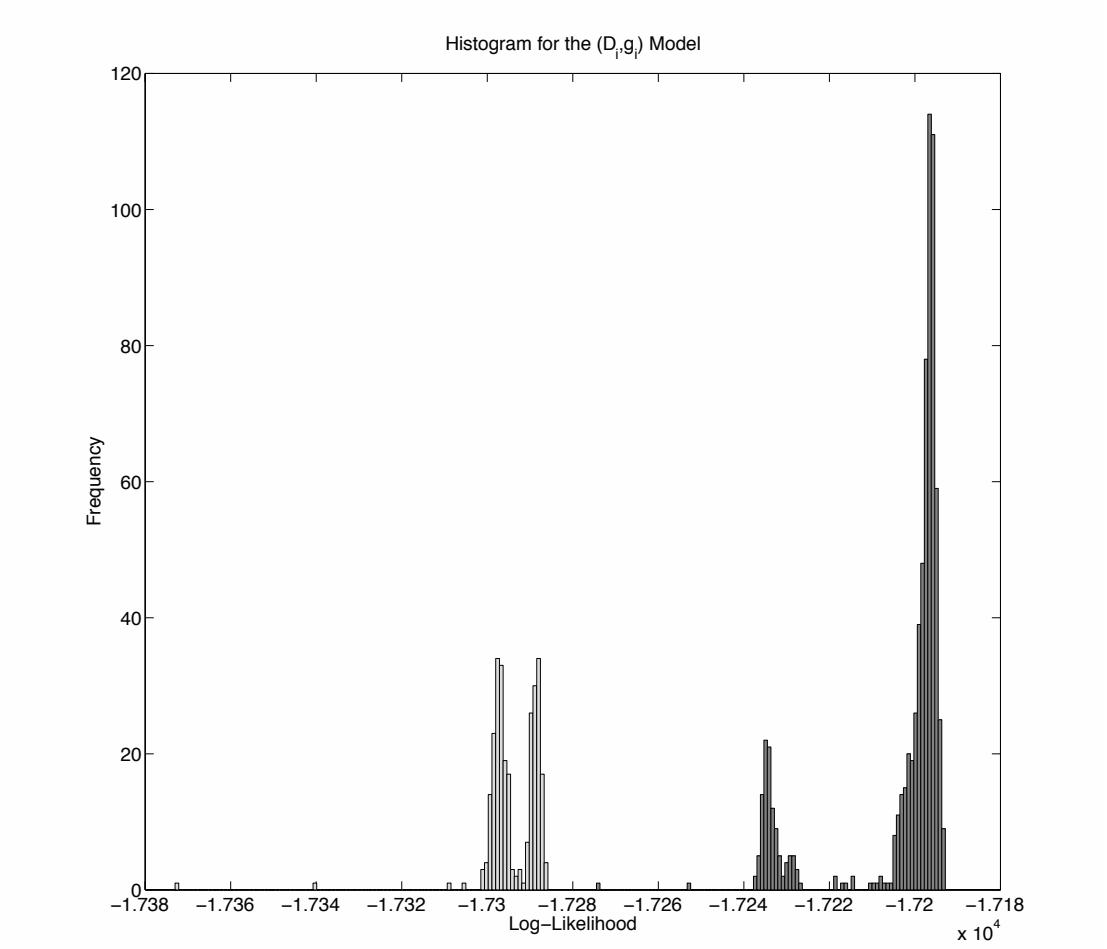

Figure 4.1: Shown here is a histogram of the Log-Likelihood of fits for the multicultural (D_i, g_i) model.

right of each group, especially of the darker group. The second cluster of answer keys leaves open the possibility that the data support at least two distinct answer keys.

We then calculated the information criteria (i.e., AIC and BIC) using the maximum likelihood estimate for each model. For the parameter count, we omit the answer key since it is the same number of parameters for each model, and thus it contributes nothing to the comparative analysis. The results of this are summarized in Table 4.1 and Figures 4.2 and 4.3.

The model selected by AIC was the full subcultural model where each individual has

Model	# of Observations	# of Parameters	log-likelihood	AIC	BIC
$C = 1$					
(D, g)	3515	2	-19247.976	38500	38517
(D, g_i)	3515	150	-18496.2674	37293	38562
(D_i, g)	3515	150	-18669.5880	37639	38908
(D_i, g_i)	3515	298	-17660.3813	35917	38439
UCINet	3515	150	-18774.9725	37848	39109
$C = 3$					
(D, g)	3515	5	-18879.644	37769	37812
(D, g_c)	3515	7	-18844.8708	37704	37763
(D, g_i)	3515	153	-18113.7009	36533	37828
(D_c, g)	3515	7	-18800.3179	37615	37674
(D_c, g_c)	3515	9	-18784.0910	37586	37662
(D_c, g_i)	3515	155	-17551.8450	35414	**36726**
(D_i, g)	3515	153	-18246.2754	36799	38093
(D_i, g_c)	3515	155	-18216.5735	36743	38077
(D_i, g_i)	3515	301	-17183.4269	**34969**	37516

Table 4.1: Data for model comparison for both AIC and BIC. The bold values are the models selected based on each criteria. In either case multicultural models were chosen and the fit provided by UCINet was the worst according to BIC and the second worst based on AIC.

her own competency and bias to how she guesses, (D_i, g_i). BIC chooses the subcultural model where the competencies are culture specific and the guess biases are individualized, (D_c, g_i). Thus, regardless of criteria used, there is justification for individual bias in guessing yes, something that is not included in standard implementations of the model. Moreover, depending on the criteria used for model selection, there may or may not be justification for between-individual differences in competence in this case.

To compare the maximum likelihood estimates described here with the estimates from the standard implementation, here calculated in UCINET, we also estimated the likelihood of the data given the model estimate from UCINET [27]. It fares worse than all models when using BIC and nearly worse than all models according to the AIC.

With this data, the model with three related answer keys was selected over the model with only one answer key, corroborating past findings. Moreover, several pieces of

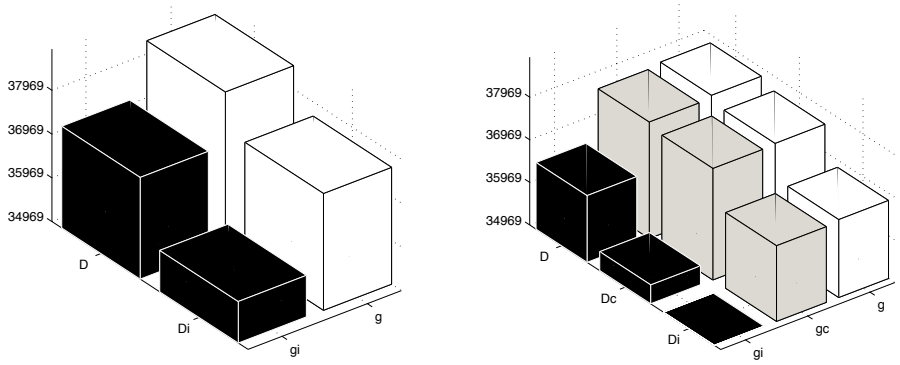

Figure 4.2: Comparison of AIC values for 13 models. The left figure is for a single culture model and the right is for the multicultural model.

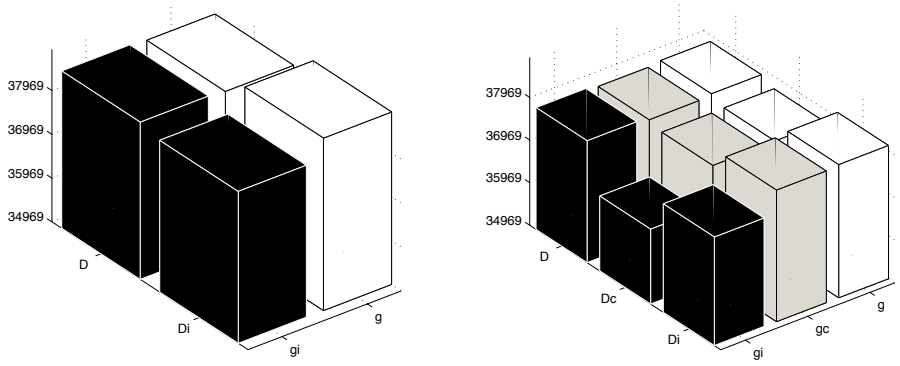

Figure 4.3: Comparison of BIC values for 13 models. The left figure is for a single culture model and the right is for multicultural model.

evidence suggest that there is insufficient data to support a single supercultural answer key. First, in the ensemble of estimates for the (D_i, g_i), and the (D_c, g_i), model, we could delineate two classes of answer keys. Answer keys within these two sets all had fewer than 7 differences, 9 for the (D_c, g_i), among themselves. Meanwhile, answer keys from different sets had at least 41, or 43 for (D_c, g_i), different answers. Further investigation indicated that these two classes of answer keys correspond to different sets of ϕ^c parameters for the three subcultures. In the first case, SBAs had nearly

perfect filtering from the supercultural answer key (while TBAs and Laywomen did not), henceforth the SBA model. In the second case, TBAs and Laywomen had nearly perfect filtering from the supercultural answer key (while SBAs did not), henceforth the TBA model. The maximum likelihood for the SBA model (-17273) was substantially lower than the maximum likelihood for the TBA model (-17183). However, the presence of what appears to be a second local maximum suggests that further study may indicate good posterior support for two answer keys, rather than one. A future analysis using Markov Chain Monte Carlo estimation of the posterior distribution will hopefully determine whether both of these answer keys lie within the 95% credibility interval thus indicating that a single answer key cannot be supported given the data.

4.5 Discussion

The standard implementation of the Cultural Consensus Model and common ways of interpreting its outputs pose several problems. They can lead to incorrect inferences about the existence of between individual variation in competence and of the existence of a single answer key. Moreover, the standard approach does not permit estimation of other parameters, such as guessing bias, or the possibility of multiple, related subcultural answer keys.

The model selection approach and publicly available programs described here provides a framework to address these problems. When applied to a dataset on beliefs about birth complications, we find support for between-individual variation in guessing bias (a parameter not normally considered in the model), we find equivocal support for between-individual variation in competence, and we find strong support for subcultural variation in responses. We show that the results are a dramatic improvement (in terms of likelihood) over the estimates derived from the standard factor-analytic implementation. Also, the inference of multiple, related answer keys

confirms earlier analyses that there is likely more than a single answer key responsible for respondent's answers [76]. And finally, the existence of a second local likelihood maxima with a fundamentally different answer key, suggests that the data might not support a single supercultural answer key. Of course, these results are specific to this population and domain of knowledge, and we expect data from other domains and populations will provide differing support for between-individual variation in model parameters and the existence of a single answer key.

4.6 Acknowledgements

We thank Jamie Jones for helpful comments on an earlier draft. We also are extremely grateful to Lynn Sibley for sharing the postpartum hemorrhage data which was collected with funding from Emory Global Health Institute and in collaboration between Emory University and the International Center for Diarrheal Disease Research in Bangladesh.

Chapter 5

The Mathematics of Adaptive Behavior in an Economic-Epidemiological Model [99].

Adaptive behavioral changes that affect interactions and mixing rates among individuals play a crucial role in determining the level of incidence, rate of spread, and the overall dynamic path of an epidemic. Individuals likely alter their behaviors in response to current epidemic conditions in order to prevent themselves from getting sick, but these behavioral changes are mediated by the individual's assessment of the role of factors that contribute to infection risk (e.g. activity, vaccination status, contact networks). Certain behavioral modifications such as total individual isolation or celibacy, in the case of sexually-transmitted diseases, would eliminate the risk of individual infection. However, such behavior is likely to be excessively costly to individuals. In general, individuals value health as an input to generating utility, an index of satisfaction or wellbeing. Most mathematical models of disease transmission do not explicitly incorporate utility, or goal-seeking behavior, which leads to adaptive behavioral changes [7, 30, 43, 48, 67, 86, 102]. By focusing on goal seeking models, we are able to capture truly endogenous adaptive behavioral responses. This level of analysis, within the contact structure of a population, is only beginning to emerge in epidemiological models [39, 58, 123].

The "co-evolutionary" dynamic between disease prevalence and behavioral response is a quintessential complex adaptive system, systems with non-linear interactions at multiple scales [97]. In models without the behavioral adaptive dynamic approach, one considers dynamics on a static landscape - the phase dynamics can be completely characterized irrespective of the initial condition of the system- where the future course of the dynamic flow can be well described without precise knowledge of the initial conditions. Acknowledging adaptive behaviors means acknowledging a

feedback where the landscape affects the evolution of the epidemic and the evolution of the epidemic affects the topological landscape and phase dynamics (see [59, 74] for more discussion of these types of feedbacks). The implications of this feedback can be profound for empirical work in addition to theory. Geoffard and Philipson note that while estimates of parameters in nonadaptive epidemiological models are quite robust, even with the proliferation of compartments, once adaptive behavior is introduced the robustness of estimation, and identification of mechanism, is lost [62].

Recent experiences with epidemics such as SARS [42, 43], avian influenza [124], and ebola [44] demonstrate the role of behavior in both the spread and control of the epidemics. Policy response to these outbreaks, aimed at altering pathogen dynamics, resulted in public interventions that had large impacts on the socio-economic landscape (i.e. the collection of individual social and economic statuses within the population). A rigorous theoretical epidemiological framework for modeling how human decisions, related to intentional and adaptive goal-seeking behavior, shape disease dynamics is needed to capture the influence of the protective behaviors induced by the fear of emergent or re-emergent infectious pathogens (e.g. HIV) [7, 21, 30, 32, 64, 65, 67, 69, 84, 90, 112] or from the potential deliberate release of biological agents [12, 48].

Recently, public health officials have systematically employed travel restrictions and social distancing measures to reduce disease spread [16, 40, 42, 43, 47, 56]. Policies quarantining or restricting contact among individuals may lead to the greatest reduction in cases but the implementation (particularly over sustained periods in time) of such extreme policies may induce other unforeseen private and social costs [128]. Enactment of contact related policies may close schools, restrict social and cultural activity, and even shut down major metropolitan areas. These heavy handed policies had serious impact on local and global economies specifically during the H1N1 pandemic [47, 124].

Related literature [9, 39, 58, 61, 63, 115, 117] posits that forward looking individuals aim to maximize an objective function that includes, but is not limited to, health over a planning horizon. A utility function is specified that includes health status and other goods. Maximization of this utility function induces individuals to make tradeoffs between long term health and the short term costs of avoiding infection. In these studies optimization is used as a way of modeling the goal seeking adaptive decisions of members of society (important to the understanding of behaviorally driven disease dynamics) and not for engineered public health interventions. Such understanding can ultimately help develop better normative (policy driven) disease intervention strategies (e.g. [61]). Fenichel et al. [58] simulate the effects of an individual decision making model and illustrate the implications of adaptive behavior for reproductive number theory and disease dynamics within the compartmental framework. Moreover, they show the potential for policy changes that alter the benefits and cost of disease avoidance lead to oscillatory dynamics (sometimes called waves in the epidemiological literature [40, 96]).

We provide a general mathematical foundation for including adaptive human behavior in epidemiological models by incorporating work by Blythe, Castillo-Chavez and Cooke [25] into a behavioral framework where individuals have a short-term payoff from making contacts with others. Individuals trade off between the increase in utility that results from increased contacts with the risk that additional contacts could lead to future utility loss through infection. Forgoing contacts in the present is similar to investing in future health capital as found in Fenichel *et. al.* [58]. We aim to combine the analyses of the general, nonlinear interactions between individuals during the course of an epidemic with the individual based optimization framework to construct an approach to modeling that is mechanistic with respect to social and economic considerations, *economic epidemiology*.

In Section 5.1 we develop a traditional, and general, model for the transmission of an

influenza like disease with nonlinear incidence. The economic considerations are described in detail and introduced into the model to serve as the new *adaptive behavior model*. The major difference between the two models lies in the individual's ability to adjust behavior in the economic model while all contact rates (i.e. behaviors) are fixed in the classical formulation. This serves to highlight the mathematical differences induced by the changing behavioral landscape as well as the possible richness of behavior that may arise when explicitly considering behavior. Section 5.2 compares outcomes between the two models and details the application of the theorems contained in [25] to the adaptive behavior model. Section 5.3 raises new questions about the implementation of the individual behavior regime and discusses the results of this paper.

5.1 Mathematical Formulations

Incorporating adaptive human behavior into mathematical models introduces a number of nonlinearities thereby drastically increasing the model's complexity. Here we investigate one of the simplest models describing influenza like dynamics in order to outline a basic technique of incorporating adaptive behavior. We divide the population into three compartments based on disease state to describe classic *SIR* disease dynamics [86]: susceptible to the disease, S; infected and infectious, I; recovered and permanently immune, R. Individuals are added to the susceptible class at the constant rate Λ and are removed from each health class at the per-capita rate μ. Disease recovery is modeled via a constant per capita rate γ (with $1/\gamma$ being the average length of infectiousness). Infection incidence within a population can be described as the product of four terms: the per-capita average number of contacts, c; the probability that a contact between a susceptible and infectious individual results in a new infection, β; the number of susceptible individuals who may become infected, S; and a nonlinear function $F(S,I,R)$ describing how the presence of a disease affects incidence. We assume that all system parameters are strictly positive, and that the

population mixes proportionately resulting in an incidence function given by

$$B = c\beta F(S,I,R)S\frac{I}{N}, \tag{5.1}$$

where $N = S+I+R$ as in [25]. This results in a model given by

$$\begin{aligned} \dot{S} &= \Lambda - c\beta SF(S,I,R)\frac{I}{N} - \mu S, \\ \dot{I} &= c\beta SF(S,I,R)\frac{I}{N} - \gamma I - \mu I, \\ \dot{R} &= \gamma I - \mu R. \end{aligned} \tag{5.2}$$

The number and stability of equilibrium points and whether oscillatory solutions exist follow from the basic reproductive number paradigm often applied to compartmental disease models [29].

Nonadaptive, Nonlinear Theorems

First we characterize the stability of the disease free equilibrium (DFE). Common practice is to characterize the stability of the DFE through an appeal to the concept of the basic reproductive number $\Re_0$. This quantity is interpreted as the average number of secondary infections a typical infectious individual causes in a fully susceptible population. It is known that the biological interpretation of this typical threshold quantity breaks down in more complicated models [70, 118]. If $\Re_0$ is less than unity, then the disease dies out, and the DFE is locally stable. If $\Re_0$ exceeds unity, in systems with recruitment of new susceptibles, then the rate of infection results in an endemic equilibrium level of infection, and the DFE is locally unstable. For the purpose of our analysis we assume that asymptotically the entire population reaches an equilibrium value[1] of $\frac{\Lambda}{\mu}$.

Theorem 5.1.1 ([25])**.** *If*

$$0 \le F(S,I,R) \le F\left(\frac{\Lambda}{\mu},0,0\right) \le F(\infty,0,0) = 1 \tag{5.3}$$

[1] A result easily arrived at when solving the ordinary differential equation for $\dot{N} = \dot{S}+\dot{I}+\dot{R}$.

and

$$\mathfrak{R}_0 := \frac{\beta c}{\gamma+\mu} F\left(\frac{\Lambda}{\mu}, 0, 0\right) < 1, \tag{5.4}$$

then the disease free equilibrium of System (5.2) attracts all local solutions, that is,

$$\lim_{t\to\infty}(S(t), I(t), R(t)) = \left(\frac{\Lambda}{\mu}, 0, 0\right).$$

If $\mathfrak{R}_0 > 1$*, then the disease free equilibrium is locally unstable.*

The interesting case is when the DFE is unstable. The existence of oscillations of the susceptible and infected populations is possible due to the recruitment mechanism of new susceptible individuals into the population. We can extend previous local stability claims to global ones by ruling out oscillatory behavior. To facilitate analysis of oscillatory solutions, assume that the total population has stabilized[2], e.g. $S(t)+I(t)+R(t) = \Lambda/\mu$. Applying Dulac's Criteria (e.g. see [33] for some examples and an extension) to the resulting planar system results in

Theorem 5.1.2. *[25] If* $\frac{\partial F}{\partial S} > \frac{\partial F}{\partial I}$ *for* $S > 0, I > 0$*, then System (5.2) has no limit cycles (e.g. oscillatory solutions) in the positive cone.*

The disease may persist at several endemic levels even if limit cycles have been ruled out due to the nonlinearities in F. An additional theorem supplies sufficient conditions for unique equilibria.

Theorem 5.1.3. *[25] Given System (5.2) if* $\frac{\partial F}{\partial S} \geq 0, \quad \frac{\partial F}{\partial I} \leq 0, \quad \frac{\partial F}{\partial R} \leq 0$, *and* $\mathfrak{R}_0 > 1$ *then System (5.2) has a unique endemic equilibrium.*

The condition on $\frac{\partial F}{\partial R}$ in Theorem 5.1.3 implies that the existence of recovered individuals does not result in increased transmission [25]. These theorems may be used to illustrate the effects adaptive behavior can have on the course of an epidemic.

[2]The assumption that the demographically limited dynamics match those of the original model is made valid in [38]. This does not imply that for the remainder of the paper we assume the demographic limit. Indeed, other than when discussing limiting behavior, N is varying with time.

Incidence Functions and the Adaptive Behavior Model

In epidemiological models frequency dependent mixing, or standard incidence, occurs when contact rates are assumed independent of population density. Classical standard incidence can be expressed as $F(S,I,R) = 1$, yielding an incidence function of $B = c\beta S\frac{I}{N}$.

Individuals may alter contact behavior over time with respect to the amount of infectious individuals within the population, I. Adaptive behavior implies that the rates c or β are not constant, but functions of I and potentially other state variables. If people behave adaptively, then the observed population level dynamics of an epidemic emerge from individual decision making. We abstract the measure of wellbeing, or benefit, an individual gains in the process of interacting with others (money, enjoyment, etc...) as utility. Individuals aim to minimize the loss of utility, that comes from becoming infected, during the course of an epidemic. Forgoing contacts reduces the probability of infection, but results in a loss of otherwise beneficial social contacts that may lead to infection. An individual may possibly reduce the intensity or alter the nature of activities during an epidemic to reduce the risk of infection, but in this work we only suppose social distancing, i.e. β is left constant.

The tradeoff between gaining utility through current period contacts and gaining utility through avoiding future infection implies an optimal individual strategy that adapts as the state of the epidemic changes. We adopt the phrase *utility maximization* to describe the strategies used by purposeful goal-seeking individuals to adaptively manage the benefit-risk tradeoffs tied to contact activities, a phrase that is commonly used in economics literature (e.g. Mas-Colell et. al.; Dixit and Pindyck; Adda and Cooper) [3, 52, 95].

Individuals experience a marginal increase in utility up to some point as a result of

making contacts within a unit time. We let $f_t := f(c_t)$ denote the utility function associated with the time interval $[t, t+1]$ so that f_t models the utility realized from a certain number of contacts over the unit time interval at t where c_t models the number of contacts made within this interval. Assume that $\frac{\partial f_t}{\partial c_t} > 0$ (monotonically increasing) and $\frac{\partial^2 f_t}{\partial (c_t)^2} < 0$ (concave) up to point c_t^* , where c_t^* denotes the optimal number of contacts within the selected time window $[t, t+1]$ when the disease is absent from the population. It is further assumed that $\frac{\partial f_t}{\partial c_t} < 0$ for any number of contacts greater than those achieved at c_t^*. The cost of making contacts in excess of c_t^* is prohibitive. We set $f_t(0) = 0$ and impose no further functional restrictions on f_t.

We assume the health status of an individual influences his utility function directly in situations where the population faces an epidemic outbreak. An individual's status is indicated via the subscript $m \in \{s, i, r\}$ (susceptible, infected and recovered) and it is assumed that individuals with different statuses benefit from contacts differentially. We allow the status of individuals to be an argument in the utility function, $f_t(c_{t|m}, m)$, and assert that for a given number of contacts, $c_{t|m}$, that $f_t(c_{t|i}, i) < f_t(c_{t|s}, s) \leq f_t(c_{t|r}, r)$ where f is strictly a measure of the net benefit one receives from contacts during a given time period.

In this framework all individuals are perfectly informed about the current state of the epidemic and seek to maximize their individual utility. Moreover, we assume that individuals do not care about the health of others and take the behavior of others as given; there is an absence of strategic behavior among individuals who are not susceptible to infection. Utility maximization provides a parsimonious model of goal-seeking behavior[3]. Infected or recovered individuals do not have incentives to modify behavior away from the behavior under disease free conditions as recovered individuals are immune to the disease, and we assume contacts do not affect recovery

[3]We are not making normative judgements about what that goal should be or what the individual should do to maximize his utility. Rather we are asserting that optimization provides a positive model of individuals making trade-offs.

from infection. Individuals in classes i and r choose the number of contacts per day that maximizes individual utilities, $c^*_{t|i}$ and $c^*_{t|r}$, respectively. Only susceptible individuals modify their behavior in response to changes in disease prevalence. In summary, a susceptible individual is made better off by increasing contacts over a time interval $[t, t+T]$ all else equal, where T is an arbitrary planning horizon (e.g., the upper limit of the long-term thinking associated with the individuals in the population). However, susceptible individuals face incentives to reduce contacts below the optimal disease free contact level in order to mitigate infection risk. To analyze the intertemporal tradeoff, we construct the expected utility function over the time horizon T. The expected utility function is a measure of the present and the future value of these contacts to the individual. The contacts that a user makes in the current time period have the utility (value) described above, $f_t(c_{t|s}, s)$, but also impact the expected future utility value over the remaining time horizon. As the number of infected individuals increases the expected future-utility value decreases since the probability that a susceptible individual becomes infected increases and $f_t(c_{t|i}, i) < f_t(c_{t|s}, s)$. Future utility is discounted by a factor $\delta \in [0,1]$ to account for the rate of time preference, an individual's relative preference for goods today relative to tomorrow. The discount factor serves as weight to make future utility units equal to present utility units. Thus over a planning horizon $[0, T]$ the expected utility for a susceptible, conditioned on the individual's future state, is

$$\mathbb{E}(U) = f_0(c_{0|s}, s) + \sum_{t=1}^{T} \delta^t \mathbb{E}^m(f_t(c_{t|m}, m_t)).$$

In order to arrive at an expression for the expected utility for an infected individual, we let v denote the time of infection and $\rho := \lceil 1/\gamma \rceil$ be the expected length of infection in the relevant time units rounded to the nearest integer above $1/\gamma$. Thus the expected utility for an infected individual is $\mathbb{E}(U) = \sum_{t=v}^{v+\rho} f_t(c^*_{t|i}, i)$. Note, we use $c^*_{t|i}$ because the infected individual aims to maximize his utility , but infection does not create incentives to decrease contacts. Furthermore, for a recovered individual the

utility for every time interval after recovery is $f_t(c^*_{t|r}, r)$.

The differentiated contact structure induced by the adaptive behavior grants a more complicated form to the standard incidence. Over the time interval $[t, t+1]$ the incidence is

$$B = c_{t|s}\beta S \frac{c^*_{t|i} I}{c^*_{t|s} S + c^*_{t|i} I + c^*_{t|r} R}. \tag{5.5}$$

where $c_{t|s}$ denotes the number of contacts susceptible individuals make; β is defined as before; and $\frac{c^*_{t|i} I}{c^*_{t|s} S + c^*_{t|i} I + c^*_{t|r} R}$ is the proportion of these contacts with an infected individual in randomly mixing population. Consequently, we may define the state-dependent function F as

$$F(S,I,R) = \frac{c_{t|s} c^*_{t|i}}{c^*_0} \frac{S+I+R}{c^*_{t|s} S + c^*_{t|i} I + c^*_{t|r} R}, \tag{5.6}$$

where c and c^*_0 have contextually analogous definitions (i.e. the average number of contacts at time zero of the epidemic). We therefore have two examples of the model described in System (5.2): one where $F(S,I,R) = 1$ and a second where F is defined by Equation (5.6). Noting that $c = c^*_0$ we may rewrite System (5.2) to reflect the inclusion of adaptive behavior via

$$\begin{aligned} \dot{S} &= \Lambda - c^*_0 \beta F(S,I,R) S \frac{I}{N} - \mu S, \\ \dot{I} &= c^*_0 \beta F(S,I,R) S \frac{I}{N} - (\gamma + \mu) I, \\ \dot{R} &= \gamma I - \mu R. \end{aligned} \tag{5.7}$$

5.2 Results

Theorem 5.1.1 implies

$$\mathfrak{R}_0 := \frac{\beta c^*_0 F(N,0,0)}{\gamma + \mu} = \frac{\beta c_{0|i}}{\gamma + \mu}$$

where time $t = 0$ is the time of introduction of the first primary case. A problem with this basic reproductive number is that it does not account for the behavioral adaptation of the susceptible population. If the risk of infection, β, and/or the period of infection,

$\frac{1}{\gamma}$, were large enough, then a single infected individual may cause all susceptibles to reduce contacts to 0. In this situation of extreme behavioral change the epidemic threshold of $\Re_0$ is not reflective of the system dynamics. In order to rule out oscillatory solutions in a system with adaptive behavior through application of Theorem 5.1.2 we must show that $\frac{\partial F}{\partial S} > \frac{\partial F}{\partial I}$. Let $D := c_{t|s}S + c_{t|i}I + c_{t|r}R$ and compute the difference

$$\frac{\partial F}{\partial S} - \frac{\partial F}{\partial I} = \frac{c^*_{t|i}}{c^*_0 D^2}\left\{\left[\frac{\partial c_{t|s}}{\partial S} - \frac{\partial c_{t|s}}{\partial I}\right](c^*_{t|i}I + c^*_{t|r}R)N + c_{t|s}(c_{t|s} - c^*_{t|i})(S - I - R)\right\}.$$

The first term in brackets is positive by the properties of $c_{t|s}$, the number of susceptible contacts increase with S and decrease with I. The sign of the second term, $(c_{t|s} - c_{t|i})(S - I - R)$, is ambiguous and requires further specification of the utility functions. A sufficient condition for $\frac{\partial F}{\partial S} - \frac{\partial F}{\partial I} > 0$ is that $S < I + R$ if and only if $c_{t|s} < c^*_{t|i}$. That is, in order to rule out oscillatory dynamics, it is sufficient to note that when the total population that is infected, or has been infected, exceeds the susceptible population then the susceptible individuals must each make fewer contacts than infected individuals. This condition may be easily violated heuristically. It requires that in the presence of a large recovered population, very small infection levels would induce susceptible individuals to make few contacts. This contradicts intuition associated with the utility maximization problem. The preceding description of individual behavioral does not allow for the presence of recovered individuals to have a negative impact on susceptible behavior. Alternatively, let $\hat{F}(S,I,R) = \frac{c_{t|s}}{c^*_0}$, and apply Dulac's Criteria directly to the planar system

$$\begin{aligned} \dot{S} &= \Lambda - c^*_0 \beta \hat{F}(S,I,R) S c^*_{t|i} \frac{I}{G} - \mu S, \\ \dot{I} &= c^*_0 \beta \hat{F}(S,I,R) S c^*_{t|i} \frac{I}{G} - (\mu + \gamma) I, \end{aligned} \tag{5.8}$$

where $G = c_{t|s}S + c^*_{t|i}I + c^*_{t|r}\left(\frac{\Lambda}{\mu} - S - I\right)$. This calculation gives the sufficient conditions that limit cycles do not exist provided that for all t greater than a finite t^*, $c_{t|s} < c^*_{t|i}$ and $\frac{\partial \hat{F}}{\partial S} > 0$ and $\frac{\partial \hat{F}}{\partial I} < 0$. To better understand this condition consider two

cases: a large endemic population and a small endemic population. If the steady state for the infected class, I_∞, were large then requiring $c_{t|s} < c^*_{t|i}$ for all t larger than some t^* is reasonable, because the susceptible individuals would be actively trying to avoid becoming infected. However, for a small I_∞ this requirement would imply some memory (a non-Markov behavioral model) linked to the infection that involves the susceptible population avoiding a second epidemiological peak (i.e. $\frac{dI}{dt} > 0$ after a period where $\frac{dI}{dt} < 0$). Theorem 2 does not rule out oscillatory behavior for the adaptive system in general, particularly at low endemic infected levels. This result provides a more mechanistic insight to Brauer et al.'s [21] result that information may destabilize a system (induce oscillatory behavior) when dynamics would otherwise be stable in the absence of information.

To illustrate adaptive behavior's effect on the epidemiological system we numerically solved System (5.2). This was done with $F(S,I,R) = 1$ to illustrate results from a traditional epidemiological model and with F as defined in Equation (5.6) to compute solution curves $S_e(t), I_e(t)$, and $R_e(t)$ (Figure 5.1). The utility functions are defined by the expressions $f_t(c_{t|s}, s) = f_t(c_{t|r}, r) = (bc_{t|m} - c^2_{t|m})^\nu$ with $b = 24$ and $\nu = .2$. We let $f_t(c_{t|i}, i) = 0$ and $\delta = .9986$. The epidemic parameters used are $\Lambda = 10, \mu = \Lambda/10000$ and $\beta = \gamma = 0.2$; and the initial conditions are set at $(S(0), I(0), R(0)) = (9999, 1, 0)$.

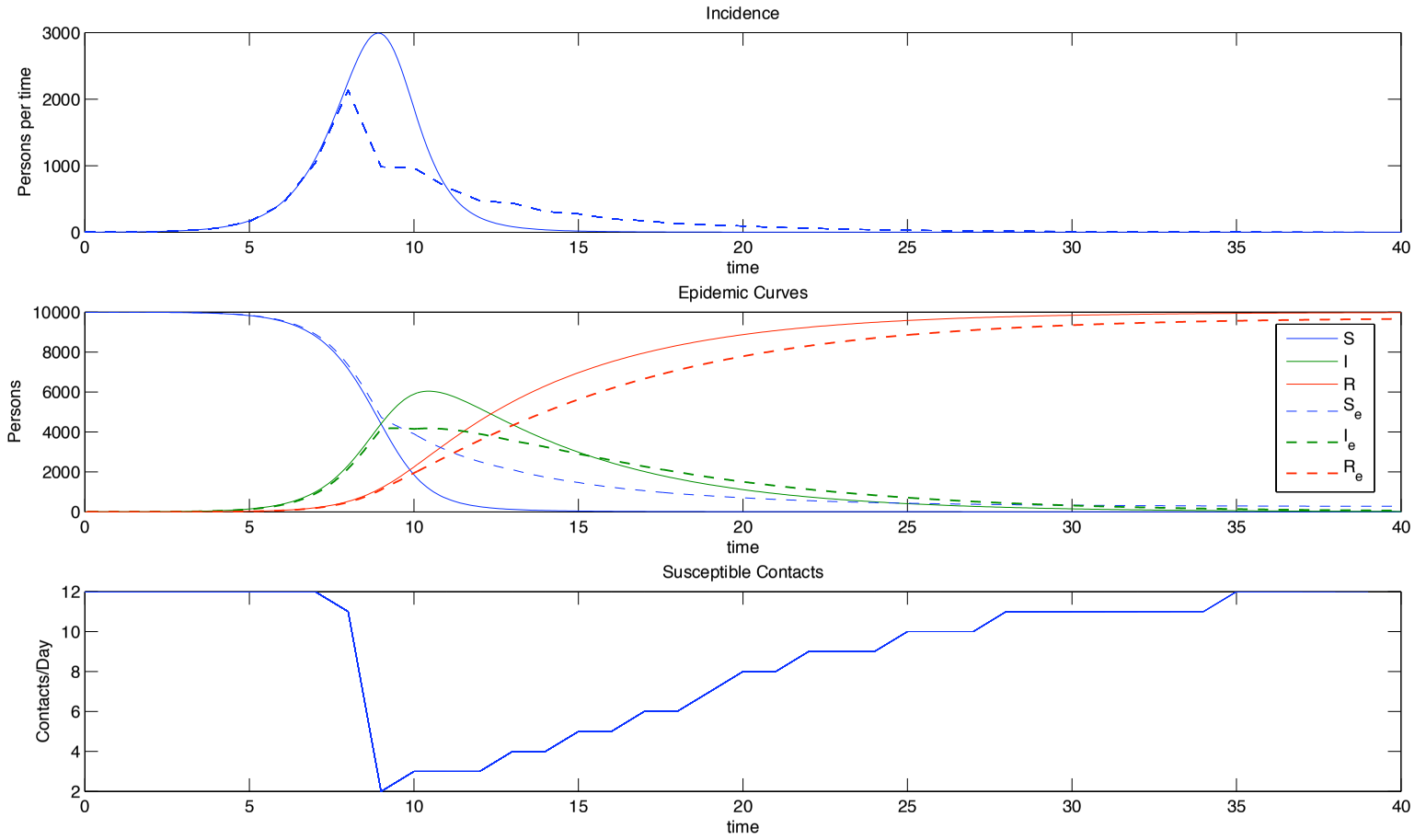

Figure 5.1: The three figures top to bottom show the incidence B (solid for F=1 and dashed for adaptive behavior), the epidemic curves S, I and R, and the contacts per day for the economically altered behavior, $c_{t|s}$.

We may summarize the effect of adaptive behavior in terms of its impact on the epidemic peak, the endemic state, and how its structure filters to system incidence. The epidemic peak, the largest number of infected individuals at a given time over the course of the epidemic, is 31% lower with adaptive behavior than in the standard epidemiological model. This reduces the immediate impact on healthcare services who may only have a few days to prepare from first reported case to peak infection level. The model with adaptive behavior results in both a slightly larger susceptible and infected endemic population than the standard model. This immediately appears counterintuitive, because the adaptive behavior returns to its pre-epidemic level by day 35 and thus the course of the adaptive epidemic should match that of non adaptive one. However, as we've shown earlier we cannot guarantee the uniqueness of the endemic equilibria for the adaptive model (of which the standard model is really a special case) and thus we may take this example as evidence of nonuniqueness (for other examples see [25, 35, 37, 55, 57, 67, 77]).

Finally, in terms of contacts per day, the susceptible behavior mirrors the infection level curve. There are a few features to note in this example: the behavior strongly influences only a finite period of time (day 8 to day 35) and that the length of periods of identical behavior (consecutive days with identical number of contacts chosen) are not monotonically increasing with $I_e(t)$'s decrease, a possibly very complex/ nonmonotonic behavioral response. If one changes Λ to 1000, by the definition of μ this does not affect the total population size and simply increases rate of the recruitment/removal in the system, a behavior structure that exhibits sustained oscillations in incidence, see Figure 2, may be observed. Interestingly in this oscillating example we have a sustained susceptible population that is large enough to support both a recovered and infectious population that are below the standard model's levels; thus, endemic prevalence in the adaptive behavior model is necessarily less than in the standard case.

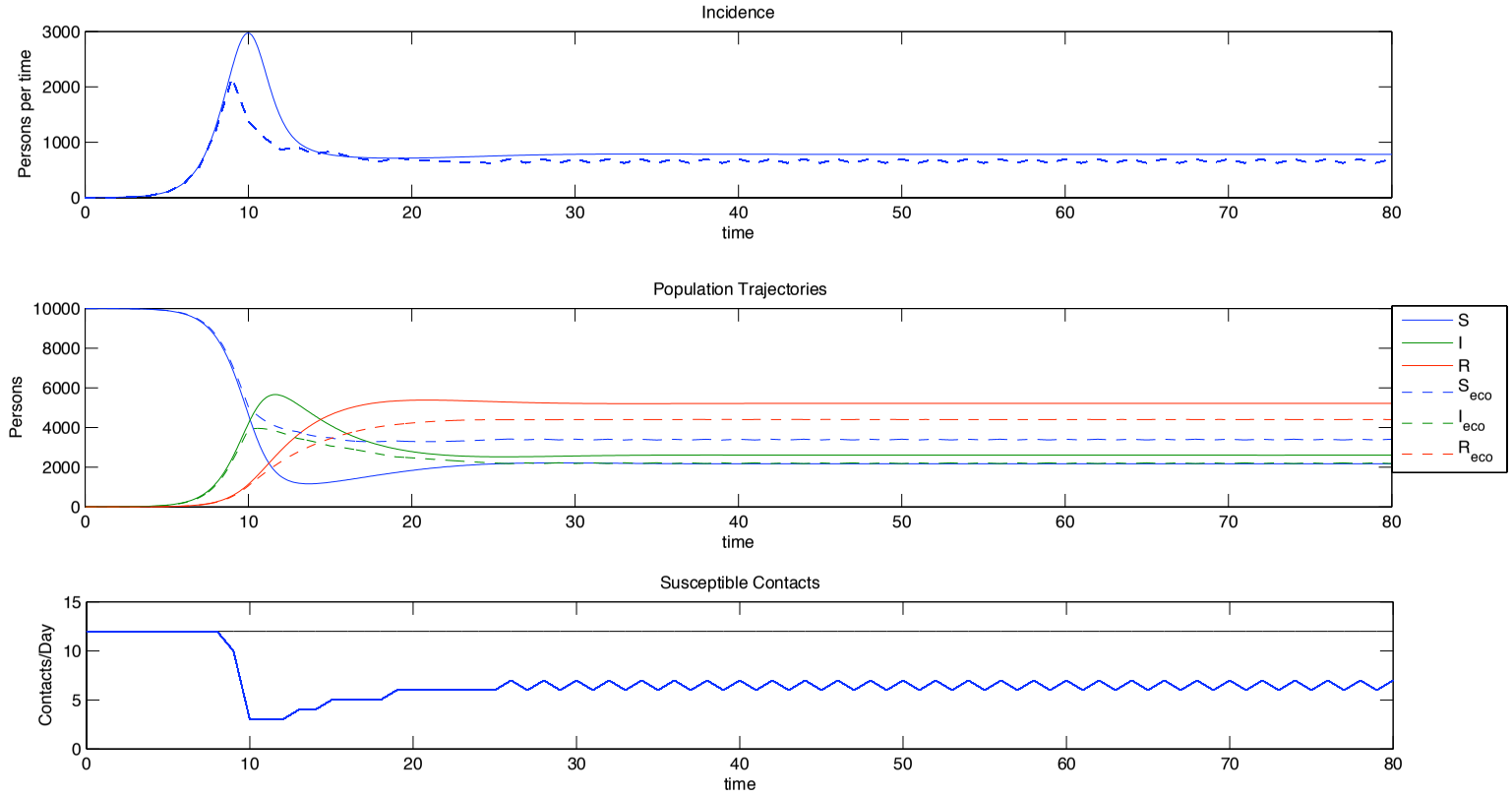

Figure 5.2: The three figures top to bottom show the incidence (solid for F=1 and dashed for adaptive behavior), the epidemic curves, and the contacts per day for the economically altered behavior.

5.3 Discussion

Public policy has focused on prevention and containment. However, public health officials must carry out these tasks in the context of challenges that include the use of reported data (biased towards severe cases and of highly variable quality, particularly at the scales of interest), the lack of "real-time" surveillance systems [19], and the absence of theoretical frameworks that assess the role of individual decisions. Infection within a population creates economic incentives that result in adaptive decisions at multiple levels of social organization and over various temporal scales [58]. Individuals gain utility from making contacts with others, but each contact incurs additional risk or exposure to disease. Tradeoffs between increased utility in the present and the risk that such contact could lead to future utility loss through infection occur on the individual level. These decisions manifest as missed work time, reduced productivity and health care expenses that add to the social cost of disease. In a sense individuals are involved in a dynamic game choosing strategies comprised of current and future contacts with payouts described by expected utility [117]. The actual strategies employed may involve a degree of commitment (i.e. open-looped in the short-term) or may involve regularly updated reaction function (i.e. closed looped, Markov perfect strategies) [130]. This perspective is different than viewing the control process as an effort to reduce the total number of infected.

Central to the difficulties with implementing the decision making process into the epidemic model, and thus applying the theory of nonlinear incidence, is that we lose a consistency of mixing, an important assumption in most analyses. However, the assumption that individuals of different health statuses behave identically seems overly strong. Furthermore, within the utility maximization framework the expected utility

$$\mathbb{E}(U^*) = \max_{c_{t|\star}} \left\{ f_0(c_{0|s}, s) + \sum_{t=1}^{T} \delta^t \mathbb{E}^m(f_t(c_{t|m}, m_t)) \right\}$$

must be maximized over some time horizon T, which may be infinite in principle.

This is typically reconciled via dynamic programming [3], solving backwards along an "optimal" path to arrive at the decision to be made in the now. Implementing this idea into epidemic systems is difficult because there are at least three time scales to consider: the epidemic scale, the decision scale, and the information arrival rate. As was shown in the contact plot within Figure 5.1 the period over which strategies changed were not necessarily days and did not adhere to rates of change of the system in directly identifiable ways. An appropriate and realistic time scale may be a time period of τ (that may be random) that models how long it takes to disperse information as a function of factors such as disease prevalence and severity. In this paper, we set the planning horizon to $T = 1$, and time step to 1. Then while numerically solving the ODE system (5.2) we stop at each integer time value and recalculate the expected utility over the next unit time interval. Simulating in this fashion has shown that the economic behavior can induce oscillation, in both the behavior and the epidemic trajectory, and may maintain a susceptible population at a much higher level than without the behavioral adjustment (with appropriate parameter values our simulation models have generated infection level a whole order of magnitude less). Longer time scales, likely the case with the proposed random interval model, are expected to induce similar behavior over a broader parameter range.

Over any planning horizon greater than 1, indicative of the time period over and the frequency with which decisions are made, the probabilities of being in particular states become very complicated to compute, and thus the expected utility function becomes impractical to even formulate numerically. For example if we try to compute the probability an individual remains infectious throughout the third day there are many ways he could have come to be infectious at some point in the third day and for each path, the path of another individual's infection changes (e.g. sick on day one and remain as such, be well on day one, sick on day two and remain sick or well on day one and two and become sick on day three). A time step with biological significance

that removes some of this path difficulty, but is not natural for decision making, is that of *event times*. The process then has stochastic time steps and the susceptible individuals would update their behavior as events (new infection, new recovery, new individual enters the system or an individual leaves the system) occur. Without exploring this method here it should be clear that this would produce far fewer infections than the methods in this paper because there would be a more rapid dissemination of information implied. This illustrates a critical new frontier yet to be resolved in epidemiology or the study of complex systems more generally - the need for a general way to address temporal scaling issues.

The complex adaptive system generated by the introduction of the economic behavior described here may reduce an epidemic's forecasted size and alters forecasts to suggest a spreading out of the peak of the epidemic over time while lessening its severity. This implies that individual pathogens may actually be more biologically infectious than currently believed. To strengthen the results, in order to make policy decisions using the ideas of utility maximization, a great deal of work should be put into estimating the form of $f_t(c_{t|m}, m)$ as the numerical results are sensitive to its shape. The introduction of differential contact qualities with different payouts and risks would add a level of realism and applicability (e.g. family contacts versus work contacts or monogamous versus polygamous). In addition, population wide policy decisions, such as closing public transportation, may also affect the tradeoffs with respect to $c_{t|s}$. Such policies could have unintended consequences (i.e. forming reservoir susceptible populations that may produce second epidemic peaks) if we do not explicitly consider the adaptive nature of human behavior. Despite all the challenges involved in such a complicated model we have been able to use previous techniques to prove stability and complexity of fixed points for the system, and proofs of qualitative behavior under a delay in information is underway again using theory from [25]; for completeness we have outlined the relevant theory here in the

appendices. Numerically we've been able to show that the entry and removal in the system may be used as a control, with the economic behavior structure, to destabilize and induce oscillatory behavior. To advance the applicability of epidemiological models it is imperative that we move from thinking of individuals as passive particles to beings that actively attempt to shape their own futures. In so doing, the mathematics becomes more challenging, but we enhance our chance of explaining complex disease dynamics with parsimonious models.

5.4 Acknowledgements

This project has been partially supported by grants from the National Science Foundation (NSF - Grant DMS - 0502349), the National Security Agency (NSA - Grant H98230- 06-1-0097), the Alfred T. Sloan Foundation, the National Institute of Health (NIH - Grant 1R01GM100471-01), and the Office of the Provost of Arizona State University. This work was also conducted as part of the SPIDER Working Group at the National Institute for Mathematical and Biological Synthesis (NIMBioS), sponsored by NSF, the U.S. DHS, and USDA through NSF Award #EF-0832858, with support from The University of Tennessee. This publication was made possible by grant number 1R01GM100471-01 from the National Institute of General Medical Sciences (NIGMS) at the National Institutes of Health. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of NIGMS.

5.5 Proof of Theorem 5.1.1

Suppose the conditions of the theorem hold (e.g. Equations (5.3) and (5.4)). If $I(0) = 0$, the solution has $I(t) = 0$ for all $t \geq 0$ as we see from (5.8), and therefore

$R(t) \to 0, N(t) \to \Lambda/\mu$, and $S(t) \to \Lambda/\mu$. If $I(0) > 0$ then $S(t) \leq N(t)$ and

$$\begin{aligned}
\frac{1}{\gamma+\mu}\frac{dI}{dt} &= \left[\frac{\beta c}{\gamma+\mu}F(S,I,R)\frac{S}{N} - 1\right]I \\
&\leq \left[\frac{\beta c}{\gamma+\mu}F(S,I,R) - 1\right]I \leq \left[\frac{\beta c}{\gamma+\mu}F\left(\frac{\Lambda}{\mu},0,0\right) - 1\right]I \\
&= (\mathfrak{R}_0 - 1)I < 0.
\end{aligned}$$

Since $I(t)$ is decreasing, $\lim_{t\to\infty} I(t) = 0$. Then the variation of parametric formula gives

$$R(t) = R(0)e^{-(\delta+\mu)t} + \gamma\int_0^t I(s)e^{-(\delta+\mu)(t-s)}ds,$$

and it follows that $R(t) \to 0$ as $t \to 0$. And since $N(t)$ tends to Λ/μ, we deduce that $S(t) \to \Lambda/\mu$.

The Jacobian matrix of (5.8), with derivatives evaluated at $I = R = 0, S = N = \Lambda/\mu$, is

$$\begin{bmatrix} -\mu & -\beta c F_0 & 0 \\ 0 & \beta c F_0 - (\gamma+\mu) & 0 \\ 0 & \gamma & -\mu \end{bmatrix}.$$

The eigenvalues are the diagonal entries. Hence, the disease-free equilibrium is unstable if $\beta c F_0 > (\gamma+\mu)$ or equivalently when $\mathfrak{R}_0 > 1$. This completes the proof of Theorem 1. □

5.6 Proof of Theorem 5.1.2

Let $g_1(S,I)$ and $g_2(S,I)$ be the functions in the right members of (5.8) with $N = \frac{\Lambda}{\mu}$ (i.e. the autonomous version), that is

$$\frac{dS}{dt} = g_1(S,I), \quad \frac{dI}{dt} = g_2(S,I).$$

Then

$$\begin{aligned} \frac{\partial}{\partial S}\left\{\frac{g_1(S,I)}{SI}\right\} + \frac{\partial}{\partial I}\left\{\frac{g_2(S,I)}{SI}\right\} &= -\frac{\Lambda\mu}{\mu S^2 I} - \frac{\beta c\mu}{\Lambda}(F_S - F_R) + \frac{\beta c\mu}{\Lambda}(F_I - F_R) \\ &= -\frac{\Lambda}{IS^2} + \frac{\beta c\mu}{\Lambda}(F_I - F_S). \end{aligned}$$

Clearly, the first term is negative. The second term is negative by hypothesis. Thus the expression is of fixed sign in the region $S > 0, I > 0, S + I \leq \Lambda/\mu$, and it follows from Dulac's Criterion test that (5.8) has no limit cycles in the region.

5.7 Proof of Theorem 5.1.3

Define

$$G(\eta) := \mathfrak{R}_0 F(h_1(\eta), h_2(\eta), h_3(\eta)) - F_0\eta, \tag{5.9}$$

where F is the nonlinear incidence function, $\mathfrak{R}_0 = \frac{\beta c}{\gamma+\mu} F\left(\frac{\Lambda}{\mu}, 0, 0\right)$ is the basic reproduction number, $F_0 = F\left(\frac{\Lambda}{\mu}, 0, 0\right)$, $\eta = \mathfrak{R}_0 \frac{F(S^*,I^*,R^*)}{F\left(\frac{\Lambda}{\mu},0,0\right)}$, and $h_i(\eta)$ are expressions of the equilibrium in terms of η. More specifically consider the equilibria of (5.2) which satisfy

$$\begin{aligned} \Lambda - \mu S^* &= c\beta F^* S^* \frac{I^*}{N^*}, \\ c\beta F^* S^* \frac{I^*}{N^*} &= (\gamma+\mu) I^*, \\ \gamma I^* &= \mu R^*, \end{aligned}$$

where $F^* = F(S^*, I^*, R^*)$ and $N^* = \frac{\Lambda}{\mu}$. Supposing that $I^* \neq 0$ we have $\frac{S^*}{N^*} = \frac{\gamma+\mu}{c\beta F^*} = \frac{1}{\eta}$. Letting $q = \frac{\gamma}{\mu}$ it also follows that $\frac{I^*}{N^*} = \frac{1}{1+q}\left(1 - \frac{1}{\eta}\right)$ and $\frac{R^*}{N^*} = \frac{q}{1+q}\left(1 - \frac{1}{\eta}\right)$.

We therefore may write that

$$G(\eta) = \mathfrak{R}_0 F\left(\frac{N^*}{\eta}, \frac{N^*}{1+q}\left(1 - \frac{1}{\eta}\right), \frac{N^* q}{1+q}\left(1 - \frac{1}{\eta}\right)\right) - F_0\eta.$$

As $\eta \to 1$ it is easy to see that $F \to F_0$ and thus $G(\eta) = F_0(\mathfrak{R}_0 - 1) > 0$. Similarly, given the conditions of F given in Theorem 5.1.1 we have that for $\eta > \mathfrak{R}_0$ that $G(\eta) < 0$. Thus by continuity there is at least one equilibrium. For uniqueness we may assume that $\mathfrak{R}_0 > 1$ and look for conditions for which $\frac{dG}{d\eta} < 0$. Straightforwardly one may show that

$$\frac{dG}{d\eta} = \mathfrak{R}_0 \frac{N^*}{\eta^2}\left(-\left(\frac{\partial F}{\partial S}\right)^* + \frac{1}{1+q}\left(\frac{\partial F}{\partial I}\right)^* + \frac{q}{1+q}\left(\frac{\partial F}{\partial R}\right)^*\right) - F_0, \tag{5.10}$$

which clearly illustrates the sufficient conditions

$$\frac{\partial F}{\partial S} \geq 0, \quad \frac{\partial F}{\partial I} \leq 0, \quad \frac{\partial F}{\partial R} \leq 0,$$

for uniqueness.

Part III

Local Heterogeneity

Chapter 6

Scale-Free Networks at time t: Degree Distribution & Epidemic Threshold.

The Albert-Barabási scale-free network model has received a great deal of attention for over a decade since preliminary work focused on determining the mechanism that produced power-law degree distributions of certain real-world networks emerged in 1999 [13]. Work rapidly ensued to apply this network structure to phenomena such as personal preferences, language structure, statistical physics (Bose-Einstein condensation), epidemics, and many biological examples [20, 34, 41, 46, 91, 93, 103, 104, 113, 127]. The intrinsic concept of *preferential attachment* (previously known as *cumulative advantage*) was first proposed by Price as an application to citation networks with a slightly more general model [116]. Despite the decades this concept has persisted, an appropriately general, explicit solution to the *degree distribution* of such a network has been absent.

Exact solutions for robust metrics (e.g., degree distribution, mean path length, eigenvalue structure, etc...) are important in the analysis of the structure of the networks and phenomena spreading on them. Dorogovtsev *et al.* explicitly calculated the path length of a small world network structure from the discrete model [53], and Newman *et al.* did the same for the mean field model [107]. Wang *et al.* addressed epidemic spread on complex networks without an explicit treatment of the specific network archetype [133]. Instead, they analyzed the epidemic threshold produced via the eigenvalue structure of complex topologies. Valente proposed a technique for investigating thresholds of spreading on social networks (either established or as they form) based on the type of adoption employed in the diffusion of innovation by individuals [131]. While he does not explicitly address the network topology itself, it is clear that a need for the distribution on the number of connections, as well as their

"quality", at any time during the spread is important to understanding the resultant diffusion.

One method used to model the dynamic evolution of scale-free networks is the doubly indexed, by degree and time, difference equation (partial difference equation) [106]

$$p_{k,t}(t_i) = \frac{k-1}{2(t-1)} p_{k-1,t-1}(t_i) + \left(1 - \frac{k}{2(t-1)}\right) p_{k,t-1}(t_i).$$

In this model new nodes are introduced at unit times with a single edge to connect preferentially to highly connected existing nodes. The solution to this model was found asymptotically, i.e. as $t \to \infty$, by Dorogovtsev *et al.* [54]. The continuous time analog was simultaneously solved by Krapivsky *et al.*, again in the "most interesting asymptotic regime ($t \to \infty$)" [89]. It is unclear if Krapivsky and his co-authors found the explicit time t solution for this paper specifically. Later work by Krapivsky and Redner contains the explicit solution at any time t (for any finite network size) of the degree distribution for a model where rewiring of connections, and a more general preferential connection structure, is modeled [88]. In this, and a previous paper, the authors investigate various correlations (e.g., between age of nodes and their connectivity) in addition to utilizing rate equations (differential equations) and the generating function (partial differential equations) to compute the form and various moments of the degree distribution [87].

Presented here is something of a side-generalization of the scale-free model that was previously solved asymptotically by Krapivsky *et al.*, again explicitly by Krapivsky and Redner and in other approximate forms by Barabási, Albert, and Jeong [14]. The model in this chapter allows new nodes to enter the network at constant rate Λ. The probability that a new node is of degree k is given by ρ_k, and naturally $\langle \rho \rangle$ is the expected value of the degree of a new node. The only requirements put on the new node distribution is that it is a discrete distribution defined on nonnegative integers, has finite expected value[1], and is chosen to not induce multiple connections to the

[1] As will be shown later, for the n^{th} moment of the degree distribution to exist then $\langle \rho^j \rangle < \infty$ for all

same node[2]. The last condition is a technical one arising from the desire to not have the distribution, ρ_k, time or state dependent (i.e., if there are only 2 nodes in the network and a new node enters with degree 3 making all of its connections immediately, then there would be a problem). Thus the third restriction, along with assuming the distribution is constant, implies that ρ_k is only non-zero for k from 1 to $N(0)$. A new node connects to k existing nodes with density dependent preferential attachement rate $\pi_k(t) = \frac{k+w}{D(t)}$, where $D(t)$ is the correct normalizer which will be derived later. The w term (identical to the λ term in Krapivsky and Redner) is an additive shift on linear preferential attachment that allows for scaling the importance of popularity in attachment rate (when $w = 0$ the attachment is identical to the Albert-Barabási model and as $w \to \infty$ the attachment becomes more random [88]). The derivation of the so-called rate equation(s) for degree k nodes is well handled in a number of review papers [4, 26, 105], and is given for this model by

$$\dot{N}_k(t) = \Lambda\rho_k + \Lambda\langle\rho\rangle\left[\pi_{k-1}(t)N_{k-1}(t)(1-\delta_{k,0}) - \pi_k(t)N_k(t)\right], \tag{6.1}$$

for $k \in \mathbb{N}$. The model considered here is dynamically less-rich (the process of rewiring is excluded), but considers a more general "introduction regime" (the newly introduced nodes are sampled from a general distribution) compared to the model studied by Krapivsky and Redner.

While scale-free networks match the form of many established structures, such as the World Wide Web ([15]), they do not properly match the evolution of such structures [2]. A goal here is to produce explicit, time t solutions which may be used in addressing questions which take place during the evolution of such network (e.g., synchronicity or "worst-case" spread of disease on networks as considered in [108] and [41] respectively). The ability to compare these time t solutions with specific application data may allow for researchers to more accurately propose better models

$j \leq n$.

[2]Loops and paths between two nodes of length two are not allowed by the algorithm.

for the phenomena in question.

Analysis of the mean field approximation of a simple epidemic model, where individuals are either susceptible, infectious or immune to the disease, on a network (heterogeneous) topology involves computation of the moments of the *degree distribution*. Such a model has a so-called effective reproductive number, $\mathscr{R}$, which, when less than one for all time, implies that an epidemic outbreak will not occur within the population and the infectious population will exponentially decay to 0 [71]. If β is the rate of infection given a contact between a susceptible and an infectious individual, $\frac{1}{\gamma}$ is the average period of infectiousness, and the population exhibits heterogeneous connectivity, a non-regular graph/network structure, this quantity has the form

$$\mathscr{R} = \frac{\beta}{\gamma}\langle K\rangle \frac{\langle K^2\rangle_S}{\langle K\rangle^2} \tag{6.2}$$

where $\langle K^n\rangle$ is the n^{th} moment of the degree distribution and $\langle K^n\rangle_S$ is the same but conditioned on the individuals being susceptible. The quantity $\frac{\beta}{\gamma}\langle K\rangle$ is the traditional $\mathscr{R}_0$, *basic reproductive number*, calculation for homogeneous topologies but the new term

$$\frac{\langle K^2\rangle_S}{\langle K\rangle^2}$$

involves the structure of the network, implicitly considered at time t (at the beginning of the epidemic this is approximately $\frac{\langle K^2\rangle}{\langle K\rangle^2}$ in the absence of previously implemented controls like vaccination, quarantine, etc...).

When considering computation on complex networks the scale-free model of Albert and Barabasi is an attractive option because it has a solidly derived form (see [4] and references therein) and may serve as a "simple" starting point in the analysis of more complex structures. Also, it has been shown that $\mathscr{R}$ is infinite for A-B scale-free networks that have achieved their limiting degree distribution, due to $\frac{\langle K^2 \rangle}{\langle K \rangle} \to \infty$. I wish to investigate if the structure of the network as it evolves affords a second moment small enough so that $\mathscr{R} < \infty$, and possibly less than one. This would indicate that on newly forming, or otherwise small, scale-free networks an epidemic peak may not occur so long as it is introduced before some critical time.

In Section 6.1 I go over the general recursive form of the solution found via integrating factors. Solving the generalized system presented here for base cases ($k = 0, 1,$ and 2) is straightforward and omitted for brevity. Induction is used to demonstrate that the observed pattern is consistent across all k and is in Section 6.1. The moments are directly calculated from a set of ODEs where the state variables are the moments themselves. This procedure is described in Section 6.2. In Section 6.3 I show that the solution to the more classical Albert-Barabási model may be found as well as a special case of Krapivsky and Redner's. I conclude the paper with a discussion of the results and future work.

6.1 General Solution Form

For the sake of space I define

$$R_k(t) := \pi_k(t)\Lambda\langle\rho\rangle = \frac{(k+w)\Lambda\langle\rho\rangle}{D(t)}.$$

One may derive the expression for $D(t)$ both intuitively and somewhat more rigorously. Intuitively, it is true that $D(t) = \sum_j jN_j(t) + w\sum_j N_j(t)$, or in words $D(t)$ is the sum of the "double-count" of all edges in the network with w times the total

number of nodes. Since a new node enters the system at rate Λ this gives

$$w\sum_j N_j = w(\Lambda t + N_0),$$

with an initial number of nodes $N_0 = N(0)$. This may be more rigorously shown with a straightforward application of uniform convergence to the partial sums to conclude the final summation by solving the ODE for $N(t)$, $\dot{N}(t) = \Lambda$. A proportion of these new nodes, ρ_k will enter with degree k and thus add $2k$ to the double-count of edges. Thus, with an initial double count of E_0, one may write

$$\sum_j jN_j(t) = 2\Lambda\langle\rho\rangle t + E_0,$$

again we consider the expression $\frac{d\sum_j jN_j(t)}{dt}$ and reduce the telescoping sum assuming $m\pi_m N_m(t) \to 0$ as $m \to \infty$. This gives $D(t) = (2\Lambda\langle\rho\rangle + w\Lambda)t + (E_0 + wN_0)$.

Further let

$$e^{\int R_k(t)dt} = D(t)^{\frac{\langle\rho\rangle(k+w)}{2\langle\rho\rangle+w}} =: H_{\langle\rho\rangle(k+w)}(t).$$

The general solution to $N_k(t)$ can be given recursively as

$$\begin{aligned} \dot{N}_k(t) + R_k(t)N_k(t) &= \Lambda\rho_k + R_{k-1}(t)N_{k-1}(t)(1-\delta_{k,0}), \\ \frac{d\left[N_k(t)H_{\langle\rho\rangle(k+w)}(t)\right]}{dt} &= \left(\Lambda\rho_k + R_{k-1}(t)N_{k-1}(t)(1-\delta_{k,0})\right)H_{\langle\rho\rangle(k+w)}(t), \end{aligned} \tag{6.3}$$

and finally

$$\begin{aligned} N_k(t) &= \frac{\int\left[\Lambda H_{\langle\rho\rangle(k+w)}(t)\rho_k + R_{k-1}(t)N_{k-1}(t)H_{\langle\rho\rangle(k+w)}(t)(1-\delta_{k,0})\right]dt + C_k}{H_{\langle\rho\rangle(k+w)}(t)}, \\ &= \frac{(1-\delta_{k,0})\Lambda\langle\rho\rangle(k-1+w)}{H_{\langle\rho\rangle(k+w)}(t)}\int N_{k-1}(t)H_{\langle\rho\rangle(k+w-2)-w}(t)dt \\ &+ \frac{\rho_k H_{2\langle\rho\rangle+w}(t)}{\langle\rho\rangle(k+w+2)+w} + \frac{C_k}{H_{\langle\rho\rangle(k+w)}(t)}, \end{aligned} \tag{6.4}$$

where the final step follows from the two observations (when $k+w \neq -2$)

$$\begin{aligned} \int H_\alpha(t)dt &= \int\left((2\Lambda\langle\rho\rangle + \Lambda w)t + (E_0 + wN_0)\right)^{\frac{\alpha}{2\langle\rho\rangle+w}}dt, \\ &= \frac{\left((2\Lambda\langle\rho\rangle + \Lambda w)t + (E_0 + wN_0)\right)^{\frac{\alpha+2\langle\rho\rangle+w}{2\langle\rho\rangle+w}}}{\Lambda(\alpha + 2\langle\rho\rangle + w)} = \frac{H_{\alpha+2\langle\rho\rangle+w}(t)}{\Lambda(\alpha+2\langle\rho\rangle+w)}, \end{aligned} \tag{6.5}$$

and

$$\frac{H_{\alpha+n}(t)}{H_{\alpha}(t)} \quad = \quad H_n(t). \tag{6.6}$$

The constants of integration, C_k, are defined recursively via

$$\begin{aligned} C_k \quad = \quad & -(1-\delta_{k,0})\Lambda\langle\rho\rangle(k-1+w)\left(\int N_{k-1}(t)H_{\langle\rho\rangle(k+w-2)-w}(t)dt\right|_{t=0} \\ & +N_k(0)H_{\langle\rho\rangle(k+w)}(0) - \frac{\rho_k H_{\langle\rho\rangle(k+w+2)+w}(0)}{\langle\rho\rangle(k+w+2)+w}. \end{aligned} \tag{6.7}$$

The expression for $N_k(t)$ has the general form (proven consistent for all k by induction)

$$\begin{aligned} N_k(t) = \sum_{j=0}^{k}\left(\frac{\rho_j}{\langle\rho\rangle}\frac{\Gamma\left(\frac{w}{\langle\rho\rangle}+w+j+2\right)\Gamma(w+k)}{\Gamma\left(\frac{w}{\langle\rho\rangle}+w+k+3\right)\Gamma(w+j)}H_{2\langle\rho\rangle+w}(t)\right. \\ \left.+\frac{C_j}{H_{\langle\rho\rangle(w+j)}(t)}\frac{\Gamma(w+k)}{\Gamma(w+j)(k-j)!}\right), \end{aligned} \tag{6.8}$$

with the constants of integration

$$C_j = H_{\langle\rho\rangle(w+j)}(0)\sum_{i=0}^{j}\frac{(-1)^{j-i}\Gamma(w+j)}{(j-i)!\Gamma(w+i)}\left[N_i(0)-\frac{\rho_i}{\langle\rho\rangle}\frac{H_{2\langle\rho\rangle+w}(0)}{\Gamma\left(\frac{w}{\langle\rho\rangle}+w+j+2\right)}\right]. \tag{6.9}$$

Inductive Step

To prove the general solution is consistent for all k I invoke induction, assuming that the solution is verified for cases such as $k = 0,1,2$. Suppose that for some $k > 2$ the solution given by Equations (6.8) and (6.9) hold. Utilizing Equation (6.4) we may construct the solution of $N_{k+1}(t)$. Addressing just the integral term we have

$$\begin{aligned} \sum_{j=0}^{k}\left[\int\left(\frac{\rho_j\Gamma\left(\frac{w}{\langle\rho\rangle}+w+j+2\right)\Gamma(w+k)}{\langle\rho\rangle\Gamma\left(\frac{w}{\langle\rho\rangle}+w+k+3\right)\Gamma(w+j)}H_{\langle\rho\rangle(k+w+1)}(t)\right.\right. \\ \left.\left.+\frac{C_j\Gamma(w+k)}{\Gamma(w+j)(k-j)!}H_{\langle\rho\rangle(k-j-1)-w}(t)\right)dt\right], \end{aligned} \tag{6.10}$$

which reduces, using Equation (6.5), to

$$\sum_{j=0}^{k}\left[\frac{\rho_j\Gamma\left(\frac{w}{\langle\rho\rangle}+w+j+2\right)\Gamma(w+k)}{\langle\rho\rangle\Gamma\left(\frac{w}{\langle\rho\rangle}+w+k+3\right)\Gamma(w+j)}\frac{H_{\langle\rho\rangle(k+w+3)+w}(t)}{\Lambda(\langle\rho\rangle(k+w+3)+w)}\right.$$
$$\left.+\frac{C_j\Gamma(w+k)}{\Gamma(w+j)(k-j)!}\frac{H_{\langle\rho\rangle(k-j)}(t)}{\Lambda\langle\rho\rangle(k-j+1)}\right].$$

Plugging this value in for the integral gives a solution of the form

$$N_{k+1}(t) = \frac{\rho_{k+1}H_{2\langle\rho\rangle+w}(t)}{\langle\rho\rangle\left(\frac{w}{\langle\rho\rangle}+w+k+3\right)}+\sum_{j=0}^{k}\left[\frac{\rho_j\Gamma\left(\frac{w}{\langle\rho\rangle}+w+j+2\right)\Gamma(w+k+1)}{\langle\rho\rangle\Gamma\left(\frac{w}{\langle\rho\rangle}+w+k+4\right)\Gamma(w+j)}H_{2\langle\rho\rangle+w}(t)\right.$$
$$\left.+\frac{C_j\Gamma(w+k+1)}{\Gamma(w+j)(k-j+1)!H_{\langle\rho\rangle(w+j+1)}(t)}\right]+\frac{C_{k+1}}{H_{\langle\rho\rangle(k+1+w)}(t)},$$

where each term outside of the summation follow the form of the summand when $j=k+1$. Thus for all k Equation (6.8) is indeed the solution for $N_k(t)$. The validity of Equation (6.9) follows in the exact same manner.

6.2 Moment Calculations From the Density Rate Equations

Working directly with the solutions of $N_k(t)$ to find the moments is undesirable. The size and generality of the expression proves a direct method to be a daunting task. The method employed here is relatively straightforward, but is not simply applying the definition

$$\langle K^n(t)\rangle = \sum_k k^n P_k(t).$$

I derive and solve the ODEs

$$\frac{d\langle K^n(t)\rangle}{dt} = \sum_k k^n \frac{dP_k(t)}{dt},$$

for $k=1,2$. The formalism of swapping summation and differentiation (i.e., uniform convergence of the resulting right hand side and the actual moment existing) is omitted and when convergence is a problem special cases are considered (e.g., for moments higher than one). Solving the ODEs for the first and second moments

involve a rearranging of terms and a simple application of integrating factors. To construct these differential equations first note that

$$\begin{aligned}\dot{P_k}(t) &= \frac{\dot{N_k}(t)}{N(t)} - P_k(t)\ln(\dot{N}(t)),\\ &= \frac{\Lambda}{N(t)}(\rho_k - P_k(t)) + \Lambda\langle\rho\rangle\left(\pi_{k-1}(t)P_{k-1}(t)(1-\delta_{k,0}) - \pi_k(t)P_k(t)\right).\end{aligned}$$

By a simple process of elimination, this results in

$$\frac{d\langle K(t)\rangle}{dt} = \left(\frac{\Lambda}{N(t)} + \frac{\Lambda w}{D(t)}\right)\langle\rho\rangle + \left(\frac{\Lambda\langle\rho\rangle}{D(t)} - \frac{\Lambda}{N(t)}\right)\langle K(t)\rangle.$$

The solution $\langle K(t)\rangle$ is given, using the integrating factor $N(t)H_{-\langle\rho\rangle}(t)$, by

$$\langle K(t)\rangle = \frac{D(t) - wN(t) + H_{\langle\rho\rangle}(t)C_1}{N(t)},$$

where $C_1 = \frac{N(0)}{H_{\langle\rho\rangle}(0)}\left(\langle K(0)\rangle + w - \frac{D(0)}{N(0)}\right)$.

The equation for the second moment is

$$\frac{d\langle K^2(t)\rangle}{dt} = \langle K^2(t)\rangle\left[\frac{2\Lambda\langle\rho\rangle}{D(t)} - \frac{\Lambda}{N(t)}\right] + \langle K(t)\rangle\frac{(2w+1)\Lambda\langle\rho\rangle}{D(t)} + \frac{w\Lambda\langle\rho\rangle}{D(t)} + \frac{\Lambda\langle\rho^2\rangle}{N(t)}.$$

Through the integrating factor $N(t)H_{-2\langle\rho\rangle}(t)$, the solution is

$$\langle K^2(t)\rangle = \frac{D(t)(2\langle\rho^2\rangle + 2(2w+1)\langle\rho\rangle - \Lambda w(2w+1) + w) + 2w^3N(t) + 2w(H_{2\langle\rho\rangle}(t)C_2 - H_{\langle\rho\rangle}(t)C_1(2w+1)}{2wN(t)},$$

with $C_2 = \frac{N(0)}{H_{2\langle\rho\rangle}(0)}\left(-\langle K^2(0)\rangle + \frac{2w+1}{2w}\langle K(0)\rangle + w^2 + \frac{(2w+1)w}{2w} - \frac{(2w+1)(2\langle\rho\rangle - \Lambda w + 1) + 2\langle\rho^2\rangle + w}{2w}\frac{D(0)}{N(0)}\right)$.

6.3 Specific Cases

The solutions above all have very complex forms mostly due to the undefined nature of the new-node entry distribution ρ_k. In the classic A-B scale free model it is assumed that the initial conditions are described by a complete graph of $m+1$ nodes (each with degree m). Dynamically, new nodes enter at rate 1 with m edges, and each of these edges are attached with degree-preferential attachment. This implies that $N(0) = N_m(0) = m+1$, $\Lambda = 1$, $\rho_k = \delta_{k,m}$, $\langle\rho\rangle = m$, and $w = 0$. The expressions for $N_k(t)$ and C_k reduce to

$$N_k(t) = \frac{(m+1)\Gamma(k)}{\Gamma(3+k)}H_{2m}(t) + \sum_{j=m}^{k}\binom{k-1}{j-1}\frac{C_j}{H_{mj}(t)}, \tag{6.11}$$

$$C_j = (-1)^{j-m} H_{mj}(0) \binom{j-1}{m-1} \left[(m+1) - \frac{H_{2m}(0)}{m(j+2)} \right], \tag{6.12}$$

with

$$H_\alpha(t) = (2mt + E_0)^{\frac{\alpha}{2m}}. \tag{6.13}$$

Evoking that $E_0 = 2m$ the solution collapses to

$$\begin{aligned} N_k(t) = \frac{(m+1)\Gamma(k)}{\Gamma(k+3)} H_{2m}(t) \quad &+ \quad \binom{k-1}{m-1} \left[\frac{(2m)^{\frac{m}{2}}(m+1)((1+t)^{\frac{1}{2}}-1)^{k-m}}{(1+t)^{\frac{k}{2}}} \right. \\ &- \quad \left. 2\frac{{}_2F_1\left(m+2, m-k; m+3; \frac{1}{(1+t)^{\frac{1}{2}}}\right)}{(m+2)(1+t)^{\frac{m}{2}}} \right]. \end{aligned} \tag{6.14}$$

The moments of the distribution also collapse somewhat:

$$\begin{aligned} \langle K(t) \rangle &= \frac{2m(t+1) + m(m-1)(t+1)^{\frac{1}{2}}}{m+t+1}, \\ \langle K^2(t) \rangle &= \frac{(m^3 + 2m^2 + m + m(m+1)\ln(t+1))(t+1)}{m+t+1} - \langle K(t) \rangle, \\ \Psi(t) := \frac{\langle K^2(t) \rangle}{\langle K(t) \rangle} &= \frac{(m^3 + 2m^2 + m + m(m+1)\ln(t+1))(t+1)^{\frac{1}{2}}}{2m(t+1)^{\frac{1}{2}} + m(m-1)} - 1. \end{aligned} \tag{6.15}$$

Note that it is straightforward to show that $\frac{d\Psi(t)}{dt} > 0$ (for the most general case here). Thus, the minimum is at time $t = 0$, $\frac{\langle K^2(0) \rangle}{\langle K(0) \rangle}$. For $\mathscr{R} < 1$ we require that $\frac{\beta}{\gamma} < \frac{\langle K(0) \rangle}{\langle K^2(0) \rangle}$. In the case of the A-B model this means $\frac{\beta}{\gamma} < \frac{1}{m} \leq 1$. In the case of a random network (i.e., $w \to \infty$) the parameters would have to be outside of biological feasibility (i.e., one would have to be negative).

6.4 Conclusion/ Discussion

The value of $N_k(t)$ is dependent on initial conditions, $N_j(0)$ from $j = 0, 1, ..., k$. The dependence of $N_k(t)$ on $N_j(0)$ decays on the order of $t^{-\frac{\langle \rho \rangle (w+j)}{2\langle \rho \rangle + w}}$ (e.g., for A-B attachment $t^{-\frac{j}{2}}$ and for random attachment $t^{-\langle \rho \rangle}$). Thus the dependence of $N_k(t)$ on the system's initial conditions, as a whole, decay on the order $t^{-\frac{\langle \rho \rangle (w+k)}{2\langle \rho \rangle + w}}$. This interestingly implies that the larger the degree considered the less the initial conditions impact the count of nodes of that degree.

We may consider $\frac{N_k(t)}{N(t)}$, the proportion of nodes that have degree k, and find that this both asymptotically agrees with previous findings of the degree distribution which do not consider transitory behavior, and with those found by Krapivsky and Redner under appropriate restrictions. Specifically for the general model presented here

$$P_k = \lim_{t\to\infty} \frac{N_k(t)}{\Lambda t + N(0)} = \sum_{j=0}^{k} \frac{\rho_j}{\langle\rho\rangle} \frac{\Gamma\left(\frac{w}{\langle\rho\rangle} + w + j + 2\right)\Gamma(w+k)}{\Gamma\left(\frac{w}{\langle\rho\rangle} + w + k + 3\right)\Gamma(w+j)}(2\langle\rho\rangle + w). \qquad (6.16)$$

The driving force to this investigation was to answer the question "Does the evolving structure of an A-B scale-free network allow for a second moment small enough such that $\mathscr{R} < \infty$, and possibly less than one?" The method that has definitively answered my question with a resounding "probably not," is very straightforward. By considering ODEs in the form

$$\frac{d\mathbb{E}([K(t)]^n)}{dt} = \sum_k k^n \frac{dP_k(t)}{dt},$$

I was able to construct the desired quantity $\Psi(t)$, the ratio of the second moment to the first. Noting that often $\Psi(0) > 1$ and $\dot{\Psi}(t) > 0$ we may conclude that there is little hope of having a control reproductive number less than one without most of the susceptible nodes being depleted (at the end of an epidemic).

Direct application of this time t solution is useful for two reasons: formation of, and dynamics on complex networks. While the exact rules that particular real-world networks follow during their formation is an intractable key to their structure, one may compare such structures (data) with the solution given here to justify if this node introduction/attachment mechanism is applicable. For dynamics on non-limiting case networks (e.g., diseases through a smaller population, spread of computer viruses on local networks, ideas spreading through *forming* social cliques, etc...) this solution may also be employed to describe the topology of the relevant environment. Results such as those found by Zhao *et. al*, on the fragility of scale-free networks to attacks on hubs, should be extended to include networks that are evolving over time [137]. It is

possible, that since at time t there is a smaller probability of hubs than in the limiting network (i.e., the tail is not fat enough yet), the network is more resilient to cascading failures than the seemingly more robust limiting network.

There are a few flaws with considering these rate equations as a model for the A-B network growth. In an arbitrarily small time interval from $t = 0$ there has been mass shifted into compartments, $N_k(t)$, with arbitrarily large degree. Indeed, the process description only allows for $\lfloor 2\Lambda mt \rfloor + E_0$ edges to exist at any time t, and thus there should be a maximum degree at any time. Pastor-Satorras and Vespignani do this to some extent for general scale-free networks with the introduction of exponentially decaying tails induced on the degree distribution [114]. This was done in the interest of investigating the epidemic threshold for finite sized networks to find structures with noncritical spread despite the scale-free property. This method however would be inadequate for an investigation into the structure of the networks as they evolve since in their work the networks were taken to be static. Instead the "cut-off" may be modeled implicitly with some sort of modified system that permits the correct number of ODEs to be acting at any time t. A delay type differential equation system containing delays similar to $t - \lfloor \frac{k-N_0}{\Lambda} \rfloor$ and initial conditions like $N_k(t) = 0$ for $t \in \left[-\lfloor \frac{k-N_0}{\Lambda} \rfloor, 0 \right)$ would perhaps do a better job at modeling the transitory dynamics of this system as they possess this sort of activation switch (i.e., they only start to accept "mass" once the delay exits the initial data). This would force a finiteness that is otherwise lost in the traditional rate equation techniques.

For any time t the continuum approximation admits that Λt new nodes have entered the network. Supposing a particular initial node η is selected once by each new node, then the degree of η at time t is at most $\lceil m + \Lambda t \rceil$. Thus at times where $\Lambda t \in \mathbb{Z}^+$ the number of ODEs required would increase by one and the system size itself would be

growing. Indeed, one may define

$$Q_k(t) = \frac{P_k(t)}{\sum_{k=m}^{\lceil m+\Lambda t \rceil} P_k(t)},$$

in order to be defining actual densities, in fact conditional probabilities with respect to the original $P_k(t)$ terms. The ODEs would then take on a peculiar form of

$$\dot{Q}_k(t) = \frac{\dot{P}_k(t)}{\left(\sum_{\ell=m}^{\lceil m+\Lambda t \rceil} P_\ell(t)\right)^2} - Q_k(t) \sum_{\ell=m}^{\lceil m+\Lambda t \rceil} Q_\ell(t).$$

Further work is also open in the gross generalization of this process. The concept of rewiring may be reintroduced to the model in addition to a suite of more general attachment kernels. Under these generalizations it is very likely that the network may not retain an asymptotic distribution of $P(k) \propto k^{-\gamma}$, especially if the attachment kernel differs from proportional. However, the true power of the A-B model should be seen in its flexibility to allow for generalizations that produces richer behavior and thus, farther reaching applications.

6.5 Acknowledgements

The author would like to thank Sidney Redner for his suggestions that broadened the discussion and intellectual merit of the paper, Laurent Hébert-Dufresne for his careful reading, and Carlos Castillo-Chavez along with Sergei Suslov for their encouragement and direction.

REFERENCES

[1] *Centers for disease control, cdc fact sheet: Gonorrhea*, (2007).

[2] L.A. Adamic and B.A. Huberman, *Power-law distribution of the world wide web*, Science **287** (2000), no. 5461, 2115.

[3] J. Adda and R.W. Cooper, *Dynamic economics: quantitative methods and applications*, The MIT Press, 2003.

[4] R. Albert and A.L. Barabási, *Statistical mechanics of complex networks*, Reviews of modern physics **74** (2002), no. 1, 47–97.

[5] RM Anderson, *Population dynamics of infectious diseases: theory and applications*, London, New York (1982).

[6] R.M. Anderson, *Population dynamics of infectious diseases: theory and applications*, Chapman and Hall, 1982.

[7] RM Anderson, SP Blythe, S. Gupta, and E. Konings, *The transmission dynamics of the human immunodeficiency virus type 1 in the male homosexual community in the united kingdom: the influence of changes in sexual behaviour*, Philosophical Transactions of the Royal Society of London. B, Biological Sciences **325** (1989), no. 1226, 45–98.

[8] R.M. Anderson and R.M. May, *Infectious diseases of humans: dynamics and control*, vol. 26, Wiley Online Library, 1992.

[9] M.C. Auld, *Choices, beliefs, and infectious disease dynamics*, Journal of Health Economics **22** (2003), no. 3, 361–377.

[10] N.T.J. Bailey, *The mathematical theory of infectious diseases and its applications*, The mathematical theory of infectious diseases and its applications 2nd edition **413** (1975).

[11] Frank Ball and Peter Neal, *A general model for stochastic SIR epidemics with two levels of mixing.*, Mathematical biosciences **180** (2002), 73–102.

[12] H.T. Banks and C. Castillo-Chávez, *Bioterrorism: mathematical modeling applications in homeland security*, Society for Industrial Mathematics, 2003.

[13] A.L. Barabási and R. Albert, *Emergence of scaling in random networks*, Science **286** (1999), no. 5439, 509.

[14] A.L. Barabási, R. Albert, and H. Jeong, *Mean-field theory for scale-free random networks*, Physica A: Statistical Mechanics and its Applications **272** (1999), no. 1-2, 173–187.

[15] _____, *Scale-free characteristics of random networks: the topology of the world-wide web*, Physica A: Statistical Mechanics and its Applications **281** (2000), no. 1-4, 69–77.

[16] R. Barrett and P.J. Brown, *Stigma in the time of influenza: social and institutional responses to pandemic emergencies*, Journal of Infectious Diseases **197** (2008), no. Supplement 1, S34.

[17] W.H. Batchelder and A.K. Romney, *Test theory without an answer key*, Psychometrika **53** (1988), no. 1, 71–92.

[18] _____, *New results in test theory without an answer key*, Mathematical psychology in progress (1989), 229–248.

[19] L. Bettencourt, R.M. Ribeiro, G. Chowell, T. Lant, and C. Castillo-Chavez, *Towards real time epidemiology: data assimilation, modeling and anomaly detection of health surveillance data streams*, Proceedings of the 2nd NSF conference on Intelligence and security informatics: BioSurveillance, Springer-Verlag, 2007, pp. 79–90.

[20] G. Bianconi and A.L. Barabási, *Bose-Einstein condensation in complex networks*, Physical Review Letters **86** (2001), no. 24, 5632–5635.

[21] SP Blythe, F. Brauer, and C. Castillo-Chavez, *Demographic recruitment in sexually transmitted disease models*, Biometrics Unit Technical Report BU-1154-M, Cornell University (1992).

[22] S.P. Blythe and C. Castillo-Chavez, *Like-with-like preference and sexual mixing models*, Mathematical biosciences **96** (1989), no. 2, 221–238.

[23] S.P. Blythe, C. Castillo-Chavez, and G. Casella, *Empirical methods for the estimation of the mixing probabilities for socially structured populations from a single survey sample*, Mathematical Population Studies **3** (1992), no. 3, 199–225.

[24] SP Blythe, C. Castillo-Chavez, and M. Palmer, *Toward a unified theory of sexual mixing and pair formation*, Mathematical Biosciences **107** (1991), no. 2, 379–405.

[25] SP Blythe, KL Cooke, and C. Castillo-Chavez, *Autonomous risk-behavior change, and non-linear incidence rate, in models of sexually transmitted diseases*, Biometrics Unit Technical Report B-1048-M (1992).

[26] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.U. Hwang, *Complex networks: Structure and dynamics*, Physics Reports **424** (2006), no. 4-5, 175–308.

[27] S.P. Borgatti, M.G. Everett, and L.C. Freeman, *Ucinet for windows: Software for social network analysis*, Harvard Analytic Technologies **2006** (2002).

[28] I. Bozicevic, K.A. Fenton, I. Martin, E.A. Rudd, C.A. Ison, K. Nanchahal, and K. Wellings, *Epidemiological correlates of asymptomatic gonorrhea*, Sexually transmitted diseases **33** (2006), no. 5, 289.

[29] F. Brauer and C. Castillo-Chavez, *Mathematical models in population biology and epidemiology*, vol. 40, Springer Verlag, 2001.

[30] F. Brauer, C. Castillo-Chavez, and J.X. Velasco-Hernandez, *Recruitment into a core group and its effect on the spread of a sexually transmitted disease*, Biometrics Unit Technical Report BU-1320-M, Cornell University, Ithaca, NY (1996).

[31] S. Busenberg and C. Castillo-Chavez, *Interaction, pair formation and force of infection terms in sexually transmitted diseases*, Mathematical and statistical approaches to AIDS epidemiology, Springer-Verlag New York, Inc., 1990, pp. 289–300.

[32] S. Busenberg and C. Castillo-Chávez, *A general solution of the problem of mixing of subpopulations and its application to risk-and age-structured epidemic models for the spread of AIDS*, Mathematical Medicine and Biology **8** (1991), no. 1, 1.

[33] S. Busenberg and P. Vandendriessche, *A method for proving the non-existence of limit cycles*, Journal of mathematical analysis and applications **172** (1993), no. 2, 463–479.

[34] G. Caldarelli, *Scale-Free Networks: Complex webs in nature and technology*, Oxford University Press, USA, 2007.

[35] C. Castillo-Chávez, *Mathematical and statistical approaches to aids epidemiology*, Springer-Verlag New York, Inc., 1990.

[36] C. Castillo-Chavez, S. Fridman, and X. Luo, *Stochastic and deterministic models in epidemiology*, World Congress of Nonlinear Analysts' 92: proceedings of the First World Congress of Nonlinear Analysts, Tampa, Florida, August 19-26, 1992, vol. 4, Walter de Gruyter, 1996, p. 3211.

[37] C. Castillo-Chavez and W. Huang, *Age-structured core group model and its impact on std dynamics*, IMA VOLUMES IN MATHEMATICS AND ITS APPLICATIONS **126** (2002), 261–274.

[38] C. Castillo-Chevez and H.R. Thieme, *Asymptotically autonomous epidemic models*, vol. 94, Mathematical Sciences Institute, Cornell University, 1994.

[39] F.H. Chen, *Modeling the effect of information quality on risk behavior change and the transmission of infectious diseases*, Mathematical biosciences **217** (2009), no. 2, 125–133.

[40] G. Chowell, S.M. Bertozzi, M.A. Colchero, H. Lopez-Gatell, C. Alpuche-Aranda, M. Hernandez, and M.A. Miller, *Severe respiratory disease concurrent with the circulation of h1n1 influenza*, New England Journal of Medicine **361** (2009), no. 7, 674–679.

[41] G. Chowell and C. Castillo-Chavez, *Worst-Case Scenarios and Epidemics*, Bioterrorism: mathematical modeling applications in homeland security (2003), 35.

[42] G. Chowell, C. Castillo-Chavez, P.W. Fenimore, C.M. Kribs-Zaleta, L. Arriola, and J.M. Hyman, *Model parameters and outbreak control for sars.*, Emerging Infectious Diseases **10** (2004), no. 7.

[43] G. Chowell, P.W. Fenimore, M.A. Castillo-Garsow, and C. Castillo-Chavez, *Sars outbreaks in ontario, hong kong and singapore: the role of diagnosis and isolation as a control mechanism*, Journal of Theoretical Biology **224** (2003), no. 1, 1–8.

[44] G. Chowell, N.W. Hengartner, C. Castillo-Chavez, P.W. Fenimore, and JM Hyman, *The basic reproductive number of ebola and the effects of public health measures: the cases of congo and uganda*, Journal of Theoretical Biology **229** (2004), no. 1, 119–126.

[45] G. Chowell, J.M. Hyman, and L.M.A. Bettencourt, *Mathematical and statistical estimation approaches in epidemiology*, Springer Verlag, 2009.

[46] N.A. Christakis and J.H. Fowler, *The spread of obesity in a large social network over 32 years*, The New England Journal of Medicine **357** (2007), no. 4, 370.

[47] J.A. Córdova-Villalobos, E. Sarti, J. Arzoz-Padrés, G. Manuell-Lee, J.R. Méndez, P. Kuri-Morales, et al., *The influenza a (h1n1) epidemic in mexico. lessons learned*, Health Res Policy Syst **7** (2009), no. 1, 21.

[48] S. Del Valle, H. Hethcote, JM Hyman, and C. Castillo-Chavez, *Effects of behavioral changes in a smallpox attack model*, Mathematical Biosciences **195** (2005), no. 2, 228–251.

[49] K. Dietz, *On the transmission dynamics of hiv*, Mathematical Biosciences **90** (1988), no. 1-2, 397–414.

[50] K. Dietz and KP Hadeler, *Epidemiological models for sexually transmitted diseases*, Journal of Mathematical Biology **26** (1988), no. 1, 1–25.

[51] _____, *Epidemiological models for sexually transmitted diseases*, Journal of Mathematical Biology **26** (1988), no. 1, 1–25.

[52] A.K. Dixit and R.S. Pindyck, *Investment under uncertainty*, {Princeton University Press}, 1994.

[53] S.N. Dorogovtsev and J.F.F. Mendes, *Exactly solvable small-world network*, EPL (Europhysics Letters) **50** (2000), 1.

[54] S.N. Dorogovtsev, J.F.F. Mendes, and A.N. Samukhin, *Structure of growing networks with preferential linking*, Physical Review Letters **85** (2000), no. 21, 4633–4636.

[55] J. Dushoff, W. Huang, and C. Castillo-Chavez, *Backwards bifurcations and catastrophe in simple models of fatal diseases*, Journal of mathematical biology **36** (1998), no. 3, 227–248.

[56] J.M. Epstein, D.M. Goedecke, F. Yu, R.J. Morris, D.K. Wagener, and G.V. Bobashev, *Controlling pandemic flu: the value of international air travel restrictions*, PLoS One **2** (2007), no. 5, e401.

[57] Z. Feng, C. Castillo-Chavez, and A.F. Capurro, *A model for tuberculosis with exogenous reinfection*, Theoretical Population Biology **57** (2000), no. 3, 235–247.

[58] E.P. Fenichel, C. Castillo-Chavez, MG Ceddia, G. Chowell, P.A.G. Parra, G.J. Hickling, G. Holloway, R. Horan, B. Morin, C. Perrings, et al., *Adaptive human behavior in epidemiological models*, Proceedings of the National Academy of Sciences **108** (2011), no. 15, 6306.

[59] E.P. Fenichel and R.D. Horan, *Jointly-determined ecological thresholds and economic trade-offs in wildlife disease management*, Natural Resource Modeling **20** (2007), no. 4, 511–547.

[60] J.D. Fortenberry, M. McFarlane, A. Bleakley, S. Bull, M. Fishbein, D.M. Grimley, C.K. Malotte, and B.P. Stoner, *Relationships of stigma and shame to gonorrhea and hiv screening*, American Journal of Public Health **92** (2002), no. 3, 378.

[61] A.P. Galvani, T.C. Reluga, and G.B. Chapman, *Long-standing influenza vaccination policy is in accord with individual self-interest but not with the utilitarian optimum*, Proceedings of the National Academy of Sciences **104** (2007), no. 13, 5692.

[62] P.Y. Geoffard and T. Philipson, *The empirical content of canonical models of infectious diseases: The proportional hazard specification*, Biometrika **82** (1995), no. 1, 101–111.

[63] _____, *Rational epidemics and their public control*, International Economic Review (1996), 603–624.

[64] A.B. Gumel, C. Castillo-Chávez, R.E. Mickens, and D.P. Clemence, *Mathematical studies on human disease dynamics: emerging paradigms and challenges: Ams-ims-siam joint summer research conference on modeling the dynamics of human diseases: emerging paradigms and challenges, july 17-21, 2005, snowbird, utah*, vol. 410, Amer Mathematical Society, 2006.

[65] S. Gupta, R.M. Anderson, R.M. May, et al., *Networks of sexual contacts: implications for the pattern of spread of hiv.*, AIDS (London, England) **3** (1989), no. 12, 807.

[66] KP Hadeler, *Modeling aids in structured populations*, Bull. Int. Stat. Inst **53** (1989), 83–99.

[67] K.P. Hadeler and C. Castillo-Chavez, *A core group model for disease transmission*, Mathematical biosciences **128** (1995), no. 1-2, 41–55.

[68] K.P. Hadeler and L. Elsner, *The spectral radius and the largest basic reproduction number in multi-type epidemic models with mixing*, Unpublished Manuscript (2009).

[69] R.J. Hatchett, C.E. Mecher, and M. Lipsitch, *Public health interventions and epidemic intensity during the 1918 influenza pandemic*, Proceedings of the National Academy of Sciences **104** (2007), no. 18, 7582.

[70] JAP Heesterbeek and MG Roberts, *The type-reproduction number t in models for infectious disease control*, Mathematical biosciences **206** (2007), no. 1, 3–10.

[71] H.W. Hethcote, *The mathematics of infectious diseases*, SIAM review **42** (2000), no. 4, 599–653.

[72] H.W. Hethcote and J.W. Van Ark, *Modeling hiv transmission and aids in the united states*, Springer-Verlag, 1992.

[73] H.W. Hethcote and J.A. Yorke, *Gonorrhea transmission dynamics and control*, Springer-Verlag New York, 1984.

[74] R.D. Horan, E.P. Fenichel, K.L.S. Drury, and D.M. Lodge, *Managing ecological thresholds in coupled environmental–human systems*, Proceedings of the National Academy of Sciences **108** (2011), no. 18, 7333.

[75] Thomas House, Geoffrey Davies, Leon Danon, and Matt J. Keeling, *A motif-based approach to network epidemics*, Bulletin of Mathematical Biology **71** (2009), no. 7, 1693–1706.

[76] D.J. Hruschka, L.M. Sibley, N. Kalim, and J.K. Edmonds, *When there is more than one answer key: cultural theories of postpartum hemorrhage in matlab, bangladesh*, Field Methods **20** (2008), no. 4, 315–337.

[77] W. Huang, K.L. Cooke, and C. Castillo-Chavez, *Stability and bifurcation for a multiple-group model for the dynamics of hiv/aids transmission*, SIAM Journal on Applied Mathematics (1992), 835–854.

[78] J. Joo and J.L. Lebowitz, *Pair approximation of the stochastic susceptible-infected-recovered-susceptible epidemic model on the hypercubic lattice*, Physical Review E **70** (2004), no. 3, 36114.

[79] E.H. Kaplan and R. Brookmeyer, *Snapshot estimators of recent hiv incidence rates*, Operations Research (1999), 29–37.

[80] G. Karabatsos and W.H. Batchelder, *Markov chain estimation for test theory without an answer key*, Psychometrika **68** (2003), no. 3, 373–389.

[81] GP Karev, *Dynamic theory of non-uniform population and global demography models*, J. of Biological Systems **13** (2005), 83–104.

[82] G.P. Karev, *Dynamics of heterogeneous populations and communities and evolution of distributions*, Arxiv preprint q-bio/0507028 (2005).

[83] _____, *Replicator equations and the principle of minimal production of information*, Bulletin of mathematical biology **72** (2010), no. 5, 1124–1142.

[84] GT Kasseem, S. Roudenko, S. Tennenbaum, C. Castillo-Chavez, A. Gumel, C. Castillo-Chavez, DP Clemence, and RE Mickens, *The role of transactional sex in spreading hiv/aids in nigeria: A modeling perspective*, Mathematical Studies on Human Disease Dynamics: Emerging Paradigms and Challenges. A. Gumel, C. Castillo-Chavez, DP Clemence, and RE Mickens, American Mathematical Society **410** (2006), 367–389.

[85] J.T. Kemper, *The effects of asymptomatic attacks on the spread of infectious disease: A deterministic model*, Bulletin of Mathematical Biology **40** (1978), no. 6, 707–718.

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

[87] P.L. Krapivsky and S. Redner, *Organization of growing random networks*, Physical Review E **63** (2001), no. 6, 066123.

[88] PL Krapivsky and S. Redner, *Finiteness and fluctuations in growing networks*, Journal of Physics A: Mathematical and General **35** (2002), 9517.

[89] P.L. Krapivsky, S. Redner, and F. Leyvraz, *Connectivity of growing random networks*, Physical Review Letters **85** (2000), no. 21, 4629–4632.

[90] C.M. Kribs-Zaleta, M. Lee, C. Román, S. Wiley, and C.M. Hernández-Suárez, *The effect of the hiv/aids epidemic on africa's truck drivers.*, Mathematical biosciences and engineering: MBE **2** (2005), no. 4, 771.

[91] K. Lewis, J. Kaufman, M. Gonzalez, A. Wimmer, and N. Christakis, *Tastes, ties, and time: A new social network dataset using Facebook. com*, Social Networks **30** (2008), no. 4, 330–342.

[92] J. Li, Z. Ma, S.P. Blythe, and C. Castillo-Chavez, *Coexistence of pathogens in sexually-transmitted disease models*, Journal of mathematical biology **47** (2003), no. 6, 547–568.

[93] Zonghua Liu, Ying-Cheng Lai, and Nong Ye, *Propagation and immunization of infection on general networks with both homogeneous and heterogeneous components*, Phys. Rev. E **67** (2003), no. 3, 031911.

[94] R.M. Lopez, B.R. Morin, and S.K. Suslov, *Logistic models with time-dependent coefficients and some of their applications*, Arxiv preprint arXiv:1008.2534 (2010).

[95] A. Mas-Colell, M.D. Whinston, J.R. Green, and Universitat Pompeu Fabra. Facultat de Cičncies Econňmiques i Empresarials, *Microeconomic theory*, vol. 1, Oxford university press New York, 1995.

[96] M.A. Miller, C. Viboud, M. Balinska, and L. Simonsen, *The signature features of influenza pandemicsimplications for policy*, New England Journal of Medicine **360** (2009), no. 25, 2595–2598.

[97] M. Mitchell, *Complexity: A guided tour*, Oxford University Press, USA, 2009.

[98] Y. Moreno, R. Pastor-Satorras, and A. Vespignani, *Epidemic outbreaks in complex heterogeneous networks*, The European Physical Journal B-Condensed Matter and Complex Systems **26** (2002), no. 4, 521–529.

[99] Benjamin R. Morin, Eli P. Fenichel, and Carlos Castillo-Chavez, *The mathematics of economic-epidemiology*, unpublished manuscript (2012).

[100] Benjamin R. Morin and Daniel Hruschka, *Model selection in test theory without an answer key with multiple, fixed* a priori *cultures*, unpublished manuscript (2010).

[101] B.R. Morin, C. Castillo-Chavez, S.F.H. Schmitz, A. Mubayi, and X. Wang, *Notes from the heterogeneous: a few observations on the implications and necessity of affinity*, Journal of Biological Dynamics **4** (2010), no. 5, 456–477.

[102] B.R. Morin, L. Medina-Rios, E.T. Camacho, and C. Castillo-Chavez, *Static Behavioral Effects on Gonorrhea Transmission Dynamics in a MSM Population*, Journal of Theoretical Biology (2010).

[103] Adilson E. Motter, Alessandro P. S. de Moura, Ying-Cheng Lai, and Partha Dasgupta, *Topology of the conceptual network of language*, Phys. Rev. E **65** (2002), no. 6, 065102.

[104] Adilson E. Motter, Takashi Nishikawa, and Ying-Cheng Lai, *Range-based attack on links in scale-free networks: Are long-range links responsible for the small-world phenomenon?*, Phys. Rev. E **66** (2002), no. 6, 065103.

[105] MEJ Newman, *The structure and function of complex networks*, SIAM Review **45** (2003), no. 2, 167–256.

[106] M.E.J. Newman, A.L. Barabasi, and D.J. Watts, *The structure and dynamics of networks*, Princeton Univ Pr, 2006.

[107] M.E.J. Newman, C. Moore, and D.J. Watts, *Mean-field solution of the small-world network model*, Physical Review Letters **84** (2000), no. 14, 3201–3204.

[108] Takashi Nishikawa, Adilson E. Motter, Ying-Cheng Lai, and Frank C. Hoppensteadt, *Heterogeneity in oscillator networks: Are smaller worlds easier to synchronize?*, Phys. Rev. Lett. **91** (2003), no. 1, 014101.

[109] A.S. Novozhilov, *On the spread of epidemics in a closed heterogeneous population*, Mathematical biosciences **215** (2008), no. 2, 177–185.

[110] National Institute of Allergy, Department of Health Infectious Diseases, National Institutes of Health, and Human Services, *Workshop summary: Scientific evidence on condom effectiveness for sexually transmitted disease (std) prevention.*, (2000).

[111] JD Oriel and AH Hayward, *Sexually-transmitted diseases in animals.*, The British journal of venereal diseases **50** (1974), no. 6, 412.

[112] J.S. Palmer, C. Castillo-Chavez, and S.P. Blythe, *State-dependent mixing and state-dependent contact rates in epidemiological models*, Biometrics Unit Technical Report BU-1122-M, Cornell University (1991).

[113] R. Pastor-Satorras and A. Vespignani, *Epidemic spreading in scale-free networks*, Physical review letters **86** (2001), no. 14, 3200–3203.

[114] _____, *Epidemic dynamics in finite size scale-free networks*, Physical Review E **65** (2002), no. 3, 035108.

[115] T. Philipson, *Economic epidemiology and infectious diseases*, Handbook of health economics **1** (2000), 1761–1799.

[116] D.J. Price, *NETWORKS OF SCIENTIFIC PAPERS.*, Science (New York, NY) **149** (1965), 510.

[117] T.C. Reluga, *Game theory of social distancing in response to an epidemic*, PLoS computational biology **6** (2010), no. 5, e1000793.

[118] MG Roberts and JAP Heesterbeek, *A new method for estimating the effort required to control an infectious disease*, Proceedings of the Royal Society of London. Series B: Biological Sciences **270** (2003), no. 1522, 1359.

[119] A.K. Romney, *Culture consensus as a statistical model*, Current Anthropology **40** (1999), 103–115.

[120] A.K. Romney, S.C. Weller, and W.H. Batchelder, *Culture as consensus: A theory of culture and informant accuracy*, American anthropologist **88** (1986), no. 2, 313–338.

[121] S.F.H. Schmitz, *Some theories, estimation methods and applications of marriage functions and two-sex mixing functions in demography and epidemiology*, Cornell University, January, 1994.

[122] S.F.H. Schmitz and C. Castillo-Chavez, *Parameter estimation in non-closed social networks related to dynamics of sexually transmitted diseases*, Modeling the AIDS Epidemic: Planning Policy and Prediction, New York, Raven Press Limited (1994).

[123] E. Shim, G.B. Chapman, and A.P. Galvani, *Decision making with regard to antiviral intervention during an influenza pandemic*, Medical Decision Making **30** (2010), no. 4, E64–E81.

[124] E. Shim and A.P. Galvani, *Evolutionary repercussions of avian culling on host resistance and influenza virulence*, PloS one **4** (2009), no. 5, e5503.

[125] J.C. Shlay, M.W. McClung, J.L. Patnaik, and J.M. Douglas Jr, *Comparison of sexually transmitted disease prevalence by reported condom use: errors among consistent condom users seen at an urban sexually transmitted disease clinic*, Sexually transmitted diseases **31** (2004), no. 9, 526.

[126] L.M. Sibley, D. Hruschka, N. Kalim, J. Khan, M. Paul, J.K. Edmonds, and M.A. Koblinsky, *Cultural theories of postpartum bleeding in matlab, bangladesh: Implications for community health intervention*, Journal of Health, Population, and Nutrition **27** (2009), no. 3, 379.

[127] M. Small, D.M. Walker, and C.K. Tse, *Scale-free distribution of avian influenza outbreaks*, Physical review letters **99** (2007), no. 18, 188702.

[128] R.D. Smith, M.R. Keogh-Brown, T. Barnett, and J. Tait, *The economy-wide impact of pandemic influenza on the uk: a computable general equilibrium modelling experiment*, BMJ: British Medical Journal **339** (2009).

[129] E. Stone, P. Heagerty, E. Vittinghoff, J.M. Douglas Jr, B.A. Koblin, K.H. Mayer, C.L. Celum, M. Gross, G.E. Woody, M. Marmor, et al., *Correlates of condom failure in a sexually active cohort of men who have sex with men*, JAIDS Journal of Acquired Immune Deficiency Syndromes **20** (1999), no. 5, 495.

[130] S. Tsutsui and K. Mino, *Nonlinear strategies in dynamic duopolistic competition with sticky prices*, Journal of Economic Theory **52** (1990), no. 1, 136–161.

[131] T.W. Valente, *Social network thresholds in the diffusion of innovations*, Social Networks **18** (1996), no. 1, 69–89.

[132] M. Vandeplassche, Food, and Agriculture Organization of the United Nations, *Reproductive efficiency in cattle: a guideline for projects in developing countries*, FAO, 1982.

[133] Yang Wang, Deepayan Chakrabarti, Chenxi Wang, and Christos Faloutsos, *Epidemic spreading in real networks: An eigenvalue viewpoint*, Reliable Distributed Systems, IEEE Symposium on **0** (2003), 25.

[134] J. Watmough and P. van den Driessche, *Reproduction numbers and sub-threshold endemic equilibrium for compartmental models of disease transmission*, Math. Biosci **180** (2000), 29–48.

[135] S.C. Weller, *Shared knowledge, intracultural variation, and knowledge aggregation.*, American Behavioral Scientist (1987).

[136] _____, *Cultural consensus theory: Applications and frequently asked questions*, Field Methods **19** (2007), no. 4, 339–368.

[137] Liang Zhao, Kwangho Park, and Ying-Cheng Lai, *Attack vulnerability of scale-free networks due to cascading breakdown*, Phys. Rev. E **70** (2004), no. 3, 035101.

BIOGRAPHICAL SKETCH

Benjamin Morin earned his B.A. in Mathematics from the University of Maine in Orono with honors after participating in undergraduate research on the fractal dimension of the boundaries of basins of attraction for a forced/damped pendulum. His first advanced degree was an M.A. from the University of Maine in Orono under the direction of David E. Hiebeler (Hiebeler's first graduate student). After a tour at Oregon State University he earned an M.S. in Mathematics under the direction of Mina Ossiander with a focus on Ecosystems Informatics. It was there that the precepts of Carlos Castillo-Chavez really hit home, and thus Benjamin transferred to Arizona State University in the Applied Mathematics for the Life and Social Sciences. Ben is happily married to his wife Lydia and they have a wonderful son, Ari Xavier Morin, who will be two at the time of defense.