Integrating Mathematical Models of Behavior and Infectious Disease: Applications to Outbreak Dynamics and Control

by

Michael Akira Lee Hayashi

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Epidemiological Science)
in The University of Michigan
2016

Doctoral Committee:

Assistant Professor Marisa C. Eisenberg, Chair
Professor Joseph N.S. Eisenberg
Assistant Professor David W. Hutton
Assistant Professor Rafael Meza

For Elise

ACKNOWLEDGEMENTS

To my committee, Joseph Eisenberg, David Hutton, and Rafael Meza, this work would not have been possible without your input and contributions. To my friends, both within and outside the Department of Epidemiology, my time in the doctoral program has been overwhelmingly and improbably excellent on your account. To my family, I will always be grateful for your love, support, and encouragement as I have grown from a quiet, round baby to a slightly less quiet, much less round research scientist. To my adviser, Marisa Eisenberg, truly you have been the best of all possible mentors, as well as a peerless colleague and friend. Finally, I am fortunate to have received a Francis Fellowship from the Department of Epidemiology to fund my doctoral studies and research. This funding mechanism provided me the freedom to pursue research topics and methodological developments that are currently not well covered by specific faculty research grants.

PREFACE

Chapter III has been published with Marisa C. Eisenberg in *The Journal of Theoretical Biology*. A version of Chapter IV will be submitted for publication with the following authors: Michael A.L. Hayashi and Marisa C. Eisenberg. A version of Chapter V will be submitted for publication with the following authors: Michael A.L. Hayashi, Marisa C. Eisenberg, David W. Hutton, and Joseph N.S. Eisenberg.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

Table

LIST OF APPENDICES

Appendix

CHAPTER I

Introduction

Human behavior is an essential and complex factor of infectious disease transmission systems. Individuals can change risky or protective behaviors over time in response to outbreaks or other external conditions such as political or environmental shifts. In addition, social interactions provide opportunities for pathogens to spread. Finally, epidemiological interventions often require cooperation from their target population. Indeed, almost all interventions require the consideration of behavior, changing or otherwise. Modern infectious disease control programs increasingly seek to address these concerns, yet disease modeling studies rarely include an explicit behavioral framework. To do so, mathematical models will need to be developed that can accommodate behavioral features such as adaptation that are frequently simplified in current approaches.

In higher income countries, developments in sanitation, hygiene, treatment, and vaccination have eradicated or nearly eradicated many major sources of disease-related mortality. Polio, measles, and smallpox have been controlled by aggressive vaccination campaigns, while the burden of sexually transmitted infections (STIs) such as syphilis and HIV have been reduced through the promotion of safe-sex and treatment campaigns. However, maintaining successful control is challenging, due

in part to changes in behavior over time. Recent measles outbreaks suggest that vaccine fatigue and anti-vaccination sentiment have reduced childhood vaccination rates, often counter-intuitively in wealthier, well educated communities [1–3]. Similarly, the relative incidence of treatable STIs has increased or remained high among adolescents, minorities, and men who have sex with men [4–6]. Promoting the use of prophylaxis and testing and treatment services are crucial, but adherence with both is likely to vary depending on the degree of perceived risk. In particular, when there is little risk of infection, individuals may be less inclined to practice safe sex, potentially initiating new outbreaks [7–9]. Currently, relatively few modeling studies of either vaccine-preventable or sexually transmitted diseases explicitly include behavior change. Those that do have begun to use methods from economics, specifically game theory, to handle behavior-disease feedback [10–17]. These economic models are flexible and have a well developed theoretical foundation making them suitable for integration with existing techniques in mathematical epidemiology.

Emerging disease outbreaks present important opportunities to apply mathematical models to inform public health responses. However, under these circumstances the costs of model mis-specification are high, both due to the humanitarian and economic ramifications of mis-prediction as well as the degree of public attention focused on new outbreaks. The 2014 Ebola virus epidemic in West Africa reflects both the value and hazards of modeling in the early stages of an outbreak. Model projections were used by global health agencies to estimate the scope of transmission and allot burial teams and treatment units. However, many of the early models predicted a larger number of cases than actually occurred [18–22]. Even though projections were typically reported as worst-case scenarios, mathematical modeling was criticized for producing inaccurate and alarmist forecasts. While mis-prediction for Ebola does

not necessarily reflect a fundamental fault in methodology, later findings do suggest the importance of considering behavioral factors [23]. Scaling the population by a factor to represent the population at risk and reporting rate improves model estimation and short-term forecasting [24]. A feedback loop between the number of cases, the strength of control responses, and changes in protective behavior among the populations affected may also have contributed to errors in predictions from Ebola models.

Epidemic and emerging diseases are both important, challenging classes of problems for infectious disease epidemiology. A third class, the control of endemic diseases in lower income countries is an equally difficult but crucial objective of global public health. Among these, waterborne diarrheal diseases remain the second leading cause of death among children under five years [25]. Diarrheal disease also incurs large costs in terms of lost productivity and reduced quality of life for both children and adults. Improving drinking water security is therefore a major objective in the United Nations (UN) Sustainable Development Goals (SDG) for 2030 [26]. Doing so will require large scale expansions of existing water treatment programs as well as the development of new intervention strategies.

From an intervention design and implementation perspective, it is necessary to consider whether control measures are compatible with regional attitudes and customs. Decentralized water, sanitation, and hygiene (WASH) interventions exemplify these challenges. Household water treatment (HWT) can achieve high efficacy, removing over 99.9999% of pathogens in small scale trials. However, HWT often underperforms when implemented at a population level [27–30]. Low uptake and continued use may be due to undesirable usability and reliability characteristics of HWT methods. For example, boiling requires a large amount of energy and does

not scale well with large volumes of water. Chemical treatment can alter the taste of water such that it becomes unpalatable [31, 32], while highly efficacious filters can have slow flow rates and are prone to clogging [33–35]. Designing and evaluating interventions without considering these factors may perpetuate poor compliance and impair progress toward development and health targets.

The work presented here demonstrates how mathematical models of behavior can be integrated with transmission models to address the three infectious disease contexts above. This dissertation is structured as follows: Chapter II contains a summary of mathematical modeling approaches for infectious disease and human behavior. Chapter III investigates the effects of adaptive prophylaxis use for the prevention of sexually transmitted infections by integrating Bayesian game theory and evolutionary dynamics with the susceptible-infected-susceptible (SIS) model. Chapter IV addresses changes in burial practices during the 2014 Ebola virus outbreak in West Africa. This work addresses parameter estimation, model selection, and forecasting for coupled behavior-disease models. Chapter V presents a decision theory augmented quantitative microbial risk assessment (QMRA) model for the evaluation of household water treatment (HWT) interventions with incomplete compliance. This model connects behavioral modeling and risk assessment to intervention policy evaluation. Finally, Chapter VI contains concluding remarks. Throughout the following chapters we focus on both methodological developments as well as insights gained regarding infectious disease dynamics and the implementation of control programs.

CHAPTER II

Background

2.1 Modeling infectious disease

2.1.1 Compartmental transmission models

Mathematical models are frequently used in infectious disease epidemiology to study transmission processes. The SIR compartmental model,first developed by Kermack and McKendrick in 1927 [36], is a fundamental approach from which numerous others have been derived. The population is divided into compartments that represent stages of a disease and individuals move between compartments based on rate equations. These can be represented as a system of differential equations such as the example given below:

$$\begin{aligned} \dot{S} &= \mu - \beta SI - \mu S \\ \dot{I} &= \beta SI - \gamma I - \mu I \\ \dot{R} &= \gamma I - \mu R \end{aligned} \tag{2.1}$$

where S, I, and R are the fractions of the population that are susceptible, infectious, and recovered respectively. β represents the transmission rate, the product of the contact rate and the probability of transmission per contact. γ is the recovery rate. This model assumes that new individuals are born into the susceptible compartment at rate μ, which is also the per-capita death rate, maintaining a constant population size.

The popularity of compartmental models is due in part to their flexibility. Disease natural history can be captured by extending the infectious compartments, an approach often used for HIV, Tuberculosis, and recently, Ebola. Contact patterns can similarly be addressed by classifying groups according to their activity levels.

Most infectious disease transmission models are nonlinear and cannot be solved analytically. However, many algorithms exist to numerically solve systems of differential equations [37]. In addition, it is possible to calculate some epidemiologically relevant values such as steady states without solving the full system of equations.

The basic reproduction number ($\mathcal{R}_0$) The basic reproduction number is a key concept in infectious disease epidemiology [38]. This term characterizes the expected number of secondary infections caused by a single infected individual in a completely susceptible population over the duration of their infection [38–40]. When $\mathcal{R}_0 < 1$ there are not enough new cases to sustain transmission. By contrast when $\mathcal{R}_0 > 1$ the number of cases grows over time, leading to an outbreak. Given that this threshold acts as a condition for the stability of the disease free equilibrium (DFE) in a compartmental model, it is useful for outbreak control. Interventions which can reduce $\mathcal{R}_0$ below 1 are expected to prevent the spread of infection through a population. $\mathcal{R}_0$ can be calculated by linearizing the system of differential equations about the DFE, or using the Next-Generation Matrix approach [40] for more complex models.

2.2 Modeling human behavior

Mathematical models of human behavior have been a subject of interest among social and physical scientists as early as the 17th century. Contributions by Blaise Pascal, Daniel Bernoulli, the Marquis de Condorcet, and others integrated proba-

bility theory, enlightenment philosophical concepts, and Newtonian physics to form the basis of modern decision and social choice theory. Later, John von Neumann and Oskar Morgenstern developed the formal specifications for decision and game theory in their seminal work *The Theory of Games and Economic Behavior* [41]. These methods are widely used in economics and political science, but have been adopted by other fields as well. The development of evolutionary game theory [42] provided an extended framework to model animal competition and natural selection for theoretical ecology. Both classical and evolutionary game theory have been used to represent behaviors in transmission models. For example, Reluga and Galvani developed a differential game approach to vaccination and social distancing [11, 43, 44], while Bauch, Reluga, and others have used evolutionary dynamics to model changes in vaccination behavior [10, 13]. In general however, integrating economic models of behavior into epidemiological models remains uncommon.

2.2.1 Decision theory

Decision problems represent situations where an individual actor must make a choice between some set of alternatives. Formally, a decision problem is characterized by the following features:

- Actions: the choices available to an individual.
- Outcomes: the result of each possible action.
- Preferences: the ranking an individual assigns to the set of outcomes.

We denote the set of actions A and the set of outcomes X. Often the mapping from A to X is one-to-one and actions and outcomes are described interchangeably. Formally, preferences are characterized by a *preference order* R. We use the notation $\succeq$ to indicate a *preference relation* under R. For a pair of outcomes $x, y \in X$, $x \succeq y$

indicates that x is at least as good as y. Intuitively a preference order should allow an individual to decide which action or actions lead to the best outcome. As such, a preference order should satisfy two properties: *completeness* and *transitivity*.

Definition II.1. A preference order R is **complete** if and only if for any pair of outcomes $x, y \in X$, either $x \succeq y$, $y \succeq x$ or both.

Completeness states that when an individual is presented with two potential choices they must rank them according to how much they like each outcome. This allows individuals to be indifferent by ranking two alternatives equally, but does not allow them to be indecisive by refusing to provide a ranking.

Definition II.2. A preference order R is **transitive** if and only if for any three outcomes $x, y, z \in X$, $x \succeq y$ and $y \succeq z$ implies that $x \succeq z$.

Transitivity precludes the possibility of cycles in a preference order. This imposes consistency between binary comparisons and the preference order as a whole, i.e., if an individual would choose x over y, they should also choose x over anything that they prefer less than y. Together, completeness and transitivity guarantee that an individual will always have at least one most-preferred alternative. A preference order satisfying these properties is known as a *rational preference order*.

It is possible and often convenient to represent rational preference orders using a functional form denoted the utility or payoff function.

Definition II.3. A **utility (payoff) function** $u : X \to \mathbb{R}$ represents the preference order R if and only if for any pair of outcomes $x, y \in X$, when $x \succeq y, u(x) \geq u(y)$.

One key difference between utility functions and preference orders is whether alternatives are compared based on a cardinal or ordinal relation, respectively. As a result, in addition to specifying whether one alternative is preferred to another, a

utility function characterizes the degree to which it is preferred. This is advantageous in some circumstances but requires stronger assumptions about how individuals form preferences.

In order to determine which choice an individual is likely to make, decision theory assumes that actors are *rational*. In an economic context, this is defined as follows:

Definition II.4. A **rational actor** knows the set of alternatives A and outcomes X, forms a rational preference order R over these outcomes, and always chooses an action a^* that corresponds to the most preferred outcome x^* such that for any other action $y \in X$, $x^* \succeq y$.

Note that economic rationality is not identical to colloquial rationality – a most-preferred action need not be a "good" choice from any perspective other than the individual making it. However, the assumption of rationality remains actively debated as behavioral experiments often demonstrate apparently irrational choices among participants [45, 46]. While these results are noteworthy, it is not clear whether they demonstrate truly irrational behavior or an incorrect specification of available actions and preferences [47].

Like a preference order, a utility function can be used to determine which alternative an individual is likely to choose. Doing so requires finding the outcome that maximizes an individual's utility,

$$x^* = argmax_x u(x). \tag{2.2}$$

Based on the definition of a utility function, a maximum utility outcome must also be at the top of the preference order.

2.2.2 Game theory

Game theory was developed in the 1940's by von Neumann and Morgenstern [41] to address decision problems involving more than one interacting individual. In a general game, individuals (referred to as players) select actions from discrete or continuous alternatives and receive payoffs based on the outcome mapped to the set of all players' choices. As an extension of decision theory, the formal specification of a game theoretic model is similar. A *normal form game* has the following:

- A set of individuals $N = \{1, 2, ..., n\}$.
- Each individual i chooses an action a_i from a set A_i.
- A *profile* of actions for all individuals $a = (a_1, a_2, ..., a_n)$ defines an outcome.
- Each individual has a complete and transitive preference order R_i over outcomes.

Actions in a game-theoretic model typically refer to discrete alternatives. More generally, these may be referred to as strategies, where a strategy s_i can be either a specific action, or a probability distribution over multiple actions. We refer to the former as a *pure strategy* and the latter as a *mixed strategy*. A strategy profile s specifies one strategy for each player. As in decision theory, we can use a utility function to represent a player's preference order R_i. However, now the utility function takes a strategy profile as its argument to reflect that players' payoffs depend on each others' choices. Thus, $u_i : S \to \mathbb{R}$ where S is the set of all possible strategy profiles.

Matrix representation Payoffs for a normal form game can be represented using a payoff matrix M. Consider the Prisoner's Dilemma, a classic strategic form game. Two prisoners must independently decide whether to remain quiet (cooperate) or confess, implicating the other (defect). If neither prisoner defects, both receive a

sentence of one year. If one cooperates but the other defects, the defector is released while the cooperator is sentenced to ten years. If both defect, they receive a sentence of five years. This game has the following payoff matrix:

	C	D
C	(-1,-1)	(-10,0)
D	(0,-10)	(-5,-5)

where the row entries correspond to Player 1's strategies and the column entries correspond to Player 2's strategies. Each cell gives the payoff to each player of the specific strategy combination. For example, $m_{1,2} = (u_1(C, D), u_2(D, C)) = (-5, 0)$. Note that the Prisoner's Dilemma is a member of a special class of games where the payoffs to each player are symmetrical. For such games we need only specify the payoffs for one player.

Nash equilibrium Solving a game theoretic problem introduces the complication of determining which choice is rational given the choices other players may make. A fundamental solution concept for all games is the Nash equilibrium [48], a set of choices from which no player can improve their payoff by changing their action. Formally,

Definition II.5. A strategy profile s^* is a **Nash equilibrium** if and only if $u_i(s^*) \geq u_i(s'_i, s^*_{-i})$ for all players $i \in N$ and all strategies $s'_i \in S_i$.

That is, player i's utility for any other strategy s'_i is lower than s^*_i given that all other players $-i$ are playing according to s^*. This guarantees that no player has an incentive to deviate from the Nash equilibrium. Similarly, a *strict* Nash equilibrium requires that the inequalities be strict. Equivalently, we can describe a Nash equilibrium as a set of mutual best responses.

Definition II.6. A strategy s_i^* is a **best response** if and only if $u_i(s_i^*, s_{-i}) \geq u_i(s_i', s_{-i})$ for all s_i'.

I.e., the strategy that maximizes a player's utility, fixing all other players' strategies. Conveniently, every game is guaranteed to have at least one Nash equilibrium in pure or mixed strategies. Revisiting the Prisoner's Dilemma, we can solve for its Nash equilibrium by examining the payoff matrix.

	C	D
C	$(-1, -1)$	$(-10, \overline{0})$
D	$(\underline{0}, -10)$	$(\underline{-5}, \overline{-5})$

Underlined payoffs correspond to Player 1's best response strategies while over-lined payoffs correspond to Player 2's best response strategies. Regardless of whether the other player chooses to cooperate or defect, a player receives a higher utility from defecting. Thus the Nash equilibrium of the Prisoner's Dilemma is for both players to defect. This example demonstrates an important feature of Nash equilibria: Equilibrium strategy profiles do not always lead to socially optimal outcomes. Had both players cooperated, their average payoff would have been higher, but then either player would have had an incentive to switch to defect.

Subsequent developments in political economy have focused on refinements to the Nash equilibrium concept, allowing analysis of more complex games where the assumptions of a classical game may not hold [49–51]. In particular, Bayesian games allow players to have incomplete information by assigning players belief distributions over the types of other players (Appendix A.1).

2.2.3 Evolutionary game theory

While decision and game theory were developed with the intention of modeling human behavior, their structures do not preclude application to other domains. Evolutionary game theory (Maynard-Smith and Price [42, 52]) treats the problem of natural selection in interacting animal populations. Seminal examples involve competition for food resources and the evolution of cooperation. Unlike the previous methods, evolutionary game theory does not assume that agents are rational. Instead, *evolutionary dynamics* govern changes in strategies over time [53, 54]. Payoffs in an evolutionary context are referred to as fitness, often based on the number of offspring produced by animals playing a given strategy.

The replicator equation is a common form of evolutionary dynamics [53]. Broadly, a strategy will increase in frequency if its fitness is greater than the average fitness and decrease otherwise. Let x_k be the frequency of the kth strategy. Replicator dynamics for x_k are:

$$\dot{x_k} = x_k[u_k(x) - \bar{u}] \tag{2.3}$$

where $u_k(x)$ is the fitness of the kth strategy. Note that fitness is a function of the relative frequencies of all strategies. $\bar{u}$ is the average fitness. For a symmetrical game, $u_k(x) = \sum_j x_j m_{kj}$ and $\bar{u} = \sum_k u_k(x) x_k$.

Even though replicator dynamics allow players in an evolutionary game to act irrationally, a fundamental result links evolutionary dynamics to classical game theory. The Folk Theorem of Evolutionary Game Theory [53, 55] demonstrates that stable steady states of the replicator equation are Nash Equilibria and Nash equilibria are steady states of the replicator equation. This can often simplify the analysis of evolutionary games, as they can be treated as strategic form games for the purpose of determining long-term behavior.

Replicator dynamics can be extended to allow mutation, producing the replicator-mutator equation [54]:

$$\dot{x_k} = \sum_j x_j u_j(x) Q_{jk} - \bar{u} x_k \tag{2.4}$$

where Q is a matrix that specifies the probability of a random change between any pair of strategies.

While replicator (and replicator-mutator) dynamics were developed to treat generations of organisms adopting phenotypic traits, they can be interpreted in a behavioral context as well. Suppose individuals determine whether to change their strategy by observing others, switching with some probability if the payoff received by a sampled player is higher than their own. This process is known as *imitation dynamics*, and is provably equivalent to replicator dynamics with a scaling factor for the rate of sampling. Similarly, mutation can represent random changes between behaviors. Evolutionary game theory can therefore be a useful tool for modeling changing human behavior – the framework includes differential equations for strategy change and does not require rationality.

CHAPTER III

Effects of adaptive protective behavior on the dynamics of sexually transmitted infections

3.1 Introduction

In spite of advances in treatment and prevention, sexually transmitted infections (STIs) remain endemic worldwide. The CDC estimates that 20 million new cases occur annually in the United States alone [5], incurring a total cost of $16 billion for treatment and care. Globally, treatable STIs are responsible for approximately 500 million new cases per year [4], while an estimated 35 million individuals currently live with HIV. These statistics underscore the importance of understanding the dynamics that drive and sustain STI transmission. To this end, mathematical epidemiology has made substantial progress investigating the role of contact patterns such as age-structure, sexual networks, and levels of sexual activity [56, 57]. However, many open questions remain in understanding the feedback relationship between behavioral change and disease dynamics.

From a behavioral standpoint, sexually transmitted diseases are noteworthy as they require a direct and intimate interaction between individuals. As a result, many common preventative measures, such as condom use, are not determined unilaterally [58, 59]. In addition, assuming individuals form preferences over protective behaviors based on the associated costs and benefits, we would expect these behav-

iors to respond to the risk of infection as an outbreak progresses [60–62]. While changes in risky or protective behavior can amplify or suppress outbreaks, adoption of these behaviors can in turn be driven by the spread of disease, as demonstrated by increased testing and condom use among men who have sex with men in response to the HIV outbreak in the US [8, 63].

Methods from game theory provide a framework with which to capture this feedback, grounded in well-established mathematical and economic theory. The resulting economic- epidemiological models can explicitly represent the decision process of individuals either in direct interactions (e.g. sexual encounters) or population interactions (e.g. vaccination behavior) [10–13, 15, 16, 43]. Including the effects of behavioral change on STI dynamics has been primarily motivated by the HIV epidemic among men who have sex with men (MSM) in the 1980s and 1990s, but this modeling approach is relevant for the study of other pathogens and communities as well. Indeed, in 2013 the WHO highlighted the need to study behavioral change in order to design effective interventions [4].

Many economic-epidemiological models rely on two behavioral assumptions that are worth consideration. The traditional game theoretic framework assumes that all actors are fully rational, responding optimally at every stage of play [49]. While convenient, the rationality assumption remains a subject of debate in economic literature. Empirical studies note circumstances in which individual behavior appears to depart from a strict payoff maximization foundation [45–47]. In addition, most models (e.g.[14] and [17]) assume that only susceptible individuals make choices regarding protective behavior. This assumption is the result of representing the cost-benefit calculus of individuals as a tradeoff between various private costs of protective behavior and the risk of infection. However, infected individuals may have incentives

to reduce contact or use protection as well, motivated by altruism, self interest, or other factors [64, 65]. This has been the focus of several intervention strategies in practice [66, 67].

In this paper, we propose a model of combined behavioral and disease transmission dynamics that uses the outcome of sexual interactions between susceptible and infected individuals to determine the effective contact rate for a mass action model of disease transmission. The combined model bears some similarity to the behavior-disease model proposed in [15]. There are, however, several critical distinctions. We use a deterministic ODE framework, while our game-theoretic model collapses the protection-use game to a single interaction instead of a multi-stage negotiation. In addition, similar to [13] and [10] we use evolutionary dynamics to represent the process of behavioral change over time, allowing for non-optimal but potentially more realistic behaviors. Unlike the inductive reasoning game developed by Breban et al. [68], our behavioral dynamics only explicitly consider the current state instead of a history of actions. While this approach loses some realistic features, it still allows us to relax the assumption of full rationality while also providing a convenient mathematical formulation for the combined model [42, 53].

3.2 Model

The Susceptible-Infectious-Susceptible (SIS) model has been studied extensively as a simplified representation of bacterial sexually transmitted diseases [38, 69, 70]. The model equations are

$$\begin{aligned} \dot{S} &= \gamma I - \beta SI, \\ \dot{I} &= \beta SI - \gamma I, \end{aligned} \tag{3.1}$$

where S is the fraction of the population which is susceptible, I is the fraction infected, β is the effective contact rate; the product of the rate of sexual partner ac-

quisition and the probability of disease transmission from an infected to a susceptible partner, and γ is the rate of recovery or treatment. The basic reproductive ratio is

$$\mathcal{R}_0 = \frac{\beta}{\gamma}. \tag{3.2}$$

The disease free equilibrium (DFE) occurs if $\mathcal{R}_0 < 1$. Otherwise the endemic prevalence is

$$I^* = 1 - \frac{\gamma}{\beta}. \tag{3.3}$$

In order to capture the potential for individuals to adapt their protective behaviors over the course of an outbreak, we define a Bayesian game [51] between a pair of players. In A.1, we give a brief overview of the definitions and structure of Bayesian games, with more complete treatments given in [49–51]. The payoffs for the game depend on the disease states of both players, which are considered private information. Players must infer the type of their partner, reflecting realistic uncertainty about serostatus [7, 9, 65, 71, 72]. Consistent with the notation for the SIS model, a player may be one of two types chosen from the type space $\Theta = \{S, I\}$. Each player chooses between using protection (P) or no protection (U) for a given sexual encounter. If both players select the same action, the outcome of the game is the same as the chosen action. We assume that if both players select different actions the encounter does not proceed and the effective contact rate for the pair of players is 0. In order to characterize the strategy space for this game, it is convenient to use the type-contingent notation $\sigma_j(\theta_i) = a_i \in \{P, U\}$[1] to denote the action player i would choose if she was of type θ_i under the jth strategy. A complete strategy for a player then specifies a pair of actions $\sigma_j := \sigma_j(\theta_i = S)\sigma_j(\theta_i = I) \in \{PP, PU, UP, UU\}$. Without loss of generality, the type-dependent payoff (utility) to player 1 for a given pair of actions and types is written $u_1(\sigma_j(\theta_1), \sigma_k(\theta_2), \theta_1, \theta_2)$ (where the first argument

[1] Players in this game are interchangeable so we do not specify distinct strategy sets for each player.

specifies player 1's action assuming type θ_1 under strategy j and the second entry player 2's action assuming type θ_2 under strategy k). Then player 1's overall expected payoff for a particular strategy profile (i.e. for a pair of type-contingent strategies σ_j and σ_k for each player) is the double expectation of $u_1(\sigma_j(\theta_1), \sigma_k(\theta_2), \theta_1, \theta_2)$ over both players potential types [49], that is,

$$E[u_1(\sigma_j(\theta_1), \sigma_k(\theta_2), \theta_1, \theta_2)] = \sum_{\theta_1} Pr(\theta_1)[\sum_{\theta_2} p_1(\theta_2|\theta_1) u_1(\sigma_j(\theta_1), \sigma_k(\theta_2), \theta_1, \theta_2)] \quad (3.4)$$

where $E[u_1(\sigma_j(\theta_1), \sigma_k(\theta_2), \theta_1, \theta_2)]$ is sometimes written more simply as $E[u_1(\sigma_j, \sigma_k)]$. The probabilities for each player being of either type, $Pr(\theta \in \{S, I\})$, are given by the distribution of susceptible and infected individuals in the population. This distribution acts as the common prior for the Bayesian game. We assume that partner selection is not assortative in disease type, so the belief for player 2 being type θ_2 given player 1 being type θ_1 is $p_1(\theta_2|\theta_1) = Pr(\theta_2)$.

Fig. 3.1 shows the two possible payoff matrices for the protected sex game that determine the type-dependent payoff terms $u_i(\sigma_j(\theta_1), \sigma_k(\theta_2), \theta_1, \theta_2)$ in Equation 3.4. For a concrete example, we compute the expected payoff to player 1 of playing strategy $\sigma_j = PU$ if player 2 picks $\sigma_k = UU$ when 30% of the population is infected and we choose $a = 1, b = 0.75, c = 0.5, d = 0.25$. From Equation 3.4 we have

$$\begin{aligned} E[u_1(PU, UU)] &= 0.7(0.7 \times 0.5 + 0.3 \times 0.25) \\ &\quad + 0.3(0.7 \times 0.25 + 0.3 \times 1) \\ &= 0.44. \end{aligned} \quad (3.5)$$

In general, we suppose the specific payoff entries satisfy $a > b > c > d$. Thus, individuals prefer unprotected sex to protected sex with a partner of the same disease status, preferring either to action pairs resulting in no sexual encounter. However, individuals prefer protected sex to all other outcomes when their partner is of a dif-

		Player 2 P	Player 2 U
Player 1	P	b	c
	U	c	a

(a)

		Player 2 P	Player 2 U
Player 1	P	b	c
	U	c	d

(b)

Figure 3.1: The two variants of the protected sex game for the cases a: $\theta_i = \theta_j$ and b: $\theta_i \neq \theta_j$.

ferent disease type. This is intended to capture the notion that both susceptible and infected individuals have some incentive to avoid infection or transmission respectively. We explore the alternate case where infected individuals do not distinguish between susceptible and infected partners in A.5.

Consequently, for the protected sex game, the time-varying payoff matrix M is a 4×4 square matrix with elements given by Eq. (3.4). We can compute the Bayes-Nash equilibrium for any given choice of a, b, c, d, S, I. One case of particular interest is at the disease free equilibrium, $I^* = 0$. Here the game reduces to the 2×2 game with payoffs as in Figure 3.1a. This game has two pure strategy Nash equilibria, (P, P) and (U, U). However, the (U, U) equilibrium in the reduced game is both payoff and risk dominant. Since the $\sigma(\theta_i = I)$ actions do not contribute to the expected payoff, any mixture of the type-contingent strategies UP and UU (similarly PP and PU) is a Nash equilibrium in the full model. We will revisit this point shortly.

Since we are ultimately interested in the dynamics of behavior at more than a single disease state, we use methods from evolutionary game theory to couple the dynamics of strategy change over time to the disease trajectory. This allows us to close the feedback loop between behavior and disease dynamics. In particular, we use replicator-mutator dynamics with a linear fitness function [53, 73]. The distribution of strategies in the population using the replicator-mutator approach depends both on the existing distribution of strategies and their expected payoffs, as well as a

small degree of random strategy choice. This allows us to capture the notion that individuals may not respond immediately or strictly optimally to the presence of an outbreak, as payoff-suboptimal strategies may remain frequent in the population for some time. In addition, the possibility of mutation prevents any strategy from becoming extinct. In the behavioral context, this can be thought of as allowing for a small amount of random choice. This feature is particularly important as the replicator equation has stable steady states at strict pure strategy Nash equilibria, such that behavior cannot change after fixation on a particular Nash equilibrium even if disease conditions change. Both players in the protected sex game have the same action and type sets, so the relative frequency of the jth strategy in a large population is $f(\sigma_j, t)$. The replicator-mutator equation for a given strategy σ_j is

$$\dot{f}(\sigma_j) = s[\sum_k f(\sigma_k, t) q_{kj} \phi(\sigma_k, t) - \bar{\phi}(t) f(\sigma_j, t)], \tag{3.6}$$

where q_{kj} is the probability that an individual playing σ_j switches to σ_k, $\phi(\sigma_j, t) = \sum_k E[u(\sigma_j, \sigma_k)] f(\sigma_k, t)$ is the fitness of σ_j, $\bar{\phi}(t) = \sum_j \phi(\sigma_j, t) f(\sigma_j, t)$, and $s \in [0, \infty)$ is a scaling term that determines the speed of behavior change. In vector-matrix form, this can be written as

$$\dot{\mathbf{f}} = s[D_{\mathbf{f}} Q(M\mathbf{f}) - (\mathbf{f}^T M \mathbf{f})\mathbf{f}], \tag{3.7}$$

where M is the 4×4 payoff matrix with elements $m_{jk} = E[u_1(\sigma_j, \sigma_k)]$, $\mathbf{f} = (f(PP,t), f(PU,t), f(UP,t), f(UU,t))$, Q is the mutation matrix and $D_{\mathbf{f}} = diag(\mathbf{f})$. Note that when $Q = I$, the evolutionary dynamics are equivalent to standard replicator dynamics. However, for our simulations we use a mutation probability $\mu = 0.03$ so

$$Q = (1 - \frac{4}{3}\mu)I + \frac{\mu}{3} \cdot \mathbf{1}. \tag{3.8}$$

In a large population, the aggregate effective contact rate is determined by the average outcome over all pairs, so $\beta = \frac{\beta_b S_U I_U}{SI}$ where β_b is the baseline effective contact rate, $S_U = S \times (f(UP,t) + f(UU,t))$ is the proportion of susceptible individuals playing U, and $I_U = I \times (f(PU,t) + f(UU,t))$ is the proportion of infected individuals playing U.

The combined model equations with evolutionary behavioral dynamics can be written as

$$\begin{aligned} \dot{S} &= \gamma I - \beta_b S_U I_U, \\ \dot{I} &= \beta_b S_U I_U - \gamma I, \\ \dot{\mathbf{f}} &= s[D_{\mathbf{f}} Q(M\mathbf{f}) - (\mathbf{f}^T M \mathbf{f})\mathbf{f}]. \end{aligned} \tag{3.9}$$

The system above has six compartments. Of these, only four are strictly necessary as $S + I = 1$ and $\sum_j f(\sigma_j, t) = 1$. In a completely susceptible population, action pairs where both susceptibles play P or U are Bayes-Nash equilibria. In order to restrict the domain of possible initial conditions, we take the risk dominant equilibria, or mixtures between UP and UU. This can be interpreted as individuals in a disease free state preferring unprotected sex because it best hedges against uncertainty in partner actions. While the $\sigma(\theta_i = I)$ actions are never realized at the DFE, the type-contingent strategy framework implies that individuals must be able to specify an action that they would take if they became infected. The effective contact rate for a single infected individual is $(f(UU,0) + f(PU,0))\beta_b$, where $f(UU,0), f(PU,0) \in [0,1]$. As a result, the basic reproductive rate of the combined model is

$$\mathcal{R}_0 = \frac{(f(UU,0) + f(PU,0))\beta_b}{\gamma}. \tag{3.10}$$

Thus, $\mathcal{R}_0$ depends on the behavioral initial condition, unlike the standard SIS model. The value of the above expression will lie within the interval $[0, \beta_b/\gamma]$. The payoff and risk dominance of the (U, U) Nash equilibrium at the initial disease-free steady state (discussed above) suggests that $f(PU, 0)$ is likely to be low, and indeed for the remainder of this paper we will assume that the initial infected plays the strategy UU, while the underlying susceptible population at $t = 0$ will be assumed to take on a mixture of the UU and UP strategies. It is nonetheless of interest to note that the dependence of $\mathcal{R}_0$ on the initial behavioral conditions suggests the potential for both high $\mathcal{R}_0$ with disease extinction and low $\mathcal{R}_0$ with endemic prevalence: if $f(PU, 0) \approx 1$ it may be possible to have high $\mathcal{R}_0$ but still have disease extinction due to all susceptibles playing the protected strategy, while conversely if $\mathcal{R}_0$ is low due to low frequency of UU and PU but otherwise has a high contact rate, it may be possible to generate endemic prevalence even for $\mathcal{R}_0 < 1$ (as explored further below).

3.3 Results

Due to the addition of the replicator-mutator equations, it is not straightforward to solve for the steady states of the combined model analytically. Instead, we conducted a range of numerical simulations to investigate the effects of behavioral dynamics using Python 2.7 with Numpy 1.9, Scipy 0.15.0, and Matplotlib 1.4.2. In order to explore a range of outcomes generated by the combined model, we focused on both long-term and short-term dynamics. Given the addition of behavior change, the short term dynamics provided useful mechanistic insights to explain the steady state properties of the model. Unless noted otherwise we used $a = 1, b = 0.6, c = 0.4, d = 0.2$ for the values of the outcome payoffs. In general the qualitative features of the model do not change substantially for different payoff values provided the overall scale and

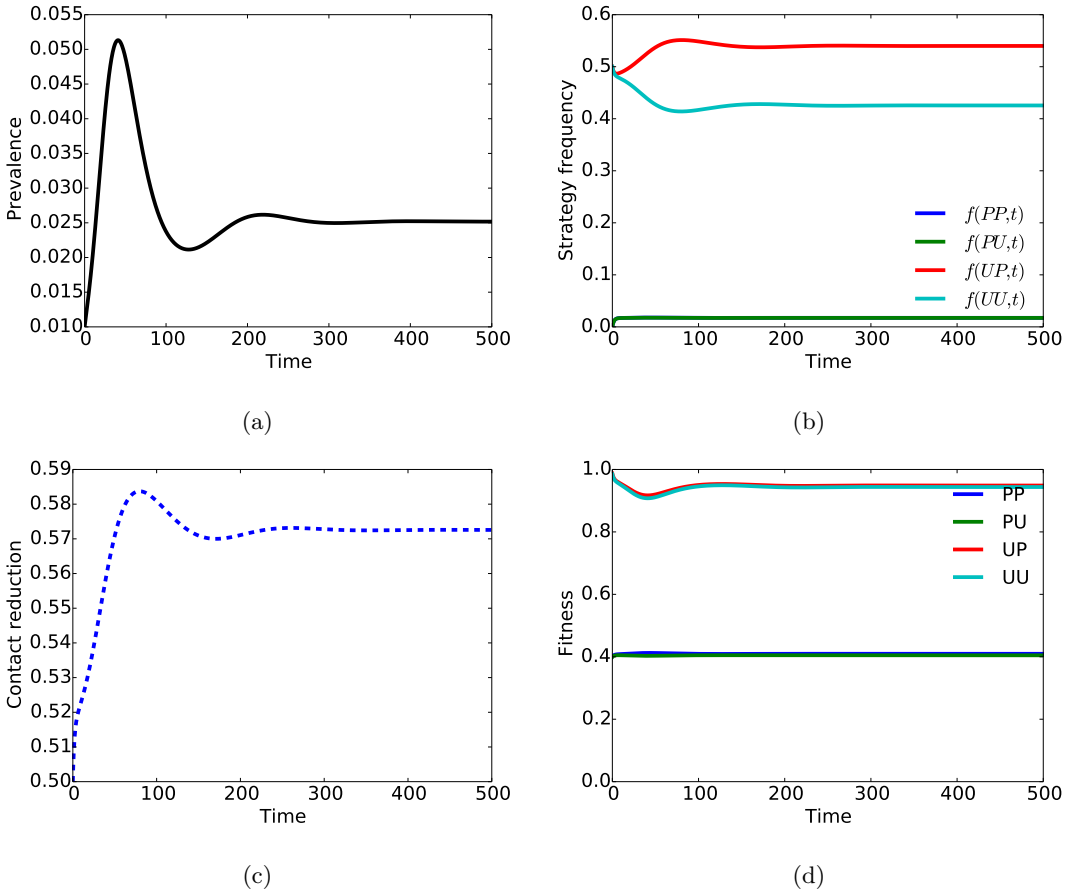


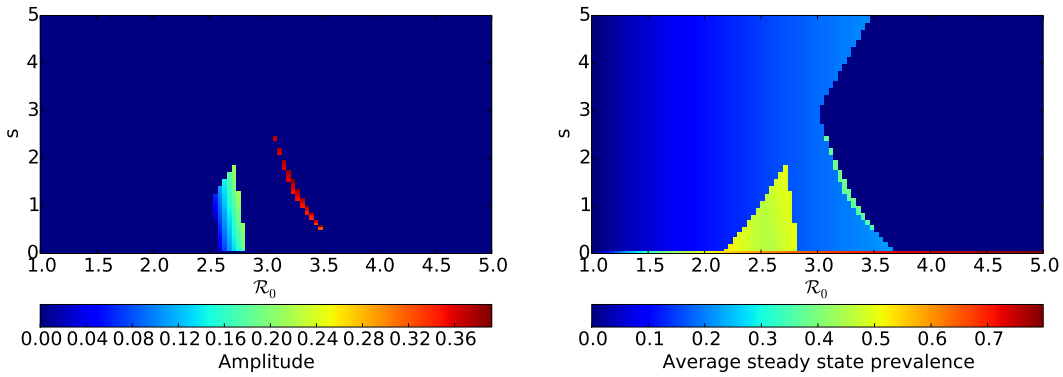
Figure 3.2: Simulated trajectories for the combined model with $R_0 = 1.2$. Disease parameters $\beta_b = 1.2, \gamma = 0.5$ with behavioral parameters $a = 1, b = 0.6, c = 0.4, d = 0.2$. (a) Disease prevalence, (b) Type-contingent strategies, (c) Fractional contact reduction, (d) Strategy fitness.

ordering are preserved. Unless otherwise specified, the disease initial conditions were $S(0) = 0.99$, $I(0) = 0.01$. In the first section below, we consider fixed initial behavioral conditions at $f(UU, 0) = f(UP, 0) = 0.5$, while in the following section we consider the effects of varying the initial behavioral conditions (for $f(UU, 0)$ between 0 and 1).

3.3.1 Model Dynamics

We began by inspecting an example set of model trajectories, shown in Fig. 3.2 (with $\beta_b = 1.2, \gamma = 0.5$, and $f(UU, 0) = 0.5$ to give $\mathcal{R}_0 = 1.2$), to illustrate the interactions between behavioral changes and disease dynamics. Comparing the prevalence and contact reduction trajectories demonstrated the impact of the adaptive behavioral dynamics on the progression of the simulated outbreak. In this example the outbreak was relatively small, so susceptible individuals did not have a large incentive to use protection. However, the initial increase in disease prevalence favored strategies where infected individuals used protection, which in turn decreased the force of infection enough to halt the initial outbreak. However, as prevalence decreased, unprotected strategies became less costly and contact reduction decreased again. Consequently there was a small secondary outbreak. The timing of the behavioral response also appeared to be influential. While contact reduction tracked prevalence over time, it did so at a delay. Thus the level of contact reduction was still relatively high at the onset of the second outbreak, preventing a large secondary peak. This preliminary exploration suggested that both the infectiousness of the disease and the speed of behavioral adaptation play important roles in the overall model dynamics.

To explore this issue further, we defined a timescale parameter s for the behavioral dynamics, given as a scaling factor on the payoff matrix A in Eq. (3.7). From a

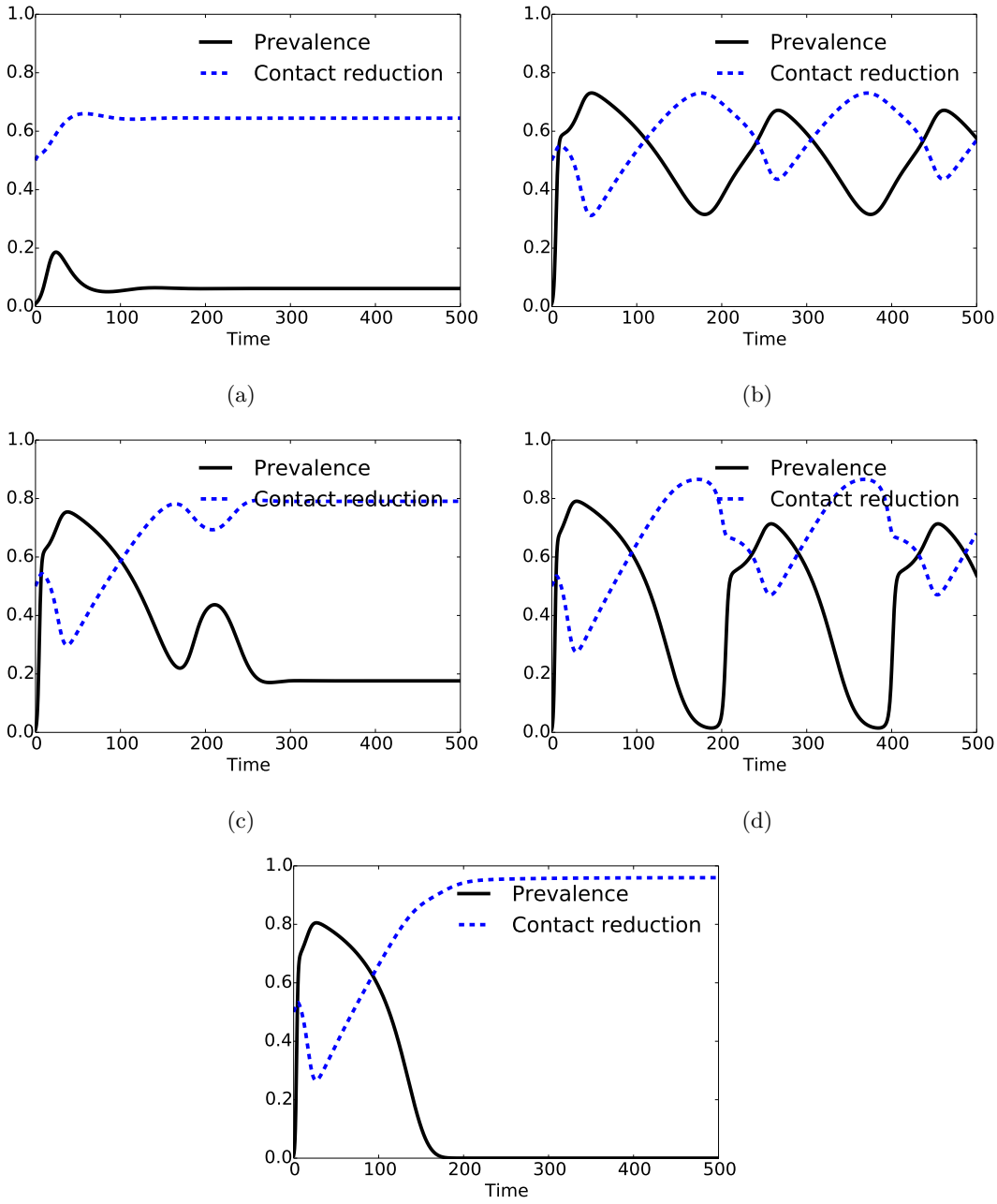


(a) (b)

Figure 3.3: Long term dynamics of the combined model with increasing values of $\mathcal{R}_0$ and the behavioral scale parameter s. $\mathcal{R}_0$ was adjusted by varying the parameter β_b in Eq. (3.10). (a) The amplitude of steady state prevalence oscillations. (b) The average prevalence at steady state.

behavioral perspective, this parameter can be thought of as controlling the speed of adaptation in the population. We then evaluated the long-term dynamics of the model in steady state for a range of values of $\mathcal{R}_0$ and s, shown in Fig. 3.3 (with all remaining parameters as in Fig. 3.2). In the triangular region about $\mathcal{R}_0 = 2.7$ and the canyon between $3 < \mathcal{R}_0 < 3.5$ the damped oscillations became sustained with a substantially higher average steady state prevalence. The triangular region of oscillations gives way to another endemic steady state region as $\mathcal{R}_0$ increases, followed by the thin band of oscillations and eventually disease extinction in the triangular region along the right edge.

Figure 3.4 and Supplementary Figure A.1 provide a more detailed examination of the dynamics in Fig. 3.3 as $\mathcal{R}_0$ increases for $s = 1$. For $\mathcal{R}_0$ sufficiently above 2, the early increase in contact reduction by infected individuals was not sufficient to stop the spread of the initial outbreak (3.4b). As a result prevalence increased, causing infecteds to begin to switch back to unprotected strategies (A.1 b). However, at higher prevalence susceptible individuals had an incentive to use protection, compensating for the behavior of the infected population. As in the initial example, this overall



(a)

(b)

(c)

(d)

(e)

Figure 3.4: Prevalence and contact reduction from the combined model with $\gamma = 0.5, \mathbf{f}(0) = (0.0, 0.0, 0.5, 0.5)$ for all simulations and (a) $\beta_b = \mathcal{R}_0 = 1.5$, (b) $\mathcal{R}_0 = 2.7$, (c) $\mathcal{R}_0 = 2.9$, (d) $\mathcal{R}_0 = 3.3$, and (e) $\mathcal{R}_0 = 3.5$.

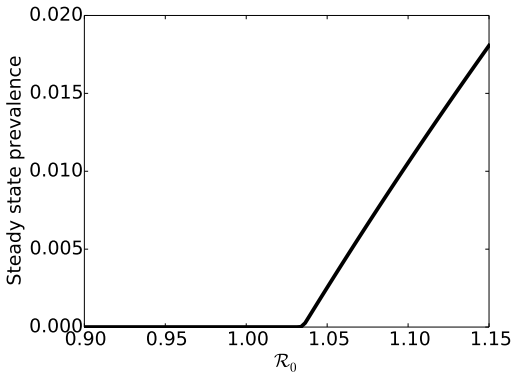
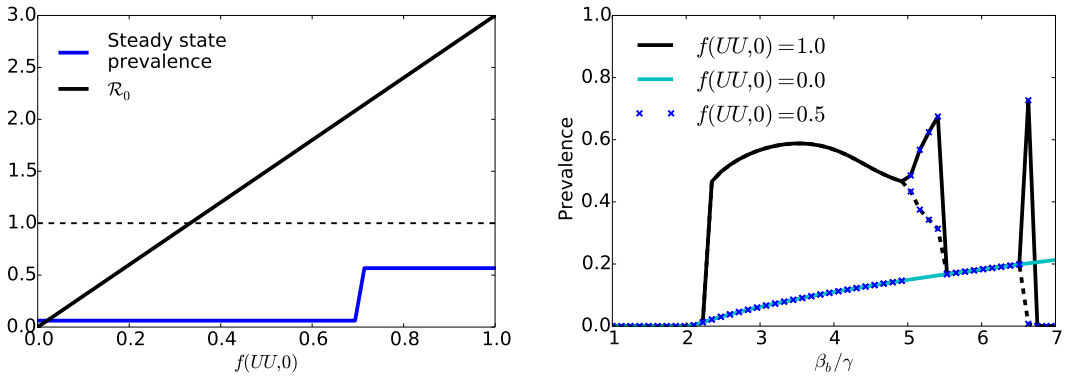


Figure 3.5: Average steady state prevalence of the combined model for increasing $\mathcal{R}_0$ with $\gamma = 0.5, \mathbf{f}(0) = (0.0, 0.0, 0.5, 0.5)$. The transition between the disease free and endemic equilibrium occurs at $1 < \mathcal{R}_0 < 1.05$.

increase in contact reduction reduced the force of infection and halted the growth of the outbreak. The magnitude of this response appeared to determine the long term dynamics of the model. When prevalence was reduced to a moderate level neither protective nor unprotected strategies were able to gain a stable foothold in the population, leading to sustained oscillations in both contact reduction and prevalence (3.4b, A.1 b). A stronger contact reduction response pushed the disease to a level where infected individuals again had a strong incentive to use protection. The fitness advantage of this strategy was sufficient to survive the second outbreak, leading to an endemic steady state (3.4c, A.1 c). Notably, when $\mathcal{R}_0$ was high enough that the resulting behavioral response nearly drove the disease to extinction, sustained oscillations were again possible (3.4d, A.1 d). Subsequent outbreaks in this scenario were contained by increasing contact reduction from a higher baseline due to the initial outbreak, but grew rapidly due to the infectiousness of the disease. Finally, when the initial outbreak was extremely large, essentially all remaining susceptibles (including those who recently recovered) played protective strategies, driving the disease to extinction (3.4e, A.1 e). Since the boundaries of both oscillatory regions shifted with both s and $\mathcal{R}_0$, it appeared that the timescales of infection and adap-



(a) (b)

Figure 3.6: The effect of changing the initial condition $f(UU,0)$ with $\gamma = 0.5$. (a) Steady state prevalence and $\mathcal{R}_0$ for $\beta_b/\gamma = 3$ (Dashed line indicates where $\mathcal{R}_0 = 1$). (b) Steady state prevalence of the combined model as β_b/γ increases for three choices of $f(UU,0)$. In regions of β_b/γ where sustained oscillations occurred, the solid black line shows the value of the crest while the dashed line shows the value of the trough.

tation must align in order to produce sustained oscillations (e.g. Figure 3.4b and A.1b). In particular, oscillatory dynamics appear only to be possible for relatively slow adaptation speeds, suggesting that the lag induced in the behavioral dynamics contributes to the potential for oscillations.

Using the same parameters as in Fig. 3.4, in Figure 3.5 we also examined steady state prevalence near $\mathcal{R}_0 = 1$, to evaluate whether behavior changes may affect the threshold for generating an outbreak. We found that steady state prevalence remained zero even when $\mathcal{R}_0$ was greater than 1, suggesting that behavioral dynamics were able to extinguish the disease even when the initial growth rate would have generated and epidemic.

3.3.2 Behavioral Initial Condition Analysis

For our next set of simulations, we tested the effect of varying the initial behavioral conditions. Figure 3.6a shows the average steady state prevalence across $f(UU,0) \in [0,1]$ when $\beta_b/\gamma = 3$ (where the remaining population was assumed to take strategy UP as noted above). As suggested by the form of $\mathcal{R}_0$, an endemic steady state

persisted even for $\mathcal{R}_0 < 1$. This counterintuitive phenomenon occurred because, while the disease initially declined, it did not immediately become extinct (Fig. A.2). The declining prevalence led to increased adoption of UU, driving the effective reproductive rate back above one. Figure 3.6b depicts the effect of different levels of $f(UU, 0)$ on the steady state behavior of the combined model for increasing β_b/γ. A wide range of dynamics were observed as β_b/γ varied, including a single stable equilibrium, bistability, and oscillations. When $f(UU, 0)$ was low (green line in Fig. 3.6b), the model only exhibited damped oscillations, and did not produce the same extinction behavior at high β_b/γ as in the case where $f(UU, 0)$ was substantially larger than zero. For $f(UU, 0) = 0.5$ (blue dotted line), the model exhibited the same bifurcation pattern as in the previous section. At high $f(UU, 0)$ (black line) the behavior was similar, however the endemic prevalence before the first oscillatory region was substantially greater. From this, $f(UU, 0)$ appeared to act as a switch between possible steady state regimes of long term dynamics.

For the constant steady state regions, the endemic prevalence took one of two values for a given β_b/γ depending on $f(UU, 0)$, similar to Fig. 3.6a. More broadly, for any fixed β_b/γ, there were two basins of attraction corresponding to either the $f(UU, 0) = 0$ or $f(UU, 0) = 1$ case, where the unstable equilibrium dividing the two regions depended on the value of β_b/γ. This is illustrated by the $f(UU, 0) = 0.5$ curve in Fig. 3.6b, which switches between existing in the $f(UU, 0) = 0$ and $f(UU, 0) = 1$ basins around $\beta_b/\gamma = 5$. There are also regions of Fig. 3.6b where only a single steady state exists regardless of the value of $f(UU, 0)$, at very low values of β_b/γ (left portion just above $\beta_b/\gamma = 2$) and larger values (between the oscillatory regions).

As noted in the previous section, there were also multiple regions where the long term dynamics showed disease extinction even though $\mathcal{R}_0 > 1$. For $f(UU, 0) = 0.5$

and 1, the left corner of Fig. 3.6b shows extinction with $\mathcal{R}_0 > 1$, and the right corner shows the same extinction discussed in the previous section for the $f(UU, 0) = 1$ basin at high values of β_b/γ.

3.3.3 Comparison with Alternate Models

Finally, we considered how the dynamics of the model compare to other potential models of the disease dynamics, to examine how neglecting the behavioral dynamics may alter model forecasts of the epidemic trajectory or affect estimates of key epidemiological parameters such as $\mathcal{R}_0$. As a first example, we chose a fixed contact rate SIS model parameterized such that the $\mathcal{R}_0$ of both models was identical. While the trajectories were similar at very early times (at which the epidemic growth rate can still be characterized by $\mathcal{R}_0$), the combined model quickly diverged due to the reduced effective contact rate, and equilibrated at a substantially lower endemic level (Figure 3.7a). Similarly, we computed $\mathcal{R}_0 = \frac{1}{1-I_\infty}$ naively from the steady state prevalence of the combined model in Figure 3.7a, without accounting for a time-varying contact rate. The estimated value of 1.03 was substantially lower than the true $\mathcal{R}_0$ of 1.2 for the combined model. This difference can be even more severe when considering endemic steady states at higher $\mathcal{R}_0$'s (such as the area to the right of the triangular region in Figure 3.3). For example, the endemic steady state in Figure 3.4c yielded an apparent $\mathcal{R}_0$ of 1.21, when the underlying $\mathcal{R}_0$ for the combined model was 2.9. Given the simplicity of the SIS model, these discrepancies were not surprising. Nonetheless, they highlight some pitfalls of neglecting the effects of behavioral dynamics.

However, as the SIS model is known not to be able to produce oscillations, it might not be a likely choice given data that came from real-world infection dynamics similar to the combined model. Due to the delay between recovering from infection

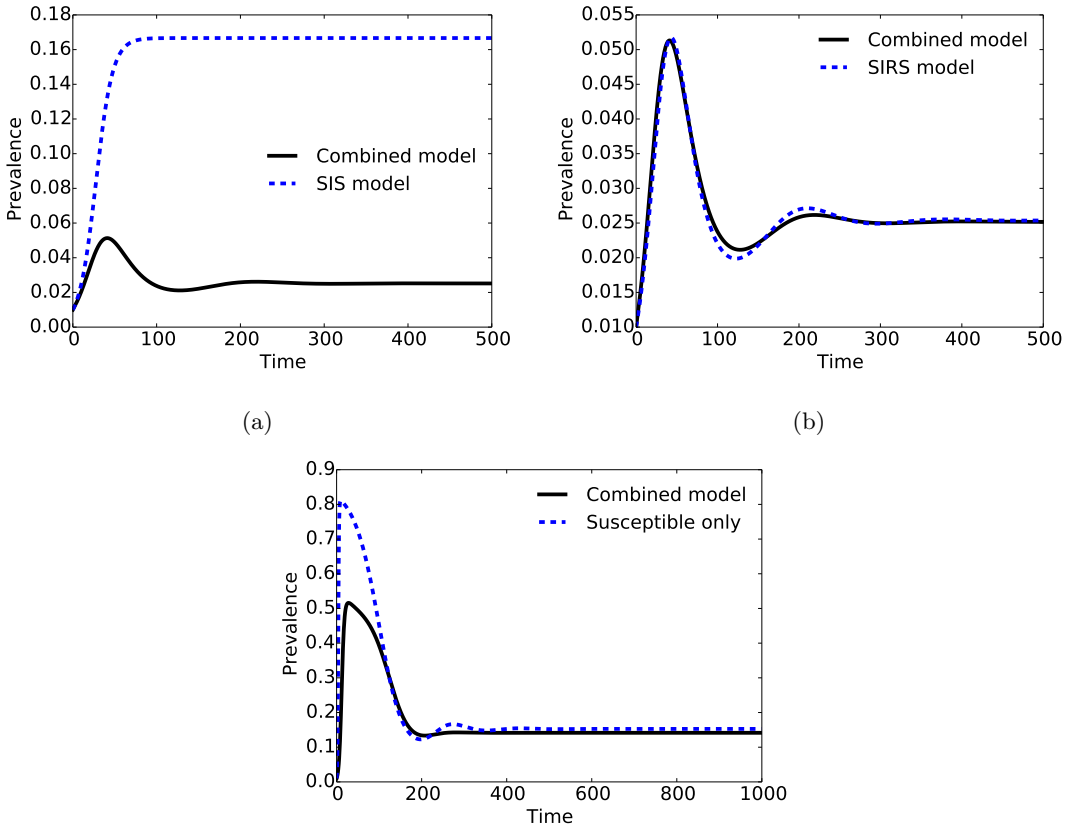


Figure 3.7: Comparison of simulations of the combined model other compartmental models. (a) SIS model contact rate $\beta = 0.6, \gamma = 0.5$ to match the $R_0 = 1.2$ of the combined model, (b) Prevalence trajectories for the combined model and an SIRS model fit to the combined model prevalence simulated with $\mathcal{R}_0 = 1.2$. SIRS best fit parameters: $\beta = 0.274, \gamma = 0.203, \delta = 0.022$. (c) A variant combined model in which only susceptible individuals adapt their behavior fit to the simulated trajectory from the full model with $\beta_b = 1.8, \gamma = 0.5$ and $\mathcal{R}_0 = 2.7$. The best fit parameter $\beta_b^* = 2.84$ for the reduced model, giving an $\mathcal{R}_0 = 5.68$.

and returning to susceptibility, the SIRS model is capable of producing damped oscillations, and so might make a reasonable initial guess for the model structure if the length of immunity was unknown. To evaluate how a more realistic but misspecified model might affect parameter estimation, we fit the SIRS model (A.4) to the prevalence trajectory of the combined model with $\mathcal{R}_0 = 1.2$, using least squares with Nelder-Mead optimization, as shown in Figure 3.7b. While the best fit SIRS model conformed well to the target trajectory, the resulting parameter estimates included both a four fold decrease in the baseline contact rate and a substantial increase in the average waiting time until a recovered individual becomes susceptible again. The $\mathcal{R}_0$ for the best fit SIRS model was 1.35, a 12.3% increase compared to the behavioral model. In this case, the feedback between behavior and transmission may make interventions more effective than would appear to be the case if we had assumed a fixed contact rate.

Finally, most previous work combining game theory with transmission models to model contact reduction has focused on the adaptive behavior of susceptibles [13, 14, 17, 43]. This is a natural formulation for vaccination, but may be less suitable for more general contact reduction behaviors. Thus, to test the effect of modeling adaptive behavior by both susceptible and infected individuals, we used a reduced variant of our combined model in which infected individuals always select the action U. As with the SIRS model, we fit this reduced model to a simulated trajectory from the full model. In this case, the reduced model overshoots the initial outbreak curve, but equilibrates to a similar endemic equilibrium to the full model (Figure 3.7c). However, the best fit parameter and $\mathcal{R}_0$ for the reduced model were nearly two times higher than the true values for the full model. For a wider range of parameter values of the full model, the reduced model generally failed to provide

qualitatively accurate fits. In addition, the reduced model did not produce the same range of dynamics as the full model. We give a full description of the reduced model in A.4 as well as an expanded set of fits and simulations.

3.4 Discussion

To explore the feedback between behavior and disease dynamics, we developed a model of disease transmission with adaptive behavior among both susceptible and infected individuals. Numerical simulations illustrated the effect of behavior-disease feedback on model dynamics and inferences about STI transmission (Figure 3.2). We found that phase transitions between damped and sustained oscillations occurred at an intermediate transmission rate and again at a higher transmission rate (Figure 3.3). Increasing the adaptation rate parameter s reduced the range of transmission rates at which sustained oscillations could occur.

Disease extinction occurred both for low and high transmission rates (Figures 3.3, 3.6). At lower transmission rates, adoption of protection by infected individuals lead to extinction above the typical $\mathcal{R}_0 = 1$ threshold (Figure 3.5). This suggests that behavior-disease feedback can create a herd-protection-like effect if the infected population can be reduced sufficiently to incentivize the use of protection among that group. At high transmission rates and low-to-mid range levels of initial risky behavior ($f(UU, 0)$) in the infected population, we observed the counterintuitive result that a large initial increase in risky behavior (UU and PU strategies) among infected individuals preceded disease extinction. However, because susceptible individuals were also switching to protective strategies (PP and PU) rapidly, the combined effect of both susceptibles and infected adopting PU resulted in sufficient growth of PU to drive the disease to extinction as the effective number of susceptible individuals

remained small and all were largely playing protective strategies (Figure 3.4, A.1). In fact, both the high $\mathcal{R}_0$ and low $\mathcal{R}_0$ extinction phenomena (for $\mathcal{R}_0 > 1$ but still low) appear to be driven by a reduction in the effective size of the 'bottleneck' population group. The low $\mathcal{R}_0$ extinction (Figure 3.5) seems largely due to increasing infected protective strategies, while the high R0 extinction seen on the right side of Figure 3.4 is largely due to increasing susceptible protective strategies. In each case, the low/high prevalence makes infecteds/susceptibles a bottleneck group, where sufficient protective response in that group appears to result in disease elimination. The feedback between disease and behavior thus often constrains endemic outcomes to scenarios where a disease is only moderately infectious, for a range of realistic levels of initial risky behavior (Figure 3.6b).

However, simulations predicted a more pessimistic outcome for changes in the initial strategic distribution. In general, the initial level at which infected individuals used protection did not determine whether the disease would become extinct. As long as the value of β_b/γ was sufficiently large, any $f(UU, 0)$ led to an outbreak and endemic disease (Figure 3.6). Instead, the initial distribution acted as a switch between two steady state regimes: one in which initial risky behavior by infecteds (high $f(UU, 0)$) lead to sustained oscillations and higher endemic prevalence could occur for values of c_b below the first oscillatory region, and another in which initial protective behavior by infecteds (low $f(UU, 0)$) yielded damped oscillations and generally lower endemic prevalence. The exception to this pattern is the region where β_b/γ is large. As noted above, in this region extinction occurs when early risky behavior by infected individuals is then countered by widespread use of protection by susceptible individuals. However, early widespread protection use by infecteds leads to a smaller outbreak that is slowed, but not eliminated by susceptible protection use. This com-

plicates the interpretation of $\mathcal{R}_0$ for the combined model, as the standard endemic threshold property often does not apply. For interventions it is thus important to understand the interaction between early behavioral patterns, the transmission rate, and treatment levels. Interventions based on these factors may yield counterintuitive results depending on the other factors, making comprehensive data collection critical when optimizing intervention strategies.

It is also interesting to note that while an analysis of the Bayes-Nash equilibria of the protected sex game at DFE indicates that the susceptible-unprotected strategies (U, U) are risk dominant, our combined model suggests that the actual initial distribution will depend on a population's previous experience with disease. That is, when the outbreak is large (i.e. for high $\mathcal{R}_0$, Figure 3.4), adoption of protective strategies results in disease elimination. However, the increased protective strategies by susceptibles mean that the model switches to the attraction basin of the non-dominant (P, P) Nash equilibrium as it goes to the new disease-free steady state A.1 (or as close as possible given the mutation rates in the replicator-mutator equation). By contrast, below the threshold for high *Ro* extinction, the model tends to reach a steady state close to the susceptible-unprotected (U, U) Nash equilibrium since endemic prevalence is low enough to favor susceptible-susceptible contact. This suggests that in general, the disease-free strategy balance between protected and unprotected depends on the magnitude of previous outbreaks, with larger outbreaks yielding disease-free protective behaviors even though they are neither payoff nor risk dominant. As a result, post-outbreak surveillance and control are crucial to prevent recurrent outbreaks due to reintroduction.

Based on comparisons between the combined model and similar disease models with simplified contact dynamics, predictions from fixed-contact rate models may

omit important dynamical features or yield misleading parameter estimates (Figure 3.7). Similarly, if an STI model is misspecified, the common method of computing $\mathcal{R}_0$ from endemic prevalence when a disease is assumed to be at equilibrium [74] is not appropriate and yields inaccurate estimates. One advantage of using a game theoretic framework to model interactive behavior is the flexibility of ordinal utility. Since protective measures can be used by susceptible or infected individuals to prevent infection or transmission, respectively, we used a preference order that gave individuals in both disease states an incentive to use protection if paired with a partner of the opposite state. Our simulation results suggest that disease dynamics differ in this scenario as opposed to when only susceptible individuals adapt their behavior, and indeed model fits using the reduced model considering only susceptible behavior change often yielded both incorrect $\mathcal{R}_0$ estimates and dynamic trajectories (Figure A.4). The role of infected behavior dynamics was particularly notable in the early stage of outbreaks when disease was not prevalent enough to induce susceptible individuals to use protection. In addition, the shift between infected and susceptible adoption of protective strategies appeared to drive sustained oscillations. To more thoroughly consider the effects of infected behavior dynamics, we also considered an alternative preference structure in which infected individuals no longer explicitly prefer protected sex with susceptible individuals, but rather have the same preferences regardless of their partner's disease status (A.5). This model only exhibits one oscillatory region and a higher endemic prevalence than the original model (Fig. A.5). However, extinction still occurs at high $\mathcal{R}_0$, confirming that susceptible protective strategies drive the high-$\mathcal{R}_0$ extinction phenomenon.

In order to focus on the effect of adaptive protective behavior, our combined model simplifies many other realistic factors that contribute to STI transmission.

The protected sex game does not include explicit negotiation, which likely biases the effective contact rate downward, since negotiation or coercion could lead to unprotected sex even when the initial action pair is (U, P). The combined model presented here still uses mass action assumptions to determine interactions between individuals. Although more complex contact patterns are known to influence transmission dynamics, we model a homogeneous population to focus on the effect of time-varying behavior. While this may not be completely implausible in a highly active group such as MSM frequenting bathhouses, it is almost certainly a poor representation of the manner in which individuals form sexual partnerships. Typically, adding contact heterogeneity increases the $\mathcal{R}_0$ of a model, so we might expect more rapid early outbreak growth for a wider range of parameters in our model, potentially expanding the regions in which sustained oscillations or even disease extinction occur. Our general framework, however, is amenable to extensions to expand the state space of the disease model to represent more complex natural histories or population structures as with standard transmission models. In addition, it is possible to model the combined process on a contact network or using a stochastic framework. One particular extension that could yield insight into spatial patterns and the effect of local information would be a simulation of the model on a regular lattice, where individuals could use either a global or local prior to estimate the probability of their partner's type.

A particularly compelling consequence of developments in economic-epidemiological models is the potential to estimate more complex behavioral parameters using traditional surveillance data. In the context of our combined model, this would result in estimates for the relative payoff values (a, b, c, d), which capture useful information about the perceptions of populations facing STI outbreaks. While subject to

as yet unknown identifiability properties, this manner of parameter estimation could provide a valuable link between game theoretic methods and the extensive empirical literature on the epidemiology of STIs.

CHAPTER IV

Modeling behavioral responses to infectious disease outbreaks: Burial practices during the 2014 Ebola outbreak

4.1 Introduction

Ebola virus disease (EVD) is a serious and frequently lethal infection often identified with the profuse internal and external bleeding observed in late stage patients [75, 76]. First identified in 1976, EVD has historically been restricted to small, self-limiting outbreaks [75, 76]. However, in 2014 West Africa experienced the largest recorded epidemic with over 25,000 total cases and 15,000 total deaths primarily distributed among Guinea, Liberia, and Sierra Leone. Unlike prior occurrences of EVD, the 2014 outbreak reached urban centers, amplifying its transmission potential. In addition, regional infrastructure was ill equipped to contain EVD, while widespread poverty and political instability further complicated control measures [77, 78].

The magnitude of the 2014 outbreak sparked a substantial amount of interest among mathematical epidemiologists, and a variety of models were developed to test control measures and forecast the trajectory in West Africa [18, 19, 21, 79]. However, early projections were criticized for frequently overestimated the final size of the outbreak, leading to a debate regarding the uses and effectiveness of mathematical models in outbreak situations [22, 80]. Much of the initial forecasting error appears to be the result of omitting an adjustment for the reporting rate and population at risk

in the output equations. Including such a correction factor improves the accuracy of short and medium-term forecasting [24]. Nevertheless, the reporting rate/population at risk adjustment does not mechanistically capture behavior changes that may be important for long-term prediction [23, 81] and intervention planning. In particular, we concentrate on funerals as a transmission route and potential source of behavioral dynamics.

Due to the pathophysiology of EVD and its high mortality rate, regional burial practices fell under scrutiny as a major source of transmission. Recent studies have indicated that Ebola virus can remain viable up to seven days postmortem [82], with high concentrations of pathogen remaining in body fluids. Traditionally, burials in West Africa involve a substantial amount of direct contact between the family of the deceased and the cadaver itself. Specifically, these practices include touching, washing, and kissing the corpse, resulting in a high probability of infection. Indeed, a single funeral in Guinea resulted in 85 new cases of EVD [83]. The deeply traditional nature of these burial practices posed a challenge to public health professionals attempting to reduce funeral transmission. Sanitary procedures such as sterilization, bagging, and disposal of corpses were viewed as an affront to the deceased and their family [84], leading to "safe and dignified" burial initiatives [85]. Of the three primary countries, Liberia instituted mandatory cremations, leading some burials to be conducted secretly. In spite of these challenges, sanitary burials have been cited as a key factor in containing the outbreak [86, 87].

We would expect the shift in burial practices from traditional to safe and dignified methods to have a measurable impact on the course of the outbreak. To capture the effect of these changes in behavior, it is necessary to address the feedback between EVD transmission and burial practices. The primary incentive to adopt sanitary

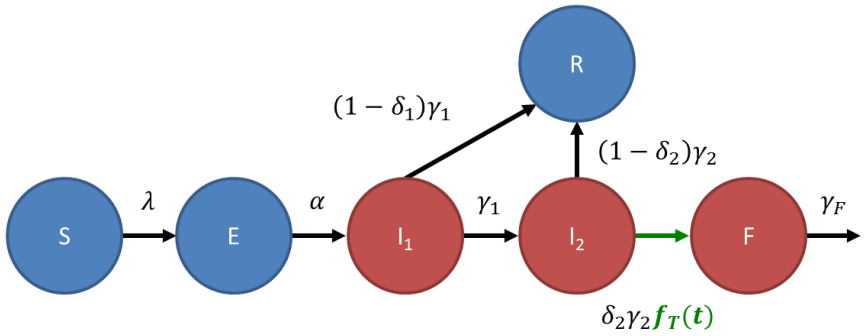


Figure 4.1: Compartmental flow diagram of the Ebola transmission model. Burial practices influence the transition between the second infected compartment I_2 and the funeral compartment F (highlighted).

burial techniques is likely related to the degree of EVD transmission, in this case the contribution to morbidity and mortality from traditional burials. Thus as an EVD outbreak grows, sanitary burials may become more common, reducing the force of infection. We explicitly model this feedback relationship using a population level game theoretic model of burial practices coupled to a compartmental model of Ebola transmission. This approach is similar to the influenza and sexually transmitted infection models of Bauch and others [10, 13, 14, 43]. Discrete individual simulations could also be used to address changing behavior, however such models are computationally expensive and can be difficult to analyze. Our deterministic model is relatively simple to implement and enables us to make inferences about behavioral processes from traditional surveillance data. In addition, the behavior-disease framework is straightforward to extend to multiple or alternative behavioral mechanisms.

4.2 Methods

Our model and analyses are implemented in Python 2.7 using Numpy, Scipy, and Matplotlib.

4.2.1 Model

Burial practices By definition, individuals cannot make a choice about the manner of their own burial at the time it occurs. However, it is possible to specify what type of burial would be preferable for family members and others in the community. In aggregate, we assume that these preferences determine the distribution of burial practices. Individual choices exist in a feedback cycle with social or cultural norms about burial. Individuals are influenced by the dominant burial practice, and continued adherence reinforces the norm. We model this feedback using evolutionary game theory. Like traditional game theory, we specify the set of choices individuals can make and the abstract payoff associated with each potential choice. However, evolutionary game theory is concerned with the dynamics of behavior in large populations of interacting individuals, which is appropriate to our analysis of changes in behavioral practices over the course of an outbreak.

As noted above, individuals choose whether they prefer a traditional or a sanitary burial. Each type of burial incurs a cost that reflects perceived advantages of the other type of burial. Individuals who prefer sanitary burials may face decreased social acceptance for acting contrary to the dominant cultural practice. However, if the contribution of burials to transmission is recognized, traditional burials may begin to be perceived as dangerous. These factors inform our payoff functions.

The payoff u_T for traditional funerals is

$$u_T(I) = -\rho I, \tag{4.1}$$

where I is the total prevalence and ρ is a constant of proportionality. This reflects perceptions regarding the risk of infection from traditional funerals which is assumed to be increasing in prevalence. Sanitary burials incur a cost C that reflects cultural

pressure, so their payoff u_S is simply

$$u_S = -C. \tag{4.2}$$

We assume that this cost does not change significantly on the timescale of an Ebola outbreak as cultural norms often change at a generational pace in the absence of major social upheaval.

The fraction of the population that prefers traditional burials at a given time is $f_T(t)$. We assume that behavior change occurs due to a social imitation process where individuals compare the payoff from their own choice to the payoffs received by others, switching if the alternative appears to be sufficiently better. This process can be represented by imitation dynamics [53]. For traditional burials:

$$\dot{f_T} = spf_S f_T \cdot (C - \rho I) \tag{4.3}$$

Individuals sample others at a rate s. When an individual encounters another with a different strategy, they may adopt the other strategy with probability proportional to the difference between payoffs $p \cdot \Delta E = p \cdot (C - \rho I)$ where p is the constant of proportionality for imitation (for $\dot{f_S}, p \cdot \Delta E = p \cdot (\rho I - C)$). We allow individuals to change strategies randomly with a small probability μ. This approach is analogous to replicator-mutator dynamics in evolutionary biology. The complete behavioral dynamics are

$$\dot{f_T} = s[pf_S f_T \cdot (C - \rho I) + \mu \cdot (f_S - f_T)]. \tag{4.4}$$

The inner term can be interpreted as the possible outcomes when an individual considers whether or not to change their burial preference. Either the individual encounters another who prefers sanitary burials and (possibly) switches their preference, or they may switch independently with a small probability. Unfortunately, s, p,and ρ

Parameter	Definition	Units	Source
β_1	First stage transmission rate	person-days^{-1}	Estimated
β_R	Second vs. first stage infectiveness ratio	unitless	Sampled [18, 19, 88, 89]
α	Average incubation period^{-1}	days^{-1}	Sampled [75, 76]
γ_1	First stage duration^{-1}	days^{-1}	Sampled [75, 90]
γ_2	Second stage duration^{-1}	days^{-1}	Sampled [75, 90]
γ_F	Burial rate	days^{-1}	Sampled [18, 75]
δ	Overall community mortality	unitless	Estimated
δ_2	Second stage mortality	unitless	Sampled [90]
k	Population at risk	unitless	Estimated
N	Total population size	unitless	Fixed [91]
c	Private/social cost of sanitary burials	unitless	Estimated
σ	Imitation sampling/adoption rate	days^{-1}	Sampled
m	Random choice rate	days^{-1}	Sampled

Table 4.1: Definitions of parameters used in the Ebola transmission model.

cannot be estimated individually. Instead we rescale the above using $\sigma = sp\rho$, $c = \frac{C}{\rho}$, and $m = \frac{\mu}{p\rho}$ giving

$$\dot{f_T} = \sigma[f_S f_T \cdot (c - I) + m \cdot (f_S - f_T)]. \tag{4.5}$$

Ebola virus transmission Clinically, EVD displays a multi-stage presentation with increasing severity and lethality over time [75, 76, 90]. The initial stage includes non-specific febrile symptoms, while later stages progressively include diarrhea, vomiting, hemorrhage, and organ failure [75, 76, 90]. Ebola virus remains viable in host fluids up to one week postmortem, enabling transmission by contact between uninfected individuals and infected cadavers [82]. We adapt the transmission model developed by Eisenberg et al. [24] by introducing the behavioral dynamics specified above. Figure 4.1 depicts the compartmental structure of our model. The full set of differential

equations are

$$\begin{aligned}
\dot{S} &= -(\beta_1 I_1 + \beta_2 I_2 + \beta_F F)S \\
\dot{E} &= (\beta_1 I_1 + \beta_2 I_2 + \beta_F F)S - \alpha E \\
\dot{I}_1 &= \alpha E - \gamma_1 I_1 \\
\dot{I}_2 &= \delta_1 \gamma_1 I_1 - \gamma_2 I_2 \\
\dot{F} &= f_T \delta_2 \gamma_2 I_2 - \gamma_F F \\
\dot{R} &= (1-\delta_1)\gamma_1 I_1 + (1-\delta_2)\gamma_2 I_2 \\
\dot{f}_T &= \sigma[(1-f_T) f_T \cdot (c - I) + m \cdot (1 - 2 f_T)]
\end{aligned} \tag{4.6}$$

S is the fraction of susceptible individuals, E is the fraction exposed but not yet symptomatic, I_1 and I_2 are individuals in the first and second stage of infection, respectively, F is the fraction who have died of EVD but have not yet been buried, and R is the fraction who have recovered and are assumed to be immune. Table 4.1 describes the parameters used in the above equations. Note that β_2, β_F, and δ_1 are derived parameters computed as follows: $\beta_2 = \beta_R \beta_1$, $\beta_F = \beta_R \beta_1$, and $\delta_1 = \delta/\delta_2$. Like Eisenberg et al. [24] we estimate δ instead of δ_1 or δ_2 as the overall community mortality rate is more likely to be available based on the case fatality rate than stage-specific mortality. We assume that all individuals practiced traditional burials before the outbreak, so $f_T(0) = 1$. For our model fitting and simulations, we set $I_2(0) = 1/kN, S(0) = 1 - 1/kN$. For Guinea, the index case was detected on December 26, 2013 [92]. For Liberia and Sierra Leone, we used the time of the first detected case.

The basic reproduction number for this model is

$$\mathcal{R}_0 = \frac{\beta_1}{\gamma_1} + \frac{\beta_2 \delta_1}{\gamma_2} + \frac{f_T(0)\beta_F \delta_1 \delta_2}{\gamma_F}. \tag{4.7}$$

This expression contains one term for each transmissible stage of EVD. However,

the burial transmission term is attenuated by the probability of a traditional burial at the disease free equilibrium. As we have noted in Chapter III the dependence of $\mathcal{R}_0$ on the behavioral initial condition can complicate its interpretation.

4.2.2 Parameter estimation and sensitivity analysis

We fit the model to cumulative incidence and mortality from the WHO situation reports (sitreps) [93] using a hybrid approach. While cumulative data can introduce estimation errors (primarily in variance estimates for deterministic models) [94], cumulative incidence and mortality were the only available data until relatively late in the outbreak. Deriving incidence (and mortality) from the reported cumulative data would not be viable, as reporting errors resulted in apparent decreases in cumulative incidence and mortality. Due to the large number of parameters in the model relative to the amount of available surveillance data, we fit a subset of parameters by numerical optimization, selecting others from plausible ranges derived from prior literature using Latin hypercube sampling (Appendix B.1). We use the following measurement equations [24] to link the model output to the cumulative sitrep data.

$$\begin{aligned} y_C &= kN \int_0^t \alpha E d\tau \\ y_D &= kN \int_0^t \delta_2 \gamma_2 I_2 d\tau \end{aligned} \tag{4.8}$$

Based on these equations and case/death data x_C/x_D respectively, our least-squares objective function is

$$Y = \sum_t (y_C(t,\theta) - x_C(t))^2 + (y_D(t,\theta) - x_D(t))^2 \tag{4.9}$$

We use local sensitivity analysis to characterize the degree of uncertainty in our parameter estimates. In particular, we take the partial derivatives of our objective function with respect to each estimated parameter [95, 96]. Formally, the relative

sensitivity coefficient for a parameter θ_i is

$$\phi_i = |(\frac{dY}{d\theta_i})(\frac{\theta_i}{Y})|_{\theta_i^*}. \tag{4.10}$$

Where the second term normalizes the sensitivity coefficient for comparisons between parameters and trajectories. θ_i^* indicates that the partial derivative is evaluated about the best fit value of a given parameter.

Given that only case and death data is available, it is possible that behavioral dynamics are not necessary to explain the epidemic trajectory. In order to determine whether this is the case, we also fit a variant of the transmission model without behavior change and compare the residual error and Akaike Information Criterion (AIC) between models.

4.3 Results

Parameter estimation Table 4.2 lists the best-fit values of β_1, δ_1, k, and c for each country as well as the LHS values for all other parameters corresponding to the best sample. The fitted transmission and mortality parameters (β_1 and δ) reflect the incidence and mortality trends in each country depicted in Figures 4.2a–4.2c. In particular, Guinea has both the highest ratio of total deaths to total cases as well as the highest estimated mortality rate δ, followed by Liberia and Sierra Leone. The behavioral parameters c, σ, and m were similar between the three countries. This could reflect the fact that traditional burial practices are also similar within the region, so we would not expect the social/cultural pressure to hold a traditional burial to vary substantially. In addition, the sampling rate σ may reflect factors such as urbanization or interpersonal connectivity that influence how frequently any given individual would be exposed to information about burial practices. The remaining best-fit sampled parameters are also relatively stable between countries with the

Parameter	Guinea	Liberia	Sierra Leone
β_1	*0.128*	*0.106*	*0.171*
β_R	2.71	4.66	3.78
α	0116	0.114	0.113
γ_1	0.158	0.363	0.189
γ_2	0.683	0.674	0.594
γ_F	0.952	0.319	0.651
δ	*0.667*	*0.454*	*0.296*
δ_2	0.825	0.820	0.887
k	8.85×10^{-3}	8.68×10^{-3}	5.25×10^{-3}
c	5.49×10^{-5}	3.92×10^{-4}	1.25×10^{-4}
σ	9.77	9.54	8.93
m	$10^{-4.08}$	$10^{-4.16}$	$10^{-4.03}$
$\mathcal{R}_0$	1.46	1.40	1.56
AIC	19775.11	224872.35	206568.72
Reduced AIC	44535.18	607758.09	477497.34

Table 4.2: Best-fit parameter values for the full Ebola transmission model. Underlined parameters were estimated by numerical optimization while all others were determined by LHS and correspond to the best sample. Model fit was computed using AIC (lower indicates better fit). The reduced AIC corresponds to a model without behavior change.

exception of the burial rate in Liberia. In general, this suggests that our selected transmission and behavioral parameters do inform differences between the outbreaks in each country. Our estimates of $\mathcal{R}_0$ range from 1.4 to 1.56. These values are relatively consistent between countries and are similar to other published estimates for Ebola [21, 24, 97]. The best-fit trajectories for each country generally match the corresponding cumulative case and mortality data from the WHO sitreps (Figure 4.2). Our model is least accurate for Liberia due to the period of apparent linear growth in the sitrep data from December, 2014 to June, 2015. It is not clear whether this phenomenon reflects actual transmission dynamics or is an artifact of more complete case detection catch-up in the later stages of the outbreak. However, our model is still able to capture the majority of the outbreak growth dynamics as well as the final size.

Sensitivity analysis indicates that the parameters we selected to fit directly do influence the quality of the model fit (Figures 4.2g-4.2i). In addition, the behavioral

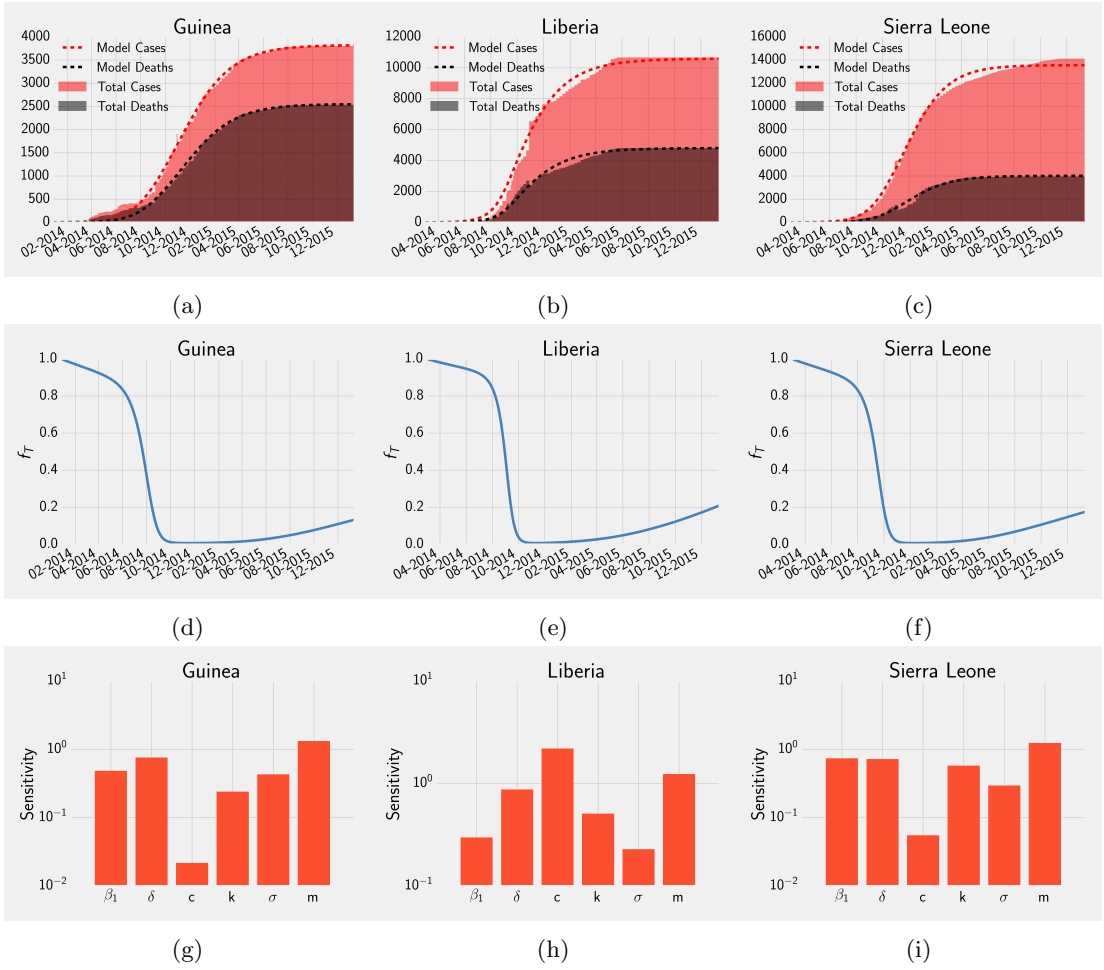
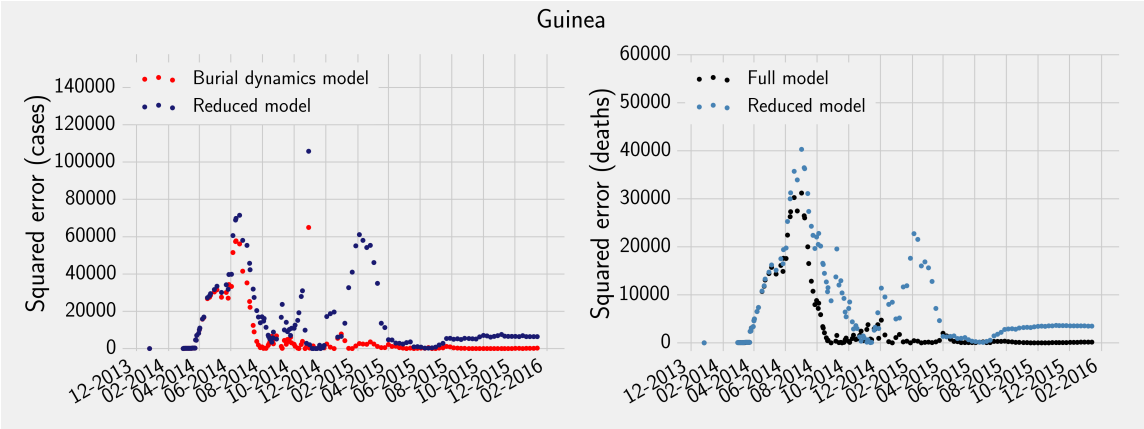


Figure 4.2: Burial dynamics model fit to cumulative cases and deaths in Guinea, Liberia, and Sierra Leone. The top three plots compare the simulated cumulative cases and deaths to the WHO sitrep data. The middle three plots show the predicted relative frequency of traditional burials from the best-fit models for each country. The bottom three panels show the sensitivity coefficients for our directly estimated parameters (β_1, δ, c, and k) as well as the remaining behavioral parameters (σ and m).

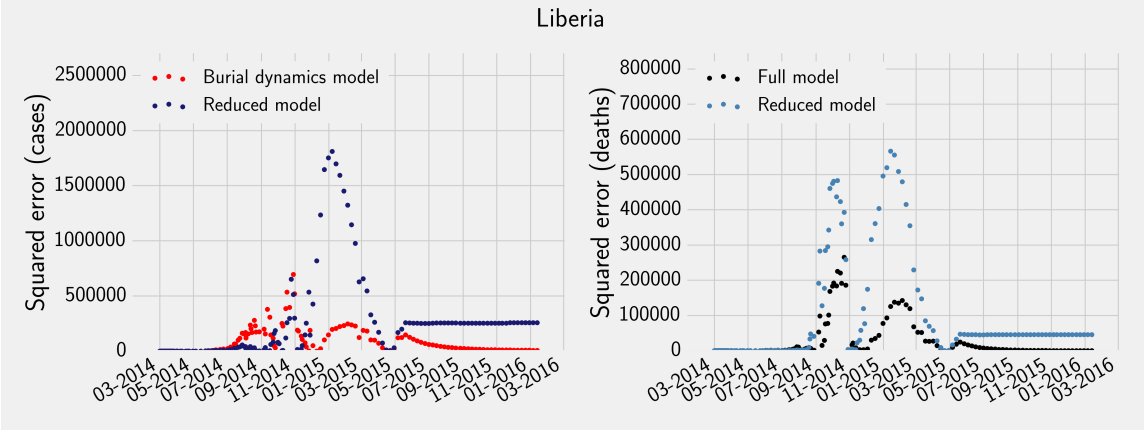
parameters σ and m appear to have individual influences on the model output. The scaled cost of traditional burials c has the lowest sensitivity coefficient when estimated for Guinea and Sierra Leone. Indeed, the range of plausible behavioral trajectories for the outbreak suggests that c can be estimated or sampled within approximately a factor of ten (Appendix B.2). Based on simulations using a range of c close to the best fit, it appears that this parameter influences the general shape of the behavioral trajectory (Appendix B.2). Thus, for practical purposes, it may be sufficient to establish which range yields dynamics that broadly match the observed outbreak.

Figures 4.2d–4.2f depict the predicted frequency of traditional burials over time from our best fit models. For each country traditional burials decline rapidly between August and October, 2014. Subsequently, our models predict that some individuals begin to revert to traditional burials once the outbreak has essentially ended. The timing of the first behavioral shift is noteworthy. While traditional burials decline somewhat during the growth phase of the outbreak, the largest change occurs after the peak simulated incidence. In addition, this period corresponds to a phenomenon observed in Eisenberg et al. [24] regarding estimates of the reporting rate/population at risk parameter k: The best fit value varied depending on the amount of data used to fit the model. In particular, k was relatively stable using data up to September, 2014, but decreased by orders of magnitude once data from October and later was included. Thus, it appears that transmission-related behavior change can explain some of the variation of the reporting rate/population at risk over time.

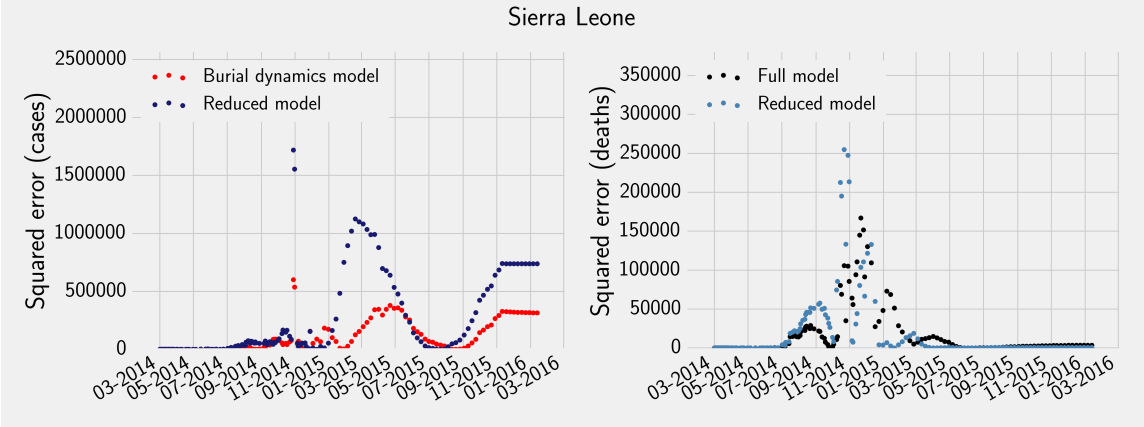
Model comparison and forecasting The reduced model without behavior change results in a worse fit both quantitatively and qualitatively. Even after adjusting for



(a)

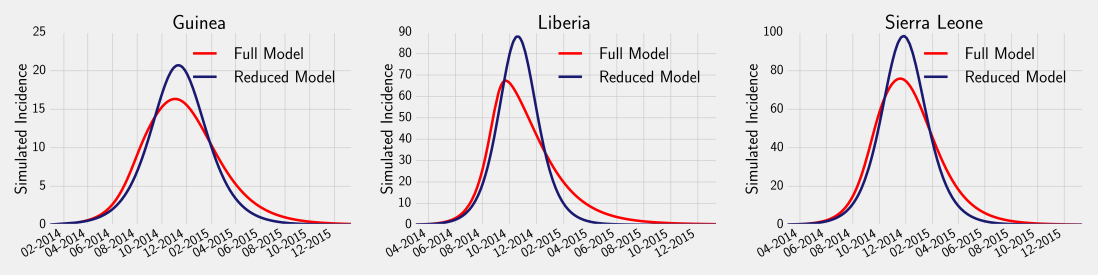


(b)



(c)

Figure 4.3: Residual error for the full and reduced models with respect to time. The left and right panels show the error contribution of cumulative cases and cumulative deaths respectively.

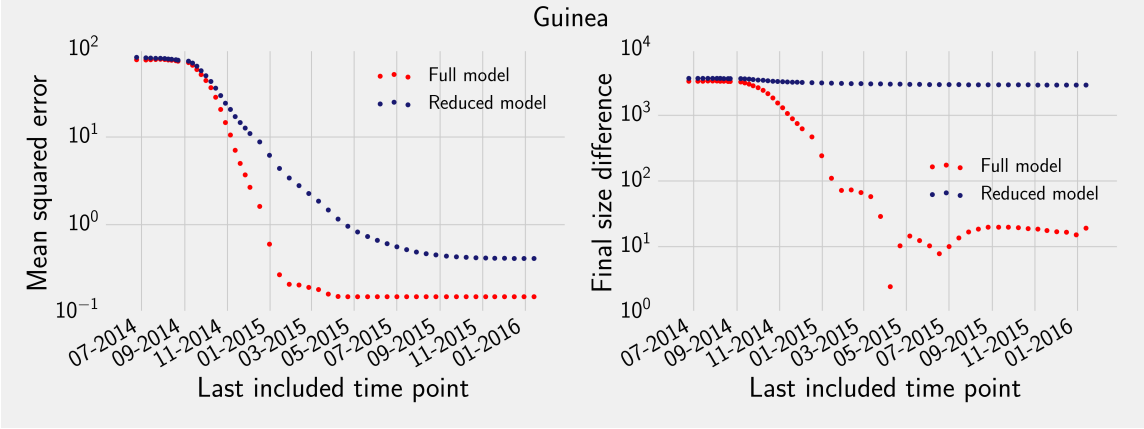


(a) (b) (c)

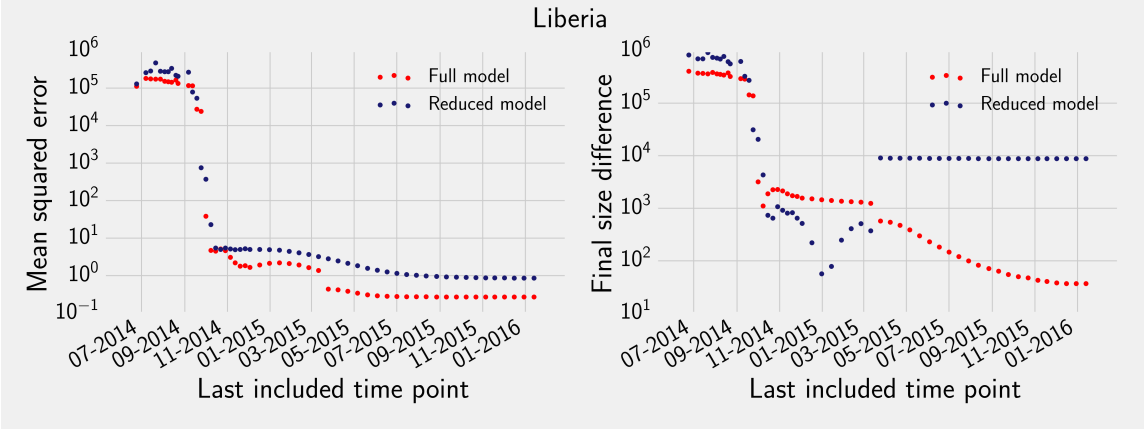
Figure 4.4: Simulated incidence from the best-fit full and reduced models for 4.4a Guinea, 4.4b Liberia, and 4.4c Sierra Leone.

an increase in the number of parameters, the full model AICs are substantially lower than those of the reduced model (Table 4.2). Figure 4.3 shows the squared residual error from both models for each time point in the sitrep data for each country. Both models perform similarly early in the outbreak, but the reduced model cannot capture features of the data after incidence peaks. This can also be seen by comparing the case and death trajectories from the reduced model to the sitrep data (Appendix B.2). As a result, the best fit model without behavior change generally mis-predicts the final size of the outbreak. This discrepancy is likely because the reduced model does not have a mechanism that can adjust transmission rates or the population at risk. Figure 4.4 compares the simulated incidence from the full and reduced models. The full model generally produces a longer tailed incidence curve but lower peak incidence. This is a consequence of the behavioral dynamics described above. When traditional funerals decrease, the force of infection from funerals also decreases. As a result, the susceptible population is depleted more slowly, allowing the outbreak to continue for a longer period of time.

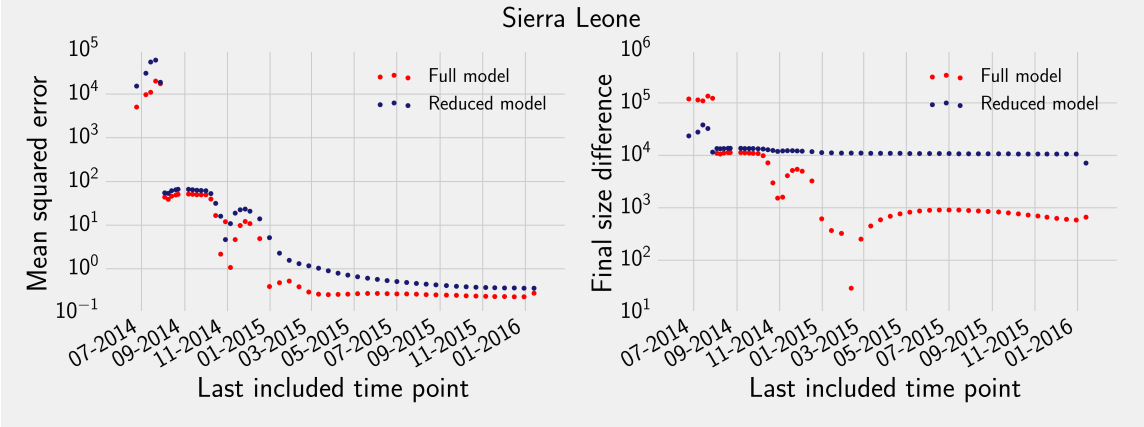
We also evaluate the performance of our model by testing its forecasting accuracy relative to the reduced model. To do this, we fit both models to truncated data from the outbreak then compute the mean squared error (MSE) for our output equations



(a)



(b)

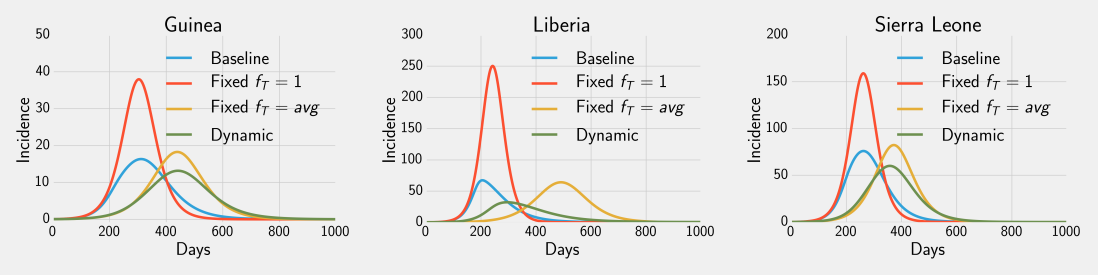


(c)

Figure 4.5: Comparison of forecasting accuracy between the full and reduced models. Left panels show the mean squared error for forecasts from the full and reduced model using data up to each included end point. Right panels compare the forecast final outbreak size from each model to the surveillance data final size.

using the full outbreak data. Thus we can test how effectively the model would have forecast the outbreak at different points in time. Figure 4.5 displays the full model's forecasting performance compared to the reduced model. We are concerned with both how accurately the models predict the full time-course of the outbreak as well as whether they predict its final size. While the forecasts are highly sensitive to the last data point included for fitting, the full model generally yields a lower MSE as well as a smaller difference between the predicted and actual final size (often by a full order of magnitude). We do note however that neither model performs well until data from October or November, 2014 is included. This corresponds to the inflection point in the outbreak data where incidence no longer increases exponentially. The quality of projections from the full model do improve significantly after this point, further suggesting that the data contains a signal of behavior change.

Alternate scenarios We compare three hypothetical scenarios to evaluate the impact of changing burial practices as a control measure. In all three scenarios, we parameterize our model using the best fit values for each country and simulate outbreaks assuming a single initial infected individual. The first scenario represents a worst case with respect to burial practices in which all funerals are traditional ($f_T = 1$) with no change over time. The second scenario considers another fixed-behavior case – the fraction of traditional burials is set to the average over the trajectory from the best-fit model. This condition can be interpreted as an approximation of the actual behavioral dynamics assuming data collection does not capture the full trajectory. For the final scenario, we set the initial fraction of traditional burials to its eventual steady state value from our fit to the 2014 outbreak and allow burial practices to change. This represents a population that has previously



(a) (b) (c)

Figure 4.6: Simulated incidence under behavioral scenarios for each country. Baseline curves represent incidence from the best-fit models shown in Figure 4.2. Total cases (by scenario) for Guinea (4.6a) Baseline: 3825, $f_T = 1$: 5836, $f_T = avg$: 4281, dynamic: 3647. Total cases for Liberia (4.6b) Baseline: 10628, $f_T = 1$: 26920, $f_T = avg$: 15459, dynamic: 8461. Total cases for Sierra Leone (4.6c) Baseline: 13567, $f_T = 1$: 19874, $f_T = avg$: 15148, dynamic: 12907

experienced a large Ebola outbreak and adapted its burial practices accordingly, but practices continue to evolve. Our model predicts that traditional burial practices do resume once the initial outbreak is over, so the new initial f_T is above zero. Figure 4.6 shows the incidence trajectories for each scenario as well as the best-fit (baseline) model. We observe a similar phenomenon as in our model comparison. Except for Liberia, both fixed-behavior scenarios have higher peak incidence than baseline, and generally symmetrical epidemic curves. All other trajectories and final outbreak sizes were higher than the dynamic scenario, indicating that prior experience with an outbreak can reduce the overall magnitude of subsequent outbreaks. Adaptation alone is not sufficient to prevent an outbreak, however, suggesting the need for additional intervention mechanisms.

4.4 Discussion

Behavior change during infectious disease outbreaks is frequently hypothesized or anecdotally reported, but rarely included explicitly in mathematical models. We used evolutionary game theory to address the population dynamics of burial practices during the 2014 Ebola outbreak in West Africa. In particular, our approach allowed

us to assess the impact of adaptive behavior change on the scope of the outbreak. Our model produced substantially more accurate fits to surveillance data than a reduced model without behavior change for all three countries. This suggests that a transition from traditional to sanitary burials is a plausible mechanism to explain the observed outbreak dynamics. In particular, this transition may be responsible for a decreasing force of infection in the later stages of the outbreak resulting in a lower total number of cases than would be predicted by a fixed-behavior model, as well as a longer time until the end of the outbreak.

A noteworthy feature of behavior-disease models is the potential to estimate behavioral parameters from surveillance data. We found that the cost term for non-traditional burials could be directly estimated and the sampling and random choice rate could be determined by LHS. While these parameters are abstractions of the true determinants of burial practices and behavior change, they can provide insight into the relative degree of resistance to non-traditional practices. Behavioral dynamics may explain observed changes in non-mechanistic parameters from fixed-behavior models. Indeed, the behavioral trajectories from our best-fit simulations correspond to previously observed changes in the reporting rate/population at risk correction factor for the reduced model when it is fit to increasing amounts of data [24]. The predicted decline in traditional burials provides evidence that behavior change may have reduced the force of infection, tipping the West Africa epidemic towards ending. However, our model also projects an increase in traditional burials once the epidemic is over as the risk of infection due to funerals is once again low. Thus while behavior change can contribute to the end of an outbreak, populations will not necessarily maintain their adherence with lower risk behaviors.

In order to account for the effect behavioral dynamics, infectious disease surveil-

lance could be supplemented with time series of risk behaviors such as the fraction of traditional burials. Behavior-disease models can use these sources with standard fitting methods. Similarly models can be designed to include behaviors that are observed to change in response to outbreaks. This may be particularly useful given that epidemics typically have similar early trajectories, characterized by a period of exponential growth. During this period it is nearly impossible to discriminate between models of varying complexity with case and mortality data alone. However, behavioral data streams may provide enough additional information to improve model selection or reduce uncertainty in estimates from a given model. In the context of Ebola, collecting the relative frequency of burial types may have improved forecasting accuracy by signaling the reduction in transmission due to increasing sanitary burials.

Our forecasting results suggest that including behavior change may enable more accurate medium and long-term projections even using only standard case and death surveillance data. Our model gave reasonable predictions of the final size of the outbreak (within 8% of the reported final size), although only once data from November, 2014 onward was included. While final size information is often sought by policymakers early in an outbreak, these results underscore how surveillance data from this period may not be sufficient to provide an accurate estimate. It is also difficult to determine ex ante whether a model's final size predictions are likely to be accurate. However, a time varying force of infection due to behavioral adaptation is a plausible element in most outbreaks. In particular, behavior change in our model leads to a reduced force of infection over time, and a lower final size as a result. While still subject to substantial uncertainty during outbreak conditions, these estimates may be more plausible than higher estimates from models without behavior change.

Our findings consistently indicate that burial practices changed significantly over the outbreak, however our model does not necessarily distinguish between changes in behavior due to interventions (e.g. increased burial team activity) and change due to social adaptation. In particular the sanitary burial strategy represents an intervention as opposed to a burial type that existed prior to the outbreak. Burial team deployment also increased as the scope of the epidemic became evident, so it is reasonable to assume that some of the predicted behavior change is capturing this phenomenon. Still, burial teams' effectiveness depended on cooperation from local citizens. As a result, our model can be interpreted as representing community-level adaptation to a behavioral intervention. We also note while the qualitative features of our model's burial trajectories appear plausible, the exact degree of behavior change may be confounded by additional behavioral or intervention activity. For example, Ebola treatment units (ETUs) were deployed in increasing numbers and individuals may have reduced their contact frequency while the epidemic was growing. As current surveillance data is not likely to be sufficient to specify further behavioral mechanisms, future work may need to integrate alternative data sources such as anthropological studies. Similarly, we omit healthcare transmission from our model to reduce its overall complexity. This may bias our estimates of transmission terms somewhat as healthcare workers in Ebola treatment units may have experienced higher risk due to their frequent contact with late-stage patients.

In spite of these limitations, our model provides a platform to test hypothetical behavioral scenarios, which can seldom be studied experimentally. For example, we compared the magnitude of simulated outbreaks between our best fit model and a combination of fixed and dynamic behavior conditions. Both scenarios with dynamic burial behavior resulted in lower peak incidence than either fixed scenario. This

suggests that adaptation can provide protection even in populations that have not experienced prior outbreaks. However, adaptation alone is not sufficient to prevent an outbreak from beginning. In general, behavioral practices do not appear to change substantially until the outbreak is near or past its peak incidence. This reflects the intuition that individuals may never react instantaneously to changing disease conditions. Thus, prevention and rapid response to newly detected outbreaks are still crucial to successful control. As emerging disease outbreaks increasingly occur in complex socio-political conditions, we argue that it is important to continue to develop methods that can provide mechanistic insights into behavioral processes as well as biological ones.

CHAPTER V

Usability, compliance, and household water treatment recommendations

5.1 Introduction

Between 2000 and 2012, the global population lacking access to safe drinking water was reduced by over 50%, fulfilling a United Nations (UN) Millennium Development Goal (MDG) for environmental sustainability three years early. Meeting this goal substantially improved quality of life by reducing the burden of diarrheal disease in low and lower-middle income (LMIC) [98]. However, many challenges remain as nations begin to address the new UN Sustainable Development Goals (SDG). As of 2012, roughly 11% of the population remained without safe drinking water while improvements to sanitation fell short of the MDG. Consequently, water, sanitation, and hygiene feature prominently in the SDG, with the ambitious target of universal access to safe drinking water and adequate sanitation by 2030. While the MDG and other initiatives have significantly reduced the burden of diarrheal disease, it remains the second leading cause of death among children under five [25]. Effective and sustainable interventions will be critical to expanding the coverage of safe drinking water. We focus our analysis on the effectiveness of household water treatment (HWT) considering both compliance behavior and treatment efficacy.

HWT has been promoted as a sustainable intervention to provide safe drinking

water to communities affected by diarrheal disease that lack central treatment systems [31, 99]. Health gains in theory can be substantial, as treatment methods are capable of removing or inactivating nearly all pathogens in drinking water. However, in practice reported risk reductions from HWT are often less than expected [27–30]. Incomplete compliance may be responsible for some of the observed inconsistency between treatment efficacy from small scale trials [100] and population health outcomes. In particular, higher treatment efficacy often compromises end-user convenience or device reliability [30]. For example, chlorine is increasingly detectable by taste as concentrations increase, and can render drinking water unpalatable [31, 32]. Water filtration devices can achieve similar pathogen reductions without altering taste, but slow flow rates and clogging can impede adoption and long-term effectiveness [33–35]. Under these circumstances, the efficacy of a HWT method does not capture its actual capacity to reduce the burden of disease, as a more efficacious treatment device may induce lower compliance and therefore be less effective than a less efficacious but more appealing intervention.

HWT efficacy is typically reported as $\log_{10}$-removal values (LRVs) that quantify the amount of pathogen removed from treated water, i.e., the proportion of pathogen remaining after treatment is 10^{-LRV}. LRVs serve as a comparative measure for treatment methods within a given class or between classes. The World Health Organization (WHO) currently recommends HWT efficacy of 4 LRV for bacteria, 5 LRV for viruses, and 4 LRV for protozoa in order to attain a "highly protective" standard in generic scenarios where contextual information about exposure levels and the population at risk are not available [31]. The specific LRVs presented in the Guidelines were computed using analytical quantitative microbial risk assessment (QMRA) with disability-adjusted-life-years (DALY) as the measure of disease bur-

den. Notably, the WHO guidelines assume perfect compliance with the treatment method. However, technological adoption is seldom complete [101]. HWT uptake in particular is variable but rarely widespread in lower income regions [102].

Current QMRA research has begun to address the impact of compliance on the effectiveness of HWT interventions [103–105]. However, these studies have treated compliance as essentially exogenous. That is, compliance is set to a range of hypothetical levels for a given intervention. By contrast, it is likely that uptake and compliance are determined by the degree to which an intervention matches individuals' preferences regarding tradeoffs between treatment efficiency and usability. Thus overall compliance depends both on the distribution of attitudes and the specific intervention proposed. We explicitly model this relationship by augmenting the QMRA approach with a decision-theoretic model of compliance. Decision theory is a mathematical framework used in economics and other social sciences to analyze scenarios where individuals must choose between multiple options subject to cost and benefit comparisons. Applications to public health have largely focused on cost analysis for institution-level interventions [106, 107]. However, decision theory is equally appropriate to modeling individual level compliance decisions. We use our combined model to evaluate current treatment recommendations and estimate recommendation levels that minimize disease burden given a range of hypothetical population attitudes toward HWT.

5.2 Methods

5.2.1 Quantitative Microbial Risk Assessment (QMRA) model

QMRA provides a framework to estimate disease risks using a set of equations that characterize how likely an individual is to become infected given their degree of exposure to a given pathogen. For waterborne disease, contaminated drinking water

acts as one of the primary transmission pathways, so exposure levels represent the quantity of viable pathogen ingested daily based on the quality of available drinking water (exposure assessment). The probability of infection per organism depends on interactions between the pathogen and the host immune system (dose-response). Empirically, infection events can be modeled using a dose-response function fit to experimental data.

QMRA can be implemented using either an analytical or (stochastic) simulation-based approach [103, 104]. Using an analytical approach risk is directly calculated using data for exposure levels. For a stochastic QMRA simulation, risk is estimated from an ensemble of simulation runs. In a given simulation, exposure levels can vary, and infection is determined randomly according to the probability distribution specified by the dose-response function. We use both methods to address the effect of compliance on HWT interventions. Specifically, we use the analytical version of our model to compute optimal recommendations and the stochastic model to assess the burden of disease.

Calculating exposure levels An individual's daily volume of pathogen ingested is

$$d_i = \begin{cases} wv \times 10^{-\hat{x}} & \text{with HWT compliance} \\ wv & \text{otherwise} \end{cases}, \tag{5.1}$$

where w is the concentration of pathogen per liter of untreated water, v is the volume of water consumed per day, and $\hat{x}$ is the $\log_{10}$-reduction value of the recommended HWT method. Individuals use their HWT device (comply) with probability $Pr_i(use)$. Here, we choose to model compliance based on individuals' attitudes toward the specific implemented HWT, by developing a decision-theoretic model. While not implemented in this analysis, it is also possible to treat compliance as

a dynamic element, responding to disease burden, environmental changes, or social changes.

The expected dose ($E[d_i]$) across all individuals in a population can be characterized as:

$$E[d_i] = wv[(1 - E[Pr_i(use)]) + E[Pr_i(use)]10^{-\hat{x}}] \tag{5.2}$$

where $E[Pr_i(use)]$ is the expected compliance. On average, individuals are exposed to fully contaminated water when they do not use their treatment device ($wv(1 - E[Pr_i(use)])$) or reduced pathogen content when they do ($wv(E[Pr_i(use)]10^{-\hat{x}})$. Alternatively, expected compliance can be interpreted as the fraction of a given day's water that is effectively treated. These interpretations yield identical analytical results, but do alter the disease outcomes in an explicit simulation.

Dose-response We use an exponential or approximate beta-Poisson dose-response function to compute the daily probability of infection for a given quantity of pathogen. Both of these functions assume that a single pathogenic organism has a non-zero probability of causing an infection, essentially treating infection as the outcome of Bernoulli trials. The exponential dose-response function is

$$Pr_i(infection) = 1 - e^{-kd_i} \tag{5.3}$$

with rate parameter k. Mechanistically, this function implies that the dose is Poisson distributed and that each unit of pathogen has an identical probability of surviving to reach the target site ($1/k$) and of causing an infection.

An exact beta-Poisson function is often computationally unstable due to its use of the confluent hypergeometric function. As a result, the approximate form is often

used instead.

$$Pr_i(infection) = 1 - (1 + \frac{d_i}{\beta})^{-\alpha} \tag{5.4}$$

$$\beta = \frac{N_{50}}{2^{1/\alpha} - 1}. \tag{5.5}$$

where α controls the slope and N_{50} is the dose required to infect 50% of a population. This approximation is appropriate when $\alpha << \beta$ and $\beta >> 1$ which are satisfied by our parameter values for *E. coli* and *rotavirus* (Table 5.1). The mechanistic interpretation of the beta-Poisson model is similar to that of the exponential model, however in this case the probability that a pathogen survives to infect (i.e., infectivity) is assumed to be given by a beta distribution. The choice of dose-response function is typically made based on both biological and statistical considerations. We use the exponential dose-response function for *Cryptosporidium* [108] while the beta-poisson dose-response function is used to characterize *E.coli* [109] and rotavirus [110].

5.2.2 Decision-theoretic compliance

Our QMRA model can accommodate variable individual compliance levels. In particular, we are interested in the impact of recommendations on compliance. We construct the following decision-theoretic model to determine the distribution of individual compliance based on hypothetical attitudes toward recommended HWT levels.

Suppose we have individuals $i \in N$ who must each select a probability of compliance with a recommendation chosen from the intervention space $X \subseteq \mathbb{R}$. This space represents the range of possible HWT levels quantified by their LRV. We assume that an individual's attitude toward HWT properties map to the LRV of any given treatment method. Each individual has a most preferred LRV denoted $x_i \in X$. We will refer to the distribution of these points as the *preference distribution*. For our

analysis we assume that preferences are distributed according to a truncated normal distribution with mean μ and variance σ^2, bounded by $[0, 6]$ (the range from no intervention to the highest current recommendation). We label the potential recommended intervention $\hat{x}$. In our model, the choice of a specific optimal recommendation level can be a function of both the LRV of the device and population preferences. By contrast, the current WHO recommendations for HWT are based solely on the microbiological characteristics of a device.

We frame the following decision problem: Given their preferred LRV, individuals must choose the degree to which they comply with the recommended HWT. As noted above, this choice is over the probability of compliance as opposed to the binary choice of compliance on a specific day. This is because we assume that conditions informing compliance do not change enough between days to alter an individual's choice. Instead, individuals set $Pr_i(use)$ when the intervention is implemented and draw their daily compliance accordingly, analogous to a mixed strategy in game theory. To represent this problem we construct a *utility function*, u_i, that represents an individual's cost/benefit evaluation of the recommendation $\hat{x}$. Intuitively, this function should yield a larger value for compliance with an intervention that is similar to the preferred LRV and a smaller value for compliance when the recommended LRV is dissimilar to the preferred LRV. We use the following general form for our compliance model:

$$u_i(Pr_i(use)|x_i, \hat{x}) = -(1 - Pr_i(use) - \Delta x)^2 \tag{5.6}$$

where Δx represents the distance between x_i and $\hat{x}$. The choice of a distance measure may depend on prior knowledge of how individuals compare HWT alternatives. We use two similar measures to characterize different potential situations. The first

measure is the squared Euclidean distance

$$\Delta x = \frac{(x_i - \hat{x})^2}{(max(X) - min(X))^2}.$$

Where $min(X)$ and $max(X)$ are the lowest and highest feasible LRVs, respectively. This measure implies that individuals dislike treatments that are either more or less efficient than their ideal preference. Chemical treatment such as chlorination may be an example of an intervention for which individuals apply a symmetrical distance measure. This may be because an individual who prefers some level of chlorination would be unwilling to treat their water at levels that they do not perceive as effective (lower LRV than preferred), and may not want to treat at high concentrations due to taste issues [31, 100].

Alternatively, we can use an asymmetrical piecewise distance measure

$$\Delta x = \begin{cases} \frac{(x_i - \hat{x})^2}{(max(X) - min(X))^2} & \text{if } \hat{x} \geq x_i \\ 0 & \text{otherwise} \end{cases}. \tag{5.7}$$

Unlike the first utility function, this variant implies that individuals dislike recommendations greater than their ideal point, but treat lower recommendations as equally favorable. This asymmetry may be appropriate for filtration methods, where individuals may not distinguish between a lower LRV device and the recommended treatment level, especially if changes in the aesthetic qualities of the filtered drinking water do not vary between devices. They may begin to become non-compliant, however, with a high LRV filter due to slow water flow or increased breakage rate. If X is normalized to the $[0, 1]$ interval, the above equations simplify to

$$\Delta x = (x_i - \hat{x})^2 \tag{5.8}$$

and

$$\Delta x = \begin{cases} (x_i - \hat{x})^2 & \text{if } \hat{x} \geq x_i \\ 0 & \text{otherwise} \end{cases} \tag{5.9}$$

Decision theory posits that individuals will choose actions that maximize their utility. The inner term of Equation 5.6 implies that when the distance between an individual's preference and the recommendation is large, utility is maximized by adopting a low probability of compliance. By contrast, when the distance is small, utility is maximized by adopting a high probability of compliance. Thus, maximizing Equation 5.6 with respect to $Pr_i(use)$ results in the following probability of compliance that approaches one when an individual's preference matches the recommendation.

$$Pr_i(use) = 1 - \Delta x \tag{5.10}$$

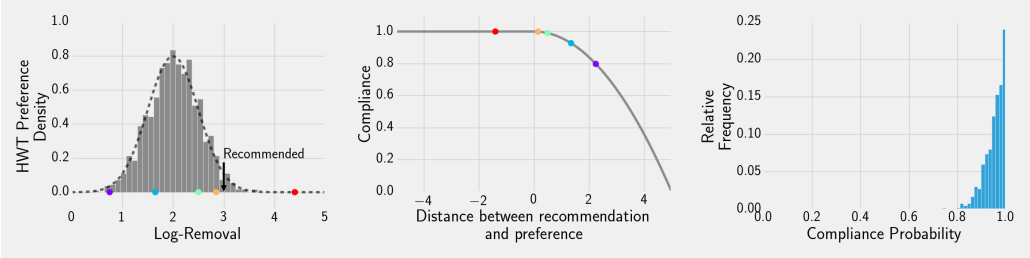
Equation 5.10 implies that an individual will comply perfectly with their most preferred HWT intervention. This is not realistic in the sense that even widespread centralized water treatment never achieves 100% usage. When the maximum possible compliance is known, it is straightforward to modify Equation 5.10 using a scaling factor as follows:

$$Pr_i(use) = c_{max}(1 - \Delta x) \tag{5.11}$$

where c_{max} is the maximum possible compliance. This does not impact our estimates of optimal recommendations but does have an impact on our risk estimates. In particular we would expect a lower c_{max} to attenuate the risk reduction for any intervention. This is because risk responds monotonically to compliance for a given intervention level, so lowering the maximum compliance probability acts as an offset.

Applying the decision theoretic model to home water filtration Household water filters include a range of specific technologies including biosand [111], and ceramic [112] devices. As a result, filters vary with respect to their filtration efficacy for viral, bacterial, and protozoan parasites as well as properties such as flow rate, capacity, and durability. While individuals facing endemic diarrheal disease are likely to value improved pathogen removal, the usability of a filter may significantly impact whether an individual is willing to treat their drinking water, particularly in resource-limited settings where individuals may not prioritize water treatment in the face of other basic survival concerns. In particular, more efficient filters may have reduced usability as smaller pores decrease the flow rate and are more likely to clog. As a result, attitudes toward filters are not likely to strictly improve with efficacy. Indeed, the trade-off between efficacy and usability may result in decreasing favorability once a filter passes a point that an individual considers acceptable. However, filters below that threshold may be considered equally acceptable, given that LRV may not be explicitly evaluated so long as a filter reduces apparent risk.

We illustrate our decision-theoretic approach using this context. Suppose preferences regarding filter efficiency are normally distributed with a mean of 2 LRV (Figure 5.1a). We wish to determine the distribution of compliance if a 3 LRV filter is recommended. Each individual selects their compliance level with the 3 LRV device based on Equation 5.6 using the asymmetric distance measure. All else equal, this measure implies that individuals dislike filters more efficient than their ideal point but are ambivalent to filters as efficient or less. Figure 5.1b shows predicted compliance (Equation 5.10) as a function of the distance between the recommendation and any given individual's preference. In this example most of the population prefers a filter relatively similar to the recommendation, so overall compliance is rel-



(a) (b) (c)

Figure 5.1: An example of our decision-theoretic framework. The preference distribution (5.1a) specifies the frequency of each individual preference. Colored circles indicate five specific example individuals with corresponding LRV preferences. Individual maximum-utility compliance (5.1b) can be expressed as a function of the distance between an individual's preference and a given recommendation. The compliance distribution (5.1c) gives the frequency of predicted compliance for all individuals by applying the maximum-utility compliance to each individual's preference given a recommendation.

atively high (Figure 5.1c). Note that the slope of the compliance function determines the degree to which individuals dislike dissimilar interventions. When the slope is steeper, the compliance distribution is broader with more mass at lower compliance relative to a shallower slope (Appendix C.3).

5.2.3 Optimizing HWT recommendations

In an analytical QMRA approach that assumes perfect compliance, we can always compute the treatment level (efficacy) necessary to obtain a given risk threshold. With incomplete compliance it is possible that no feasible treatment level could reduce risk below current acceptable disease burden standards. However, it is still important to consider which treatment level would reduce infections the most relative to baseline conditions. An optimal recommendation from this perspective must take into account both the microbiological characteristics of the device and behavioral features of potential users. Formally, this problem can be stated as follows:

$$\hat{x}^* = argmin_{\hat{x}} Pr(infection|\hat{x}, \theta), \tag{5.12}$$

where $argmin_{\hat{x}}$ means that we search for the recommendation $\hat{x}^*$ that minimizes the average probability of infection which is a function of the treatment level $\hat{x}$ and other parameters θ (i.e. the dose response function).[1] Equivalently, when the dose response function responds monotonically to changes in dose, the optimal recommendation can be found by minimizing the function representing the expected dose (Equation 5.2). Notably, this means that the solution does not depend on pathogen or exposure characteristics beyond the effect of treatment. We compute numerical solutions to Equation 5.12 assuming normally distributed preferences. Appendix C.1 describes our procedure in greater detail. We assess the effect of the preference distribution's variance using local sensitivity analysis.

5.2.4 Simulation framework

We simulate a population of size N for T days. Each day healthy individuals may become infected based on their exposure level and probability of infection. Sick individuals recover based on times drawn from a gamma distribution. Table 5.1 describes the specific dose-response and recovery models used for each pathogen [104]. Gamma distributions characterize the expected time to recovery for diseases with multiple infectious stages assuming a Poisson process. When represented by an integer, the shape parameter denotes the number of stages. In the case where the shape parameter is one, the gamma distribution is equivalent to an exponential distribution. For *Cryptosporidium* and *E. coli*, gamma distribution parameters were drawn from existing literature on the infectious period of each disease [113, 114]. Because less data is available for Rotavirus we chose a gamma distribution with an average waiting time equal to the median recovery time of 5.2 days [115] and a shape parameter of one (making the distribution effectively exponential). We implemented

[1] This procedure is equivalent to solving for the subgame perfect Nash equilibrium for a game in which a policy maker first selects a recommendation and individuals then choose their compliance probabilities (Appendix C.2).

	Dose-Response	Dose-Response Parameters	Recovery Distribution	Recovery Parameters
Cryptosporidium	Exponential [108]	$k = 5.72 \times 10^{-2}$	Gamma	$shape = 4, scale = 2.5$
E. coli	Beta-Poisson [109]	$\alpha = 0.155, N_{50} = 2 \times 10^6$	Gamma	$shape = 1.775, scale = 1.69$
Rotavirus	Beta-Poisson [110]	$\alpha = 2.53 \times 10^{-1}, N_{50} = 6.17$	Gamma	$shape = 1, scale = 5.2$

Table 5.1: Dose-response functions and recovery time distributions for each pathogen. The average time to recovery for a gamma distribution is the product of the shape and scale parameters. For an exponential distribution the average time is equal to the scale parameter.

	E. coli (STEC)	Rotavirus*	*Cryptosporidium*
Organisms/L	1×10^3	1	0.1
Daily water consumption (L)	1	1	1
Treatment efficacy (LRV)	4	5	4
DALY/person	5.47×10^{-2}	0.482	1.47×10^{-3}
Prevalence/year	2.92×10^{-4}	0.0136	2.26×10^{-3}
Disease Burden (DALY/person-year)	1.59×10^{-5}	6.54×10^{-3}	3.32×10^{-6}
WHO threshold	1×10^{-6}	1×10^{-6}	1×10^{-6}

Table 5.2: Simulated disease burden estimates for three waterborne enteric pathogens at 6 Log_{10} reduction with complete compliance. *Rotavirus assumes 6% population at risk, low income disease burden [31, 117].

our models and analyses in Python 2.7 using Numpy, Scipy, and Matplotlib.

5.3 Results

5.3.1 The effect of imperfect compliance

We first replicated the scanarios that the WHO used to determine treatment level guidelines, a 4-$\log_{10}$ reduction in bacterial and protozoa concentration and a 5-$\log_{10}$ reduction in virus concentration with perfect compliance [116]. For each pathogen type we simulated our stochastic QMRA model for 1 year and computed the average yearly disease burden. With perfect compliance our simulations for *E.coli* and *Cryptosporidium* correspond to the WHO analytical results (Table 5.2). We use the low income country disease burden for rotavirus, resulting in a higher total disease burden. This assumes a higher probability of mortality due to infection and is more likely to reflect the burden of rotavirus in countries that will receive the greatest benefit from the 2030 SDG.

Next we relax the assumption of perfect compliance for *Cryptosporidium* (Results for other pathogens can be found in Appendix C.3). At all contamination levels over

99% compliance is necessary to reach the WHO target (Figure 5.2). This finding is a consequence of two factors: The WHO guidelines are determined by solving for the lowest efficacy that will result in tolerable disease burden, so we would not expect a less efficacious intervention to meet the threshold. Additionally, individuals face substantially higher disease risk whenever they do not use their treatment device, causing disease burden to be very sensitive to compliance. While the largest changes occur between compliance levels of 80-100%, the log-scale for disease burden compresses its apparent change with respect to compliance. For example, with 1 oocyst/L, if compliance increases from 20% to 60%, the disease burden decreases from 0.134 DALY/year to 0.0827 DALY/year.

Note that the WHO target threshold of 10^{-6} DALY/year implies near-zero endemic prevalence based on analytical QMRA. This is impossible to verify in practice due to the large population size required to detect any cases once the risk of infection is sufficiently low. Our stochastic simulation results reflect this phenomenon, as many runs with near-perfect compliance had zero cases. As a result, although the disease burden for incomplete compliance is higher than the 10^{-6} threshold, contamination levels of 0.01 and 0.001 oocysts/L cause a very small absolute number of cases on average.

5.3.2 Optimal recommendations

In our model, optimal LRV recommendations are those that most reduce risk subject to a tradeoff between compliance and device efficacy. Simulating a filter intervention, we assume that the utility function has an asymmetrical distance measure, i.e., users accept recommendations lower than their preference but are less likely to use recommended filters that have higher LRV than their preference (Figure 5.1b). Based on these assumptions optimal recommendations for filters tend to be higher

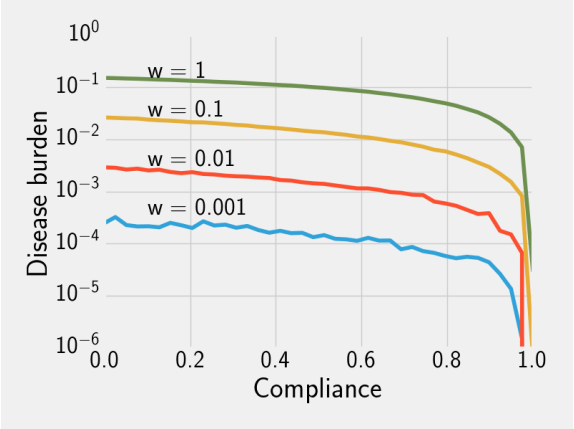
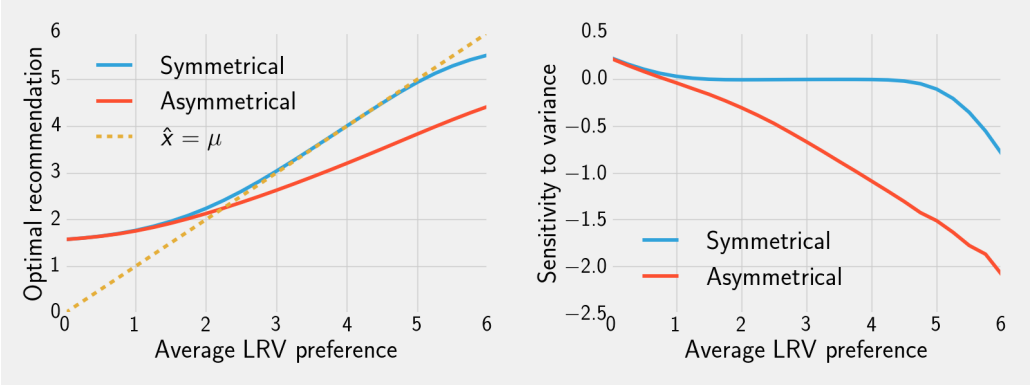


Figure 5.2: Simulated disease burden estimates for *Cryptosporidium* at varying contamination levels w as compliance with a 4 Log_{10}-reduction recommendation increases.

than the average user preference μ when the average is low and tend to be lower than the average preference when the average preference is high (Figure 5.3a). For our simulation of *Cryptosporidium* the transition point is approximately at 2 LRV. This can be seen by comparing the solution line with the dashed line indicating a recommendation set at the mean of the preference distribution (Figure 5.3a).

We also determine the optimal recommendations for a chlorination intervention. In this case we assume that the distance measure is symmetrical – individuals are less likely to treat their water if the recommended LRV is either higher or lower than their preference. Like filtration, the optimal recommendation for chlorination is higher than the average preference when the average preference is below 2. However, between 2 and 5 LRV the optimal recommendation tracks the average preference. Above an average preference of 5 LRV the optimal recommendation is slightly below the average user preference.

The symmetrical distance measure results in higher optimal LRV recommendations than the asymmetrical measure for distributions with an average preference above 2 LRV. This is because the asymmetrical measure implies that individuals



(a) (b)

Figure 5.3: Numerical solutions (5.3a) for the optimal recommendations using the decision-theoretic compliance model. We also compute the sensitivity coefficient for the optimal recommendation $\hat{x}^*$ with respect to the variance of the preference distribution (5.3b). Blue lines indicate results using the symmetrical distance function while red lines indicate results using the asymemtrical distance function. The dashed line tracks the average LRV preference μ. LRV preferences are normally distributed with variance $\sigma^2 = 3.6$.

will comply with treatments with lower LRV than their most preferred alternative. Consequently, overall compliance with any intervention under these conditions is higher than in the case where individuals also dislike less efficient treatments. Our sensitivity analysis highlights a further distinction between the two distance measures. The optimal recommendation is insensitive to the variance of the preference distribution when individuals use a symmetrical distance measure. However, variance does have a substantial effect on the optimal recommendation if individuals use an asymmetrical measure. This phenomenon is also driven by the observation that recommendations under the asymmetrical distance measure gain compliance from anyone with a greater preference. Thus, for a given average preference, increasing variance increases the expected number of people who have a larger LRV preference.

Case study Our model can also be used to estimate optimal recommendations given data about compliance with existing interventions. We demonstrate this using a hypothetical water filtration intervention. For our scenario, we assume that the

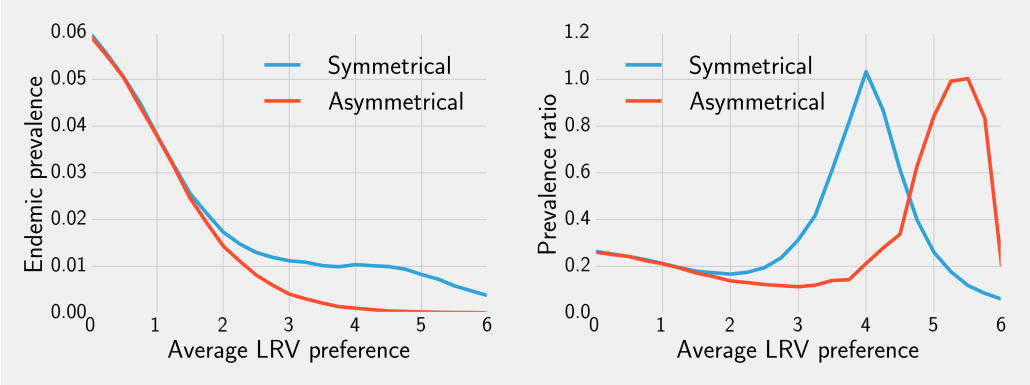
maximum compliance for any intervention is 90%. Suppose field trials of a 6 LRV filter indicate that participants do not use their device 20% of the time on average. For simplicity we assume that the variance of compliance is not available. From this we can calculate the average LRV preference $E[x_i]$ by solving Equation 5.10:

$$\begin{aligned} E[Pr_i(use)] &= c_{max}(1 - \Delta x) \\ E[x_i] &= \hat{x} - \sqrt{x_{max}^2(1 - \frac{E[Pr_i(use)])}{c_{max}})} \end{aligned} \tag{5.13}$$

where c_{max} is the maximum compliance. We use the negative square root since our model for filters assumes an asymmetric distance measure which implies that incomplete compliance is generally a product of an intervention that is more efficacious than the average user preference. Our scenario specifies that $c_{max} = 0.9$, $E[Pr_i(use)] = 0.8$, and $x_{max} = 6$ LRV. Solving Equation 5.13 gives an average preference $E[x_i] = 4$ LRV. We then solve Equation 5.12 numerically to obtain the optimal recommendation $\hat{x}^* = 4$ LRV, the same value as the average user preference. This is a conservative estimate in the sense that variance in the preference distribution would cause the optimal filter recommendation to be lower than the average preference (Figure 5.3). However, without an explicit variance estimate, the optimal recommendation solution resembles the value for a symmetrical distance measure.

5.3.3 Risk reduction

We also examine the degree to which optimal recommendations reduce the burden of disease for the specific pathogens listed in Table 5.1. Figure 5.4 depicts the absolute risk of Cryptosporidiosis in a simulated population of 10,000 individuals as well as the relative risk comparing the optimal recommendation level to the current WHO recommendation for protozoa (4 LRV). We use the stochastic simulation model for this analysis as discrete cases provide more realistic disease burden estimates, par-



(a) (b)

Figure 5.4: 5.4a: The simulated endemic prevalence when the optimal LRV is implemented comparing the symmetrical and asymmetrical distance functions. 5.4b: The prevalence ratio comparing the optimal recommendation from Figure 5.3 vs. the current 4 LRV recommendation for *Cryptosporidium* for both distance functions.

ticularly when the risk of infection is low. Our simulations assume 1 oocyst per liter and 2 liters of water consumed per day. We again compare the results using symmetrical and asymmetrical distance measures. As noted above, the symmetrical distance measure can represent attitudes toward chlorination while the asymmetrical distance measure can represent attitudes toward filters. For concreteness we will refer to the model with symmetrical/asymmetrical distance as the chlorination/filtration model, respectively.

The risk estimates in Figure 5.4a demonstrate the effect of a population's attitudes toward HWT on the intervention's capacity to reduce disease burden. We report results for a maximum compliance $c_{max} = 1$. Thus these risk estimates are optimistic in the sense that lower maximum compliance will shift the risk and risk-ratio curves downward (Section 5.2.2). We show supplementary results for a set of reduced maximum compliance scenarios in Appendix C.3. For both chlorination and filtration, optimal interventions become much more effective as the average preference increases from 1 to 2 LRV. Beyond this point, the effectiveness of chlorination

plateaus while optimal filtration continues to reduce risk for populations with greater average preference. In both cases, Figure 5.4b indicates that optimal recommendations are generally more effective than the current recommendation. The protective effect of the optimal recommendation is largest when the average LRV preference does not align with the current recommendation. Simulations for *E. coli* and Rotavirus can be found in Appendix C.3. While specific values differ, the qualitative features of the absolute and relative risk curves do not vary substantially by pathogen.

5.4 Discussion

Current HWT guidelines focus on treatment efficacy under a perfect compliance scenario. However, even nearly perfect compliance with a high LRV treatment method can result in high disease burden when drinking water is sufficiently contaminated. Studies of filter adoption have suggested that usability may decline as a function of LRV. Biosand and clay pot filters have LRV approaching 5 for bacteria, but are prone to breakage and clogging [33, 118, 119]. By contrast cloth filters have been successfully adopted for Cholera prevention in spite of a much lower (2 LRV) efficacy [120, 121]. Chemical treatment involves a similar trade-off as the taste of treated water typically becomes less palatable as the concentration of disinfectant increases [31, 32]. It is important, therefore, to determine how the interaction between microbiological and usability characteristics of HWT treatments influences their effectiveness as interventions.

Recent QMRA analyses have found evidence for diminishing health improvement returns for increasing LRVs under imperfect compliance [104, 105]. Our results are broadly consistent with these findings. Additionally, our decision-theoretic framework provides a means to examine the interaction between a population's attitudes

and the effectiveness of HWT interventions and to develop optimal recommendations that account for non-compliance. Applications of our approach are not limited to HWT contexts, indeed, our compliance model could be adapted for WASH or other interventions provided the utility function is a reasonable representation of a target population's attitudes toward interventions.

Across the range of population attitude scenarios we tested, the optimal LRV from the perspective of minimizing disease burden was nearly always lower than the current recommendation that assumes perfect compliance. In addition, we found that the risk-reduction generated by applying the optimal LRV as opposed to the current recommendation for each class of pathogen was considerable (risk-ratio: 2-15). This illustrates how assuming perfect compliance may yield overly optimistic projections of disease reduction by high efficacy treatment methods. It is important to note that our optimized interventions did not result in absolute disease risks at or under the tolerable disease burden threshold of 1×10^{-6}. These higher risk predictions are both a consequence of our explicit inclusion of compliance and the stochastic nature of our model. Given that the prevalence needed to obtain disease burdens at or lower than 1×10^{-6} DALY/year are not observable in discrete populations, stochastic QMRA may be a more realistic approach to assess disease burden and risk targets. In addition, a combination of intervention strategies (a multiple barrier approach) may be necessary to sufficiently reduce the burden of diarrheal disease in developing countries.

Willingness to adopt more efficient (but potentially less usable) HWT methods is likely to vary substantially by region. Ideally, HWT recommendations could be informed by data regarding the target group's preferences and attitudes regarding treatment. Such attitudes are complex, determined by a wide range of elements

including private costs, social contexts, and political forces. Our decision-theory model is designed to accommodate behavioral data at multiple levels of resolution. In particular, when data on compliance is unavailable or unreliable, attitude surveys and similar techniques can be used to assess the distribution of preferences and inform the selection or construction of a utility function. Alternatively, as demonstrated by our household water treatment case study, data on compliance with existing interventions can be used to infer the optimal recommendation from an effectiveness perspective.

Our QMRA model relies on a simplified representation of enteric pathogen transmission. In order to focus on the implications of intervention and preference-dependent compliance we omitted temporal variation in pathogen exposure due to seasonal or other periodic factors. Additionally, QMRA approaches rely on the simplifying assumptions that infected individuals do not shed pathogen back into drinking water sources, and that contaminated drinking water is the primary transmission pathway. Future work may address these factors by implementing a compartmental transmission model with environmental transmission similar to the EITS or SIWR models [122, 123]. However, our current approach represents a straightforward extension to QMRA that functions as a policy evaluation and data integration tool. Even in the absence of context-specific information, our analyses suggests that it may be advisable to focus on cost-efficient and readily usable treatment options. This is likely to be particularly crucial to the successful deployment of HWT in areas that continue to lack access to clean drinking water after the MDG.

CHAPTER VI

Conclusion

The contribution of human behavior to infectious disease transmission is complex, but not intractable. Decision theory and game theory provide a convenient mathematical framework to capture individual and interactive decision problems. Alongside more common tools such as deterministic compartmental models, these methods extend the reach of mathematical epidemiology. Such developments are likely to be of increasing importance as globalization, economic development, and social progress create challenges for disease control and prevention. The three behavior-disease models developed in this dissertation cover major infectious disease contexts: outbreaks in higher income, developed nations, emerging diseases, and control of endemic pathogens.

Chapter III investigated the impact of adaptive prophylaxis use on the dynamics of sexually transmitted infections. Our model explicitly represented uncertainty in sexual partner disease status by treating the prophylaxis use decision as a Bayesian game. We used evolutionary dynamics to represent non-rational behavior change over time, linked to an SIS transmission model through the contact rate parameters. We also included incentives for infected individuals to adopt protective behavior, an observed phenomenon that is often omitted from STI models – even those that treat

behavior change. The dynamics of our model depended on the infectivity of disease as well as the speed with which individuals adapted to changing disease conditions. Notably, prevalence oscillated when the disease was moderately infectious and the adaptation rate was slow, as susceptible and infected individuals switched their use of prophylaxis. These oscillations can represent the temporal trends of diseases over multi-year time periods. While recurrent outbreaks seldom occur in a single year, many diseases do exhibit seasonal patterns of incidence [124, 125]. These periodic outbreaks are often modeled using non-mechanistic forcing functions or are treated as separate instances for the purpose of model fitting. However, we have demonstrated that periodic disease trajectories can be caused by behavioral adaptations that change the effective contact rate based on disease pressure. As a consequence, unlike a fixed-behavior model, $\mathcal{R}_0$ is more complicated to evaluate as a threshold parameter. Highly infectious diseases (high $\mathcal{R}_0$) rapidly became extinct following an initial outbreak as the rapid growth of prevalence incentivized high levels of prophylaxis use. Additionally, due to its dependence on the initial fraction of individuals using protection, outbreaks were possible for $\mathcal{R}_0 < 1$. These findings suggest the importance of understanding the incentive structures that drive risk behavior. In addition, increases in unprotected sexual acts once outbreaks decline indicate that intervention programs may need to invest substantial resources to maintain sufficiently high levels of coverage to prevent recurrent outbreaks, or pursue approaches that directly target either the perceived costs or benefits of prophylaxis.

Chapter IV we constructed a behavior-disease model to address changes in burial practices during the 2014 Ebola epidemic in Guinea, Liberia, and Sierra Leone. Adding behavioral dynamics in this way gives the model a mechanistic way to change the force of infection over time without using a pre-specified forcing func-

tion. Our analysis focused on parameter estimation and forecasting to test the value of a behavior-disease model in a complex emerging disease context. Our best-fit model substantially outperformed a reduced model with fixed burial practices both in terms of reproducing the observed surveillance data and forecasting the final size of the epidemic. Simulated trajectories of the fraction of traditional burials over time show a sharp decline in traditional burials in favor of sanitary burials between August and October, 2014. This corresponds both with the period directly after the peak incidence and a change in the estimated value of the reporting rate/population at risk parameter k from the reduced model. Combined with the quality of our model fits, this suggests that behavior change did indeed cause a measurable reduction in the incidence rate over time. We note that this change was observable using only cumulative case and death surveillance data. Additional data streams that directly measure burial practices and other behavioral factors would no doubt improve our ability to forecast and make inferences about the precise extent of adaptive behavior change, however it is encouraging that the coupled behavior-disease approach can provide insight while using commonly available epidemiological data sources. Our model also projected a slow recovery of traditional burials after the outbreak. Similar to adaptive prophylaxis use, this suggests that populations do not completely sustain behaviors with disease-related incentives once there is little risk of disease. As a result, intervention activities may need to be continued for a longer period following the end of outbreaks.

While Chapters III and IV address the implications of the behavior-disease approach on dynamics and inference, Chapter V is concerned with disease control policy. We presented a novel risk model combining decision theory and QMRA to evaluate household water treatment interventions for the prevention of waterborne

diarrheal disease. Current recommendations assume perfect compliance, which is unrealistic even under ideal implementation conditions. In order to assess more realistic compliance levels, we used a rational choice approach to predict compliance with recommended interventions based on the underlying preferences of the target population. This allowed us to optimize water treatment recommendations including both the microbiological characteristics of a device and behavioral characteristics of its potential users. Recommendations selected in this way were nearly always more protective than current efficacy-based recommendations and required a lower device efficacy. These results suggest that incomplete compliance with current HWT recommendations due to usability concerns may compromise their effectiveness and sustainability relative to less efficacious but more appealing measures. Our framework was designed to be easy to integrate into existing QMRA-based intervention policy evaluations, and provides multiple possible avenues to integrate behavioral data.

These models and analyses have evaluated theoretical implications of including mechanistic behavioral models as well as applications to inference about transmission mechanisms and policy evaluation. While our models contain necessary simplifications to allow us to focus on the effect of behavioral pathways, they can be easily extended or modified for additional mechanistic detail or to address different pathogens and behaviors. Sample Python 2.7 code for each project can be found at `https://github.com/malhayashi/dissertation`.

APPENDICES

APPENDIX A

Supplementary material for Chapter III

A.1 Bayesian Games

This section provides a brief overview of static Bayesian games of incomplete information. Those desiring a more complete treatment may refer to [49],[50],and [51].

A.1.1 Definition

A normal form n-player symmetric static Bayesian game includes

- A set of players $N = \{1, 2, \ldots, n\}$.
- Actions $a_i \in A$ for each player.
- Types $\theta_i \in \Theta$ for each player.
- Pure type-contingent strategies $\sigma : \Theta \to A$. By convention, we use the notation $\sigma_j = \sigma_j(\theta_i^1)\sigma_j(\theta_i^2)...\sigma_j(\theta_i^m)$ to represent the jth strategy for a finite type space (assuming the set of strategies is the same for all players).
- Belief distributions p_i where $p_i(\theta_{-i}|\theta_i)$ is the conditional distribution on the types of other players given player i's type. $-i$ denotes the set of players except i.

- The common prior distribution $Pr(\theta_i)$ which is known by all players.
- Type-dependent payoffs $u_i : A^n \times \Theta^n \rightarrow \mathbb{R}$ for each player. The expected type-dependent payoff ($E[u_i(a_1, ..., a_n|\theta_i)]$) gives the average payoff over player i's belief regarding the other players' types conditional on player i's own type. $E(u_i(a_1, ..., a_n, \theta_1, ..., \theta_n))$ is player i's unconditional average payoff over all n player types, sometimes denoted $E(u_i(\sigma_{j_1}, ..., \sigma_{j_n}))$ in terms of strategies.

In general, a *profile* of types, actions, or strategies is defined to be a listing of the particular types/actions/strategies (respectively) assigned to each player. While the actions in a static Bayesian game take place simultaneously, it is useful to break the game down into stages as follows

1. Nature chooses a profile of types $(\theta_1, \theta_2, \ldots, \theta_n)$ from the common prior distribution.
2. Each player picks a strategy σ_{j_i}
3. Each player learns only his type θ_i.
4. Using Bayes' theorem and the common prior, each player forms beliefs $p_i(\theta_{-i}|\theta_i)$ over other players' types.
5. Players choose actions simultaneously according to their strategy $(a_1 = \sigma_{j_1}(\theta_1))$ to form a profile $(a_1, a_2, \ldots, a_n)$.
6. Players receive their payoffs $u_i(a_1, ..., a_n, \theta_1, ..., \theta_n)$ based on the action profile and the type profile.

A.1.2 Bayesian Nash Equilibrium

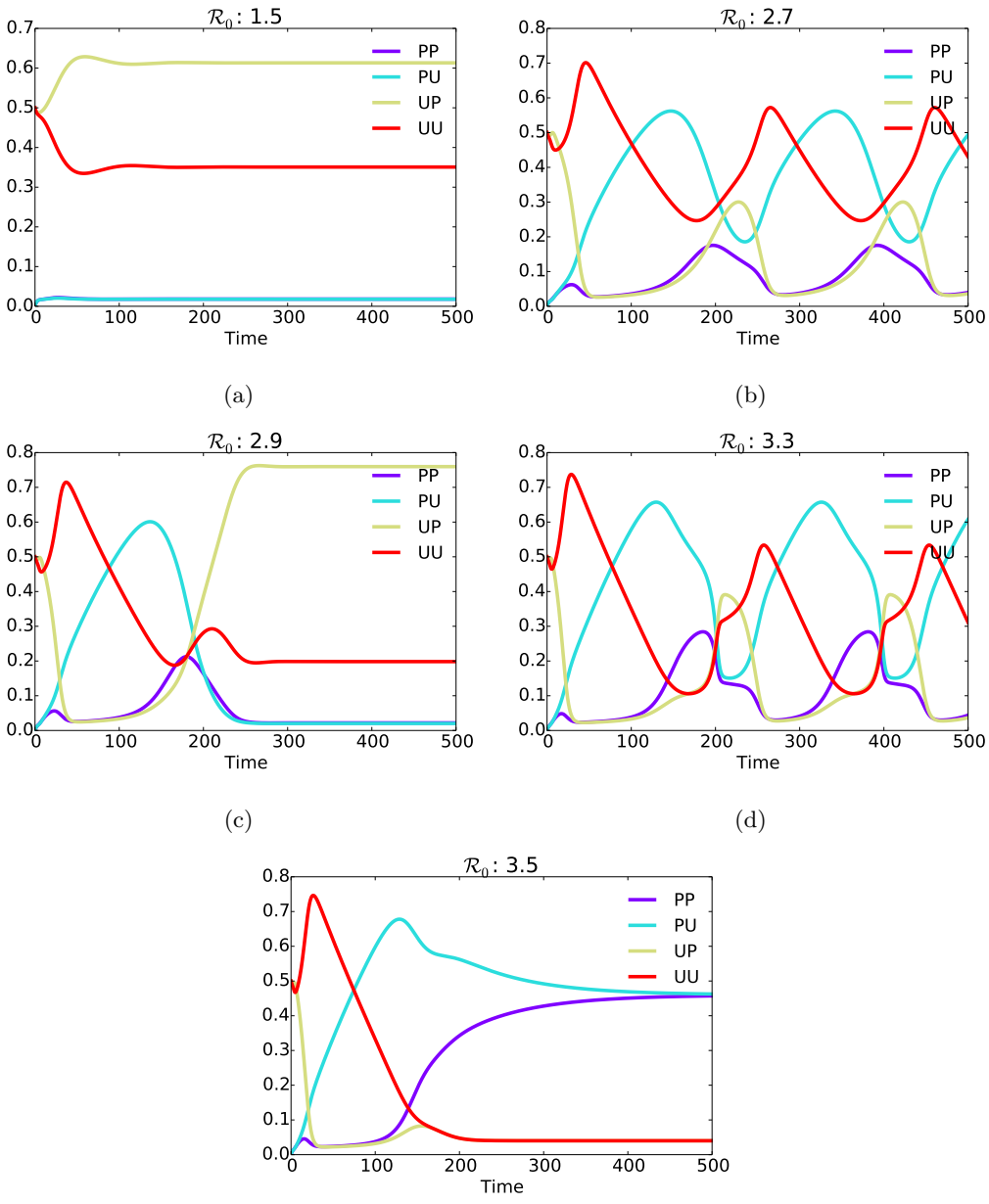
The strategy profile $(\sigma^*_{j_1}, \ldots, \sigma^*_{j_n})$ is a Bayesian Nash equilibrium if for all players and all types, $\sigma^*_{j_i}$ satisfies

$$\sum_{\theta_{-i}} p_i(\theta_{-i}|\theta_i) u_i(\sigma^*_{j_i}(\theta_i), \sigma^*_{j_{-i}}(\theta_{-i}), \theta_i, theta_{-i}) \geq \sum_{\theta_{-i}} p_i(\theta_{-i}|\theta_i) u_i(\sigma_{k_i}(\theta_i), \sigma^*_{j_{-i}}(\theta_{-i}), \theta_i, \theta_{-i}) \tag{A.1}$$

for any $\sigma_{k_i} \neq \sigma^*_{j_i}$. Equivalently

$$\sum_{\theta_i} Pr(\theta_i)[\sum_{\theta_{-i}} p_i(\theta_{-i}|\theta_i) u_i(\sigma^*_{j_i}(\theta_i), \sigma^*_{j_{-i}}(\theta_{-i}), \theta_i, \theta_{-i})] \geq \sum_{\theta_i} Pr(\theta_i)[\sum_{\theta_{-i}} p_i(\theta_{-i}|\theta_i) u_i(\sigma_{k_i}(\theta_i), \sigma^*_{j_{-i}}(\theta_{-i}), \theta_i, \theta_{-i})] \tag{A.2}$$

A.2 Behavioral Trajectories for Simulations in Figure 3.4



(a) (b) (c) (d) (e)

Figure A.1: The distribution of type-contingent strategies as a function of time with $f(UU, 0) = 0.5, \gamma = 0.5$ as in Figure 3.4.

A.3 Endemic Prevalence at $\mathcal{R}_0 < 1$

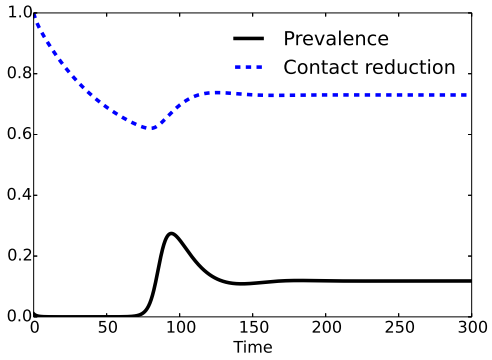


Figure A.2: Prevalence and contact reduction from the combined model for $\beta_b = 2.1, \gamma = 0.5, f(UU, 0) = 0$.

A.4 Alternate Models

A.4.1 SIRS Model

$$\begin{aligned} \dot{S} &= \delta R - \beta SI \\ \dot{I} &= \beta SI - \gamma I \\ \dot{R} &= \gamma I - \delta R \end{aligned} \tag{A.3}$$

A.4.2 Susceptible-only Behavior Change

For the reduced model, we use the fitness functions

$$\begin{aligned} \phi(P_S, t) = \quad & Pr(S)[f(P_S, t)u_1(P_S, P_S, S, S) + f(U_S, t)u_1(P_S, U_S, S, S)] \\ & + Pr(I)[f(P_I)u_1(P_S, P_I, S, I) + f(U_I)u_1(P_S, U_I, S, I)] \\ \phi(U_S, t) = \quad & Pr(S)[f(P_S, t)u_1(U_S, P_S, S, S) + f(U_S, t)u_1(U_S, U_S, S, S)] \\ & + Pr(I)[f(P_I)u_1(U_S, P_I, S, I) + f(U_I)u_1(U_S, U_I, S, I)] \end{aligned} \tag{A.4}$$

where the notation P_S denotes a susceptible player choosing P (similarly for U_S, P_I, U_I) and $f(U_I)$ and $F(P_I)$ are fixed over time. Since only susceptible players

change their strategy, $\mathbf{f}(t) = (f(P_S, t), f(U_S, t))$, the mutation matrix for this game is 2×2 and we can use the two-dimensional system

$$\begin{aligned} \dot{I} &= -\beta_b S_U f(U_I) I + \gamma I \\ \dot{f}(U_S) &= q_{U_S P_S} \phi(P_S, t) f(P_S, t) + q_{U_S, U_S} \phi(U_S, t) f(U_S, t) - \bar{\phi} f(U_S, t) \end{aligned} \tag{A.5}$$

where $S_U = S f(U_S, t)$.

Figure A.3 shows prevalence trajectories for this model at increasing baseline effective contact rates. Figure A.4 shows the best fit trajectories and parameter values for the reduced model compared to the full model.

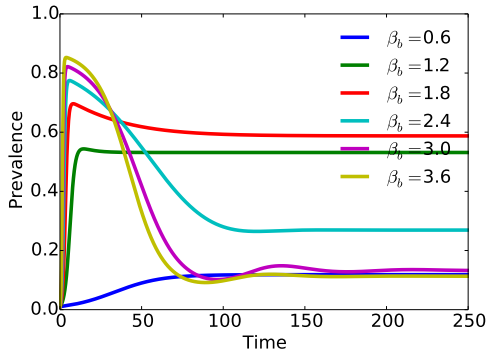


Figure A.3: Prevalence dynamics for the reduced model from Section 3.2 where only susceptible individuals adapt their behavior. For the simulations above $\gamma = 0.5$, $\mathbf{f}(0) = (0.0, 1.0)$.

A.5 Alternate preferences

We examined a case where infected individuals prefer unprotected sex over protected sex regardless of partner type, a phenomenon that has been observed empirically [126]. For susceptible individuals, the type-dependent payoffs from the alternate game are the same as in Figure 3.1. Without loss of generality, for an infected-type player 1 paired with a susceptible-type player 2, the type-dependent payoff matrix is

		Player 2	
		P	U
Player 1	P	b	c
	U	c	a

Figure A.5 shows the long-term behavior of the alternate model for a range of $\mathcal{R}_0$ and s. This model only exhibits one oscillatory region and a higher endemic prevalence than the original model. However, extinction still occurs at high $\mathcal{R}_0$.

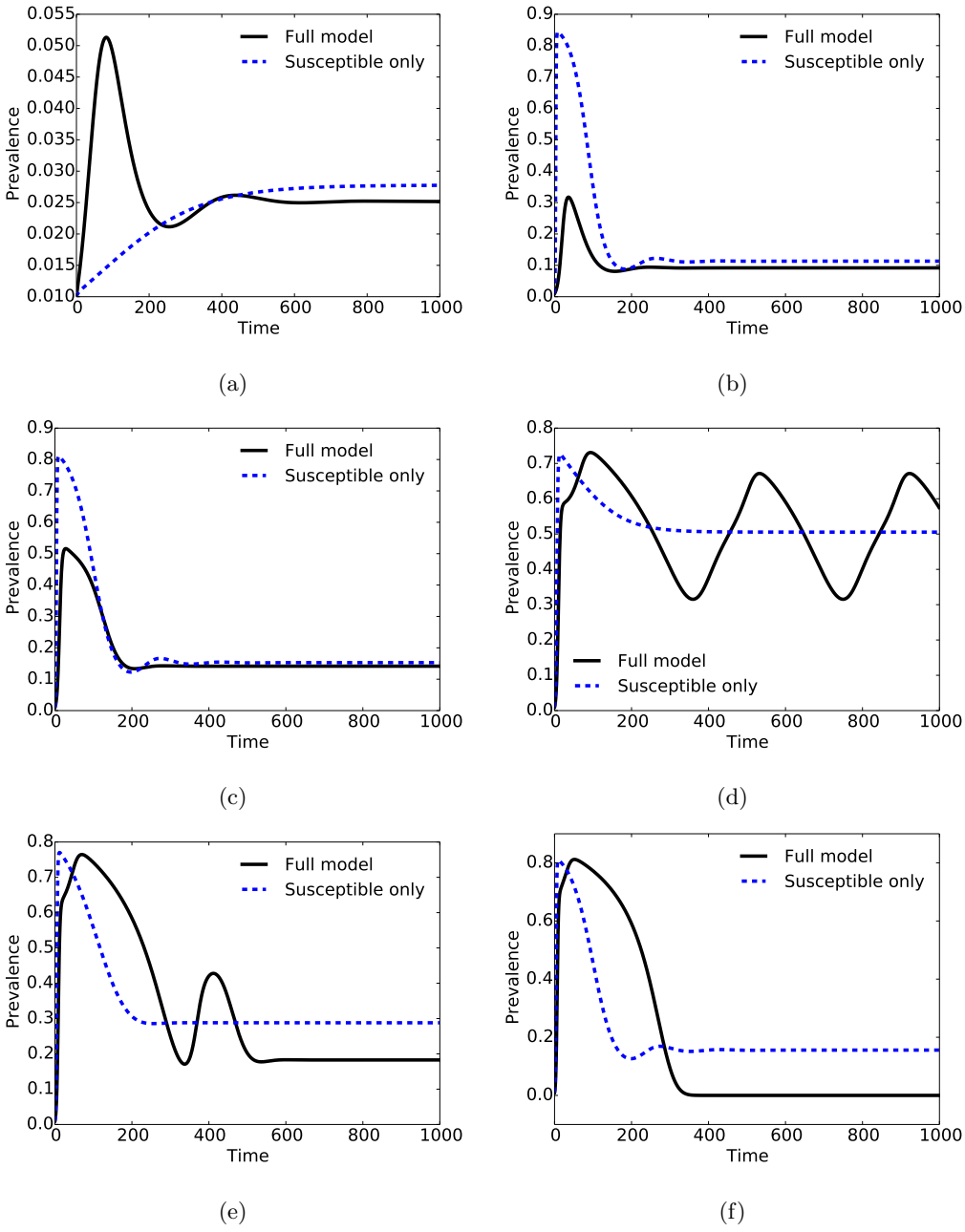
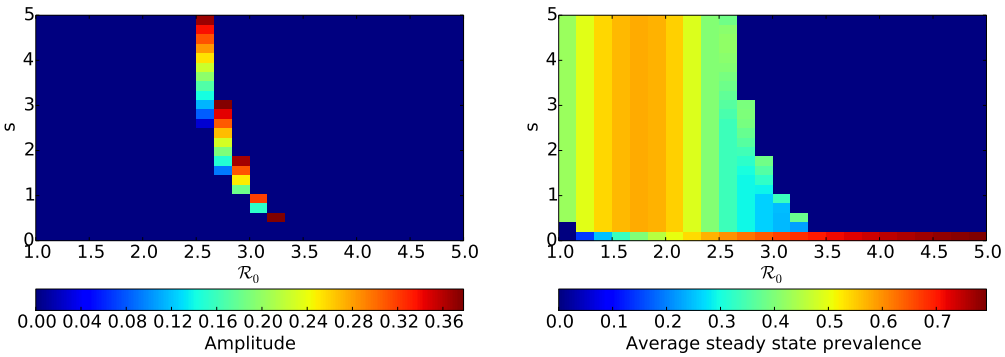


Figure A.4: The reduced model of susceptible-only behavior change fit to simulated trajectories from the full model with $\gamma = 0.5$, (a) Full model $\beta_b = 1.2$, best fit $\beta_b^* = 0.542$, (b) full model $\beta_b = 1.8$, best fit $\beta_b^* = 3.41$, (c) full model $\beta_b = 2.4$, best fit $\beta_b^* = 2.84$, (d) full model $\beta_b = 2.7$, best fit $\beta_b^* = 1.99$, (e) full model $\beta_b = 3$, best fit $\beta_b^* = 2.36$, (f) full model $\beta_b = 3.6$, best fit $\beta_b^* = 2.82$.



(a) (b)

Figure A.5: Long term dynamics of the combined model where infected individuals always prefer unprotected sex for increasing values of $\mathcal{R}_0$ and the behavioral scale parameter s. (a) The amplitude of steady state prevalence oscillations. (b) The average prevalence at steady state.

APPENDIX B

Supplementary material for Chapter IV

B.1 Parameter ranges for LHS

Parameter	Units	Range
β_R	unitless	1.5–5
α	days^{-1}	0.1–0.125
γ_1	days^{-1}	0.14–0.2
γ_2	days^{-1}	0.5–1.0
γ_F	days^{-1}	0.3–1.0
δ_2	unitless	0.8–1.0
σ	days^{-1}	1–10
m	days^{-1}	1×10^{-6}–1×10^{-4}

Table B.1: Ranges used for parameters chosen by LHS.

B.2 Supplemental results

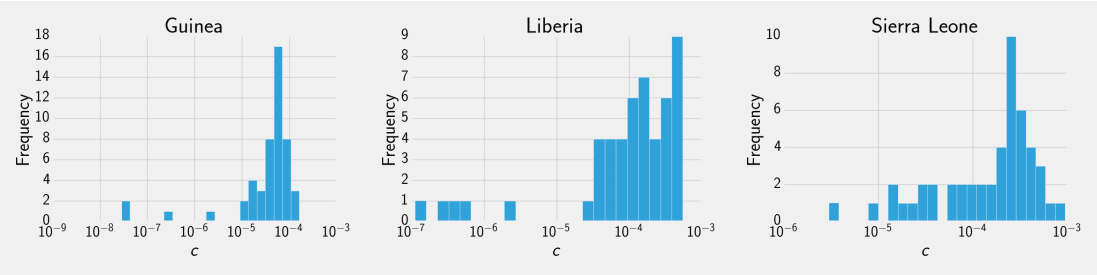
Parameter	Guinea	Liberia	Sierra Leone
β_1	(0.0811,0.166)	(0.100,0.192)	(0.119,0.206)
δ	(0.664,0.668)	(0.295,0.297)	(0.295,0.297)
k	(0.000873,0.00246)	(0.00491,0.00868)	(0.00438,0.00817)
c	$(8.28\times10^{-9}, 1.61\times10^{-4})$	$(1.12\times10^{-7}, 5.49\times10^{-4})$	$(3.01\times10^{-6}, 9.41\times10^{-4})$
σ	(5.94,9.99)	(6.37,9.95)	(5.93,9.90)
m	(4.00,4.88)	(4.00,5.35)	(4.00,5.06)

Table B.2: Parameter intervals within the top 5% of LHS fits for fit parameters (β_1, δ, k, c) and sampled behavioral parameters σ and m.

Table B.2 shows the range of best fit and sampled values for our transmission and behavioral parameters within the top 5% of LHS. The ranges for transmission parameters were relatively tight. The cost parameter c had a wider range – Figure

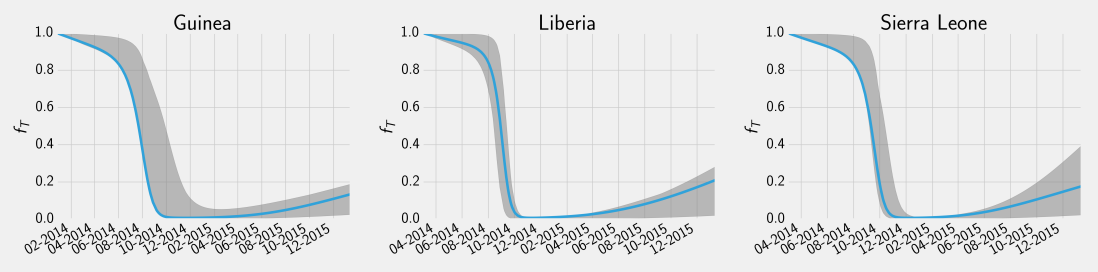
B.1 shows the distribution of best-fit values. Note that the majority of fits fall within approximately one order of magnitude. Even with this range, the top 5% of predicted behavioral dynamics appear relatively well constrained. The range of trajectories (Figure B.2) display the same general features as the best-fit. In addition, the range of cumulative case and death trajectories is tight, suggesting that the uncertainty in fit parameters does not substantially impair our ability to reproduce the outbreak dynamics (Figure B.3).

We simulated the model for multiple values of c about the best-fit for each country (Figure B.4. For all countries, changing c by an order of magnitude alters the speed with which the population reverts to traditional burials after an outbreak as well as the steady state fraction of traditional burials. For Liberia, some values of c result in oscillations in the burial trajectory. This is similar to the oscillations in vaccination and condom use observed in the models of Bauch and Hayashi, respectively [13, 127].



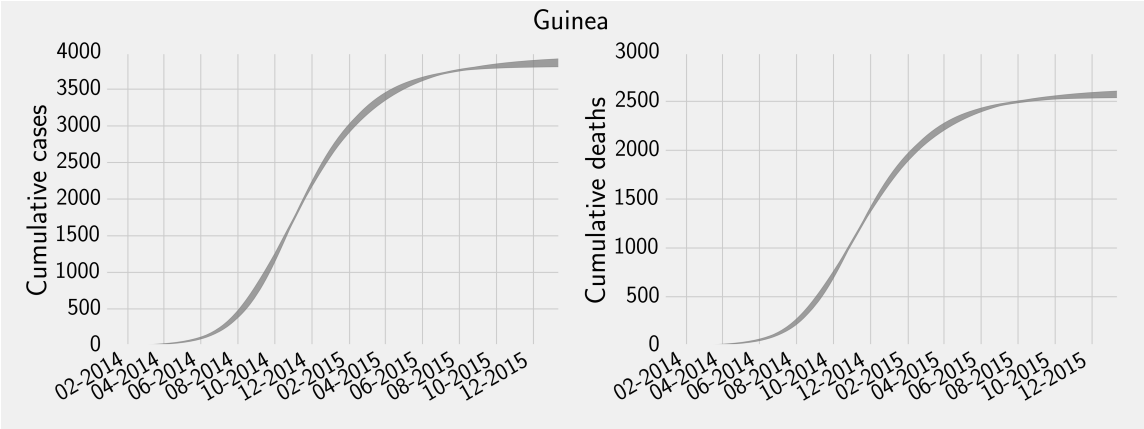
(a) (b) (c)

Figure B.1: Histograms of the best fit values of c for the top 5% of LHS.

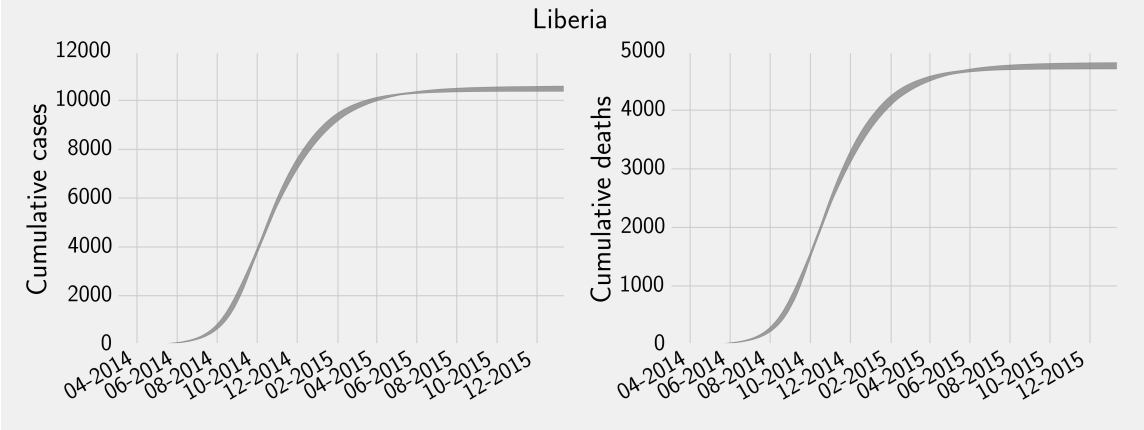


(a) (b) (c)

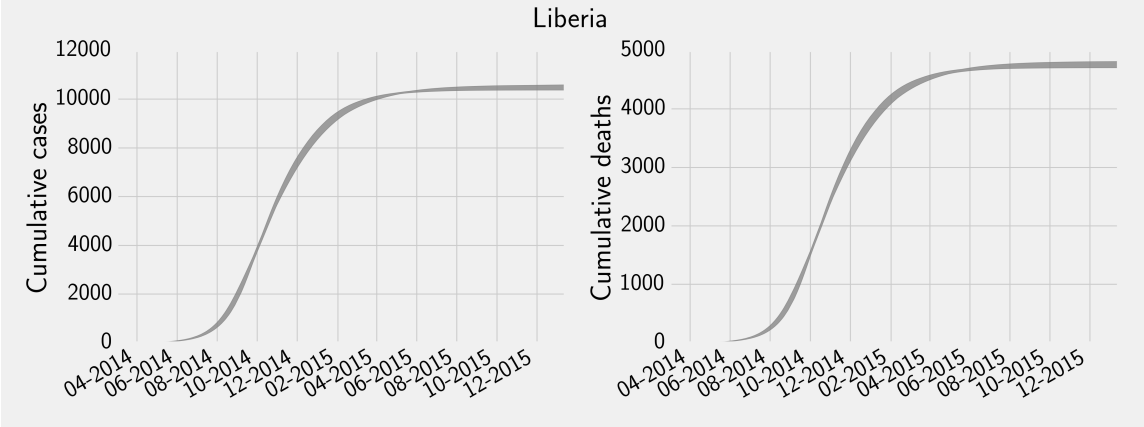
Figure B.2: Shaded areas indicate the range of traditional funeral trajectories for the top 5% of LHS.



(a)

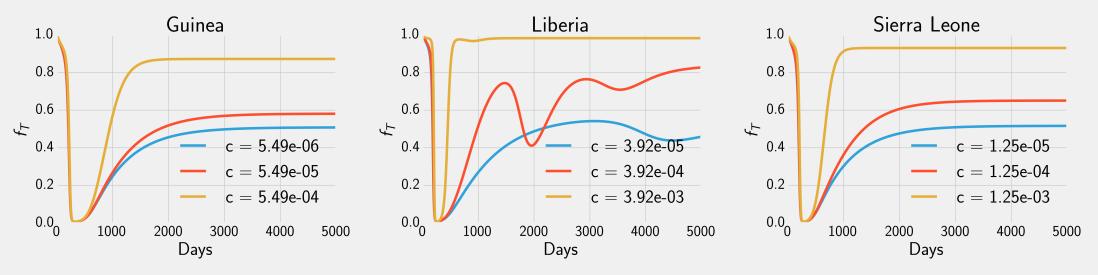


(b)



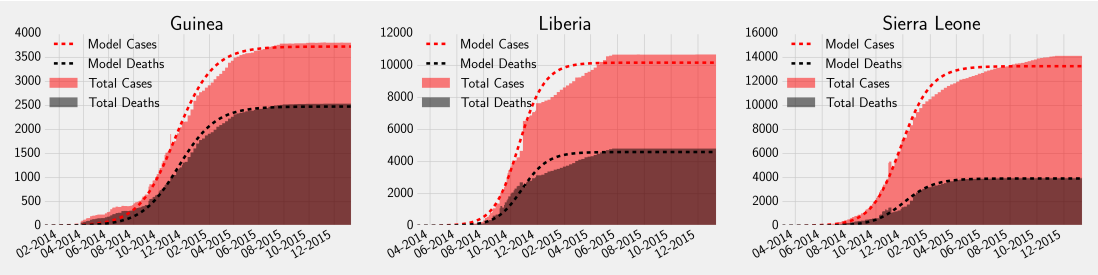
(c)

Figure B.3: Shaded areas indicate the range of case and death trajectories within the bet 5% of LHS for each country.



(a) (b) (c)

Figure B.4: Traditional burial trajectories for varying values of c for each country.



(a) (b) (c)

Figure B.5: Reduced model fit to cumulative cases and deaths in Guinea, Liberia, and Sierra Leone.

APPENDIX C

Supplementary material for Chapter V

C.1 Extended methods

Maximum utility compliance Here we prove that an individual's compliance probability is given by Equation 5.10. We seek to maximize Equation 5.6 with respect to $Pr_i(use)$. The utility function is negative quadratic in this argument, so we need only solve the first order condition

$$\begin{aligned} u_i' &= 2(1 - Pr_i(use) - \Delta x) \\ 0 &= 2(1 - Pr_i(use)^* - \Delta x) \\ Pr_i(use)^* &= 1 - \Delta x \end{aligned} \tag{C.1}$$

as desired.

Optimal recommendations The optimal recommendation for a given preference distribution is

$$\hat{x}^* = argmin_{\hat{x}} E[d_i] \tag{C.2}$$

where

$$E[d] = wv[(1 - E[Pr(use)]) + E[Pr(use)]10^{-\hat{x}}] \tag{C.3}$$

From Equation 5.10, the compliance probability is a function of individual preferences x_i, which are normally distributed. As a result, we can compute the expected value

of $Pr(use)$ (referred to here as $p(x_i)$ for simplicity) using the following

$$E[p(X)] = \int_{-\infty}^{\infty} p(x) f(x|\mu, \sigma) dx \tag{C.4}$$

where $f(x|\mu, \sigma)$ is the probability density function of the preference distribution. While an analytical solution to the minimization problem is not straightforward, numerical optimization and integration perform well. Specifically, we use Brent's algorithm for our numerical solutions. This procedure allows our optimal recommendations to take the shape of the preference distribution into account.

For our local sensitivity analysis, we compute the partial derivative of $\hat{x}^*$ with respect to the variance σ^2 for each average preference tested. We used a baseline variance of 3.6 for our results in Section 5.3.

C.2 Game theoretic representation of optimal recommendations

The process of selecting an optimal recommendation can be framed as an extensive form game between a policymaker and the N individuals in a population. This game has the following structure:

1. A policymaker selects a recommendation $\hat{x}$.
2. Next, all individuals independently determine their compliance level $Pr_i(use)$ given their LRV preference x_i.

Individual payoffs for compliance are given as in Equation 5.6 while the payoff for the policymaker is inversely proportional to the population risk of infection. We use backward induction [49, 50] to solve for the subgame perfect Nash equilibrium of this game. To do so, we determine the optimal decision for each player beginning with the last decision node and proceeding backward to the first decision node. For individuals, the maximum utility compliance level is shown in Equation 5.10. Next,

we determine policymaker's optimal decision, i.e., the recommendation that minimizes the risk of infection given that individuals will play their optimal compliance strategy. The policymaker's decision is then given by Equation 5.12.

C.3 Additional figures

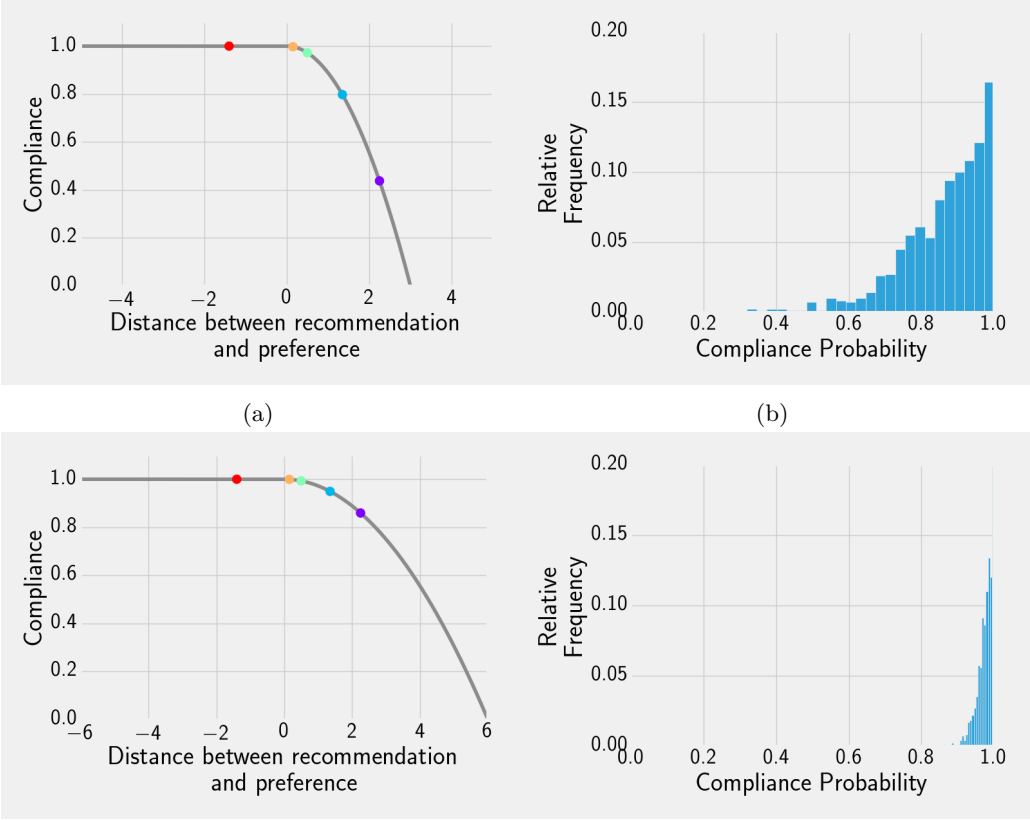
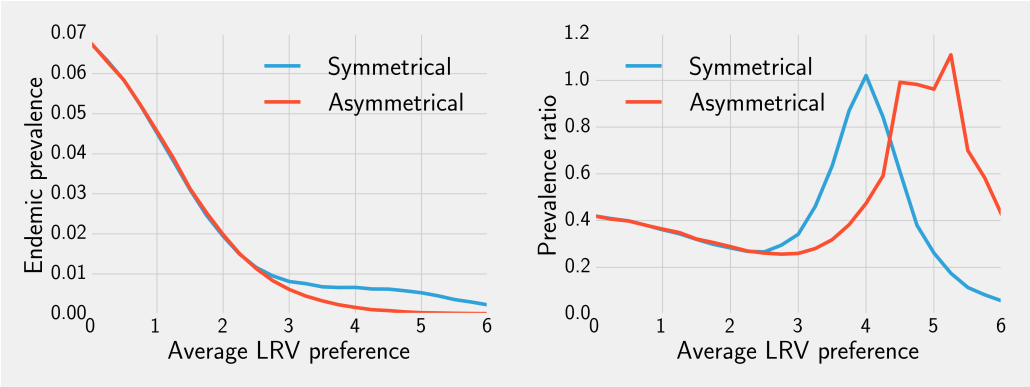
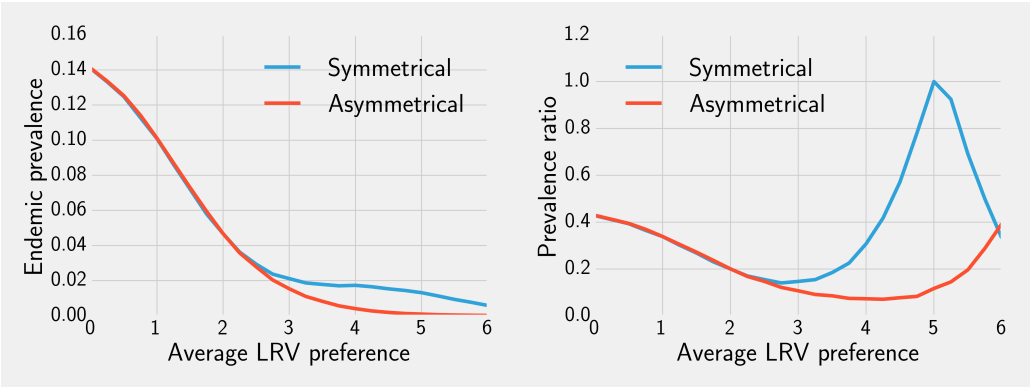


Figure C.1: The effect of the maximum-utility compliance function slope on the compliance distribution. These plots assume the same preference distribution and 3 LRV recommendation as Figure 5.1.



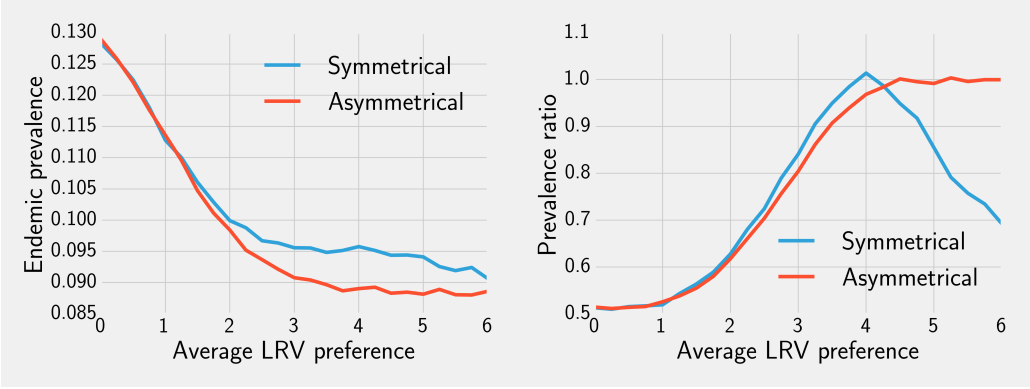
(a) (b)

Figure C.2: C.2a: The simulated endemic prevalence when the optimal LRV is implemented comparing the symmetrical and asymmetrical distance functions. C.2b: The prevalence ratio comparing the optimal recommendation from Figure 5.3 vs. the current 4 LRV recommendation for *E. coli* for both distance functions.



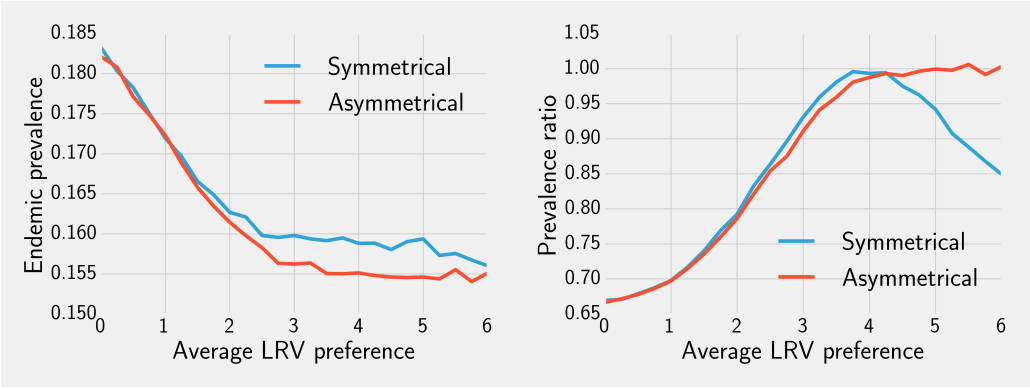
(a) (b)

Figure C.3: C.3a: The simulated endemic prevalence when the optimal LRV is implemented comparing the symmetrical and asymmetrical distance functions. C.3b: The prevalence ratio comparing the optimal recommendation from Figure 5.3 vs. the current 5 LRV recommendation for Rotavirus for both distance functions.



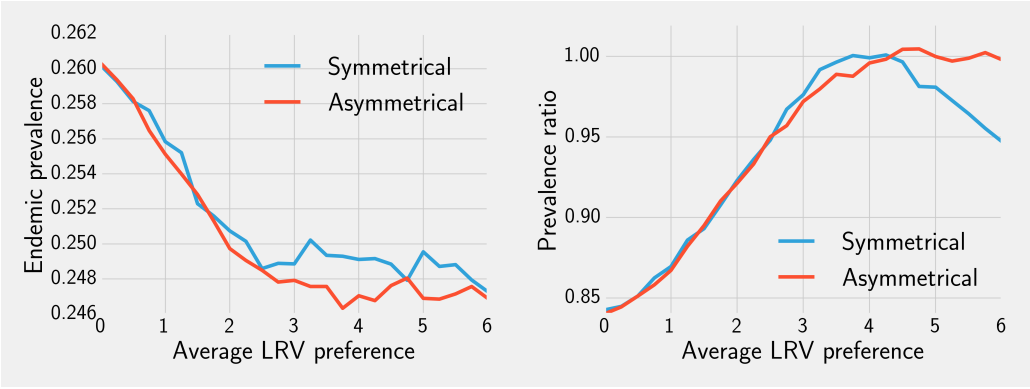
(a)

(b)



(c)

(d)



(e)

(f)

Figure C.4: Endemic prevalence (right) and prevalence ratios (left) for *Cryptosporidium* for three different maximum compliance values: $c_{max} = 0.9$ (top), $c_{max} = 0.8$ (middle), $c_{max} = 0.6$ (bottom) Prevalence ratios compare the optimized recommendation with the current 4 LRV recommendation.

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] David E Sugerman, Albert E Barskey, Maryann G Delea, Ismael R Ortega-Sanchez, Daoling Bi, Kimberly J Ralston, Paul A Rota, Karen Waters-Montijo, and Charles W LeBaron. Measles outbreak in a highly vaccinated population, san diego, 2008: role of the intentionally undervaccinated. *Pediatrics*, 125(4):747–755, 2010.

[2] Vincent AA Jansen, Nico Stollenwerk, Henrik Jeldtoft Jensen, ME Ramsay, WJ Edmunds, and CJ Rhodes. Measles outbreaks in a population with declining vaccine uptake. *Science*, 301(5634):804–804, 2003.

[3] Sonja S Hutchins, Jaime Escolan, Lauri E Markowitz, Carl Hawkins, Anita Kimbler, Richard A Morgan, Stephen R Preblud, and Walter A Orenstein. Measles outbreak among unvaccinated preschool-aged children: opportunities missed by health care providers to administer measles vaccine. *Pediatrics*, 83(3):369–374, 1989.

[4] World Health Organization Media Centre. Sexually transmitted infections. Technical report, World Health Organization, 2013.

[5] National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention. Fact sheet: Incidence, prevalence, and the cost of sexually transmitted infections in the United States. Technical report, Centers for Disease Control & Prevention, 2013.

[6] Centers for Disease Control and Prevention. Sexually transmitted diseases, 2016. URL `https://www.cdc.gov/std/`.

[7] Barbara J Kaplan and Vivian T Shayne. Unsafe sex: decision-making biases and heuristics. *AIDS Education and Prevention*, 1993.

[8] Jeffrey D Fisher and William A Fisher. Changing aids-risk behavior. *Psychological bulletin*, 111(3):455, 1992.

[9] Jill M Dawson, Ray M Fitzpatrick, Gillian Reeves, Mary Boulton, John McLean, Graham J Hart, Marion Brookes, et al. Awareness of sexual partners' HIV status as an influence upon high-risk sexual behaviour. *Aids*, 8(6):837–842, 1994.

[10] Timothy C Reluga, Chris T Bauch, and Alison P Galvani. Evolving public perceptions and stability in vaccine uptake. *Mathematical biosciences*, 204(2):185–198, 2006.

[11] Timothy C Reluga and Alison P Galvani. A general approach for population games with application to vaccination. *Mathematical biosciences*, 230(2):67–78, 2011.

[12] Chris T Bauch and David JD Earn. Vaccination and the theory of games. *Proceedings of the National Academy of Sciences of the United States of America*, 101(36):13391–13394, 2004.

[13] Chris T Bauch. Imitation dynamics predict vaccinating behaviour. *Proceedings of the Royal Society B: Biological Sciences*, 272(1573):1669–1675, 2005.

[14] Frederick H Chen. Rational behavioral response and the transmission of stds. *Theoretical population biology*, 66(4):307–316, 2004.

[15] Kirby D Schroeder and Fabio G Rojas. A game theoretical analysis of sexually transmitted disease epidemics. *Rationality and Society*, 14(3):353–383, 2002.

[16] Eli P Fenichel, Carlos Castillo-Chavez, MG Ceddia, Gerardo Chowell, Paula A Gonzalez Parra, Graham J Hickling, Garth Holloway, Richard Horan, Benjamin Morin, Charles Perrings, et al. Adaptive human behavior in epidemiological models. *Proceedings of the National Academy of Sciences*, 108(15):6306–6311, 2011.

[17] Pierre-Yves Geoffard and Tomas Philipson. Rational epidemics and their public control. *International Economic Review*, pages 603–624, 1996.

[18] Caitlin M Rivers, Eric T Lofgren, Madhav Marathe, Stephen Eubank, and Bryan L Lewis. Modeling the impact of interventions on an epidemic of ebola in sierra leone and liberia. *PLoS currents*, 6, 2014.

[19] Dan Yamin, Shai Gertler, Martial L Ndeffo-Mbah, Laura A Skrip, Mosoka Fallah, Tolbert G Nyenswah, Frederick L Altice, and Alison P Galvani. Effect of ebola progression on transmission and control in liberia. *Annals of internal medicine*, 162(1):11–17, 2015.

[20] WHO Ebola Response Team. Ebola virus disease in west africa—the first 9 months of the epidemic and forward projections. *N Engl J Med*, 2014(371):1481–1495, 2014.

[21] David Fisman, Edwin Khoo, and Ashleigh Tuite. Early epidemic dynamics of the west african 2014 ebola outbreak: estimates derived with a simple two-parameter model. *PLOS currents outbreaks*, 2014.

[22] Declan Butler. Models overestimate ebola cases. *Nature*, 515(7525):18–18, 2014.

[23] Sebastian Funk, Gwenan M Knight, and Vincent AA Jansen. Ebola: the power of behaviour change. *Nature*, 515(7528):492–492, 2014.

[24] M. C. Eisenberg, J. N. S. Eisenberg, J. P. D'Silva, E. V. Wells, S. Cherng, Y.-H. Kao, and R. Meza. Forecasting and Uncertainty in Modeling the 2014-2015 Ebola Epidemic in West Africa. *ArXiv e-prints*, January 2015.

[25] World Health Organization Media Centre. Fact sheet: Diarrhoeal disease. Technical report, World Health Organization, 2013.

[26] United Nations. Water and sanitation - united nations sustainable development goals, 2016. URL `http://www.un.org/sustainabledevelopment/water-and-sanitation/`.

[27] Benjamin Arnold, Byron Arana, Daniel Mäusezahl, Alan Hubbard, and John M Colford. Evaluation of a pre-existing, 3-year household water treatment and handwashing intervention in rural guatemala. *International Journal of Epidemiology*, 38(6):1651–1661, 2009.

[28] CE Stauber, MA Elliott, F Koksal, GM Ortiz, FA DiGiano, and Mark D Sobsey. Characterisation of the biosand filter for e. coli reductions from household drinking water under controlled laboratory and field use conditions. *Water science and technology*, 54(3):1–7, 2006.

[29] Wolf-Peter Schmidt and Sandy Cairncross. Household water treatment in poor populations: is there enough evidence for scaling up now? *Environmental science & technology*, 43(4): 986–992, 2009.

[30] Hugh Waddington and Birte Snilstveit. Effectiveness and sustainability of water, sanitation, and hygiene interventions in combating diarrhoea. *Journal of development effectiveness*, 1(3): 295–335, 2009.

[31] World Health Organization. Guidelines for drinking-water quality. Technical report, World Health Organization, 2011.

[32] Eric Mintz, Jamie Bartram, Peter Lochery, and Martin Wegelin. Not just a drop in the bucket: expanding access to point-of-use water treatment systems. *American Journal of Public Health*, 91(10):1565–1570, 2001.

[33] D Van Halem, H Van der Laan, SGJ Heijman, JC Van Dijk, and GL Amy. Assessing the sustainability of the silver-impregnated ceramic pot filter for low-cost household drinking water treatment. *Physics and Chemistry of the Earth, Parts A/B/C*, 34(1):36–42, 2009.

[34] HM Murphy, M Sampson, E McBean, and K Farahbakhsh. Influence of household practices on the performance of clay pot water filters in rural cambodia. *Desalination*, 248(1):562–569, 2009.

[35] Thomas F Clasen, Joseph Brown, and Simon M Collin. Preventing diarrhoea with household ceramic water filters: assessment of a pilot project in bolivia. *International journal of environmental health research*, 16(03):231–239, 2006.

[36] William O Kermack and Anderson G McKendrick. A contribution to the mathematical theory of epidemics. 115(772):700–721, 1927.

[37] Lawrence F Shampine. *Numerical solution of ordinary differential equations*, volume 4. CRC Press, 1994.

[38] Roy M Anderson and Robert McCredie May. *Infectious diseases of humans*, volume 1. Oxford university press Oxford, 1991.

[39] Klaus Dietz. The estimation of the basic reproduction number for infectious diseases. *Statistical methods in medical research*, 2(1):23–41, 1993.

[40] Odo Diekmann, Johan Andre Peter Heesterbeek, and Johan AJ Metz. On the definition and the computation of the basic reproduction ratio r 0 in models for infectious diseases in heterogeneous populations. *Journal of mathematical biology*, 28(4):365–382, 1990.

[41] John Von Neumann and Oskar Morgenstern. *Theory of games and economic behavior*. Princeton university press, 2007.

[42] John Maynard Smith. *Evolution and the Theory of Games*. Cambridge university press, 1982.

[43] Timothy C Reluga. Game theory of social distancing in response to an epidemic. *PLoS computational biology*, 6(5):e1000793, 2010.

[44] Alison P Galvani, Timothy C Reluga, and Gretchen 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(13):5692–5697, 2007.

[45] Daniel Kahneman. New challenges to the rationality assumption. *Journal of Institutional and Theoretical Economics (JITE)/Zeitschrift für die gesamte Staatswissenschaft*, pages 18–36, 1994.

[46] Max H Bazerman. Negotiator judgment a critical look at the rationality assumption. *American Behavioral Scientist*, 27(2):211–228, 1983.

[47] Reinhard Selten. Bounded rationality. *Journal of Institutional and Theoretical Economics (JITE)/Zeitschrift für die gesamte Staatswissenschaft*, pages 649–658, 1990.

[48] John Nash. Non-cooperative games. *Annals of mathematics*, pages 286–295, 1951.

[49] Drew Fudenberg and Jean Tirole. *Game theory*. MIT Press, 1991.

[50] Steven Tadelis. *Game theory an introduction*. Princeton university press, 2013.

[51] John C Harsanyi. Games with incomplete information played by "bayesian" players, i-iii part i. the basic model. *Management science*, 14(3):159–182, 1967.

[52] J Maynard Smith and GR Price. lhe logic of animal conflict. *Nature*, 246:15, 1973.

[53] Josef Hofbauer and Karl Sigmund. Evolutionary game dynamics. *Bulletin of the American Mathematical Society*, 40(4):479–519, 2003.

[54] Martin A Nowak. *Evolutionary dynamics*. Harvard University Press, 2006.

[55] Ross Cressman and Yi Tao. The replicator equation and other game dynamics. *Proceedings of the National Academy of Sciences*, 111(Supplement 3):10810–10817, 2014.

[56] Roy M Anderson and Geoffrey P Garnett. Mathematical models of the transmission and control of sexually transmitted diseases. *Sexually transmitted diseases*, 27(10):636–643, 2000.

[57] Ken TD Eames and Matt J Keeling. Modeling dynamic and network heterogeneities in the spread of sexually transmitted diseases. *Proceedings of the National Academy of Sciences*, 99 (20):13330–13335, 2002.

[58] Amy G Lam, Amy Mak, Patricia D Lindsay, and Stephen T Russell. What really works? an exploratory study of condom negotiation strategies. *AIDS Education and Prevention*, 16(2): 160–171, 2004.

[59] Courtney Peasant, Gilbert R Parra, and Theresa M Okwumabua. Condom negotiation: Findings and future directions. *The Journal of Sex Research*, (ahead-of-print):1–14, 2014.

[60] Don C Des Jarlais, Patricia Friedmann, Holly Hagan, and Samuel R Friedman. The protective effect of AIDS-related behavioral change among injection drug users: a cross-national study. WHO multi-centre study of AIDS and injecting drug use. *American Journal of Public Health*, 86(12):1780–1785, 1996.

[61] Gillian K SteelFisher, Robert J Blendon, Mark M Bekheit, and Keri Lubell. The public's response to the 2009 H1N1 influenza pandemic. *New England Journal of Medicine*, 362(22), 2010.

[62] G James Rubin, Richard Amlôt, Lisa Page, Simon Wessely, et al. Public perceptions, anxiety, and behaviour change in relation to the swine flu outbreak: cross sectional telephone survey. *Bmj*, 339, 2009.

[63] Simon Gregson, Geoffrey P Garnett, Constance A Nyamukapa, Timothy B Hallett, James JC Lewis, Peter R Mason, Stephen K Chandiwana, and Roy M Anderson. HIV decline associated with behavior change in eastern Zimbabwe. *Science*, 311(5761):664–666, 2006.

[64] Richard J Wolitski, Caroline J Bailey, Ann O'Leary, Cynthia A Gómez, and Jeffrey T Parsons. Self-perceived responsibility of HIV-seropositive men who have sex with men for preventing hiv transmission. *AIDS and Behavior*, 7(4):363–372, 2003.

[65] Ann O'Leary. Guessing games: Sex partner serostatus assumptions among hiv-positive gay and bisexual men. 2005.

[66] Caitlin E Kennedy, Amy M Medley, Michael D Sweat, and Kevin R O'Reilly. Behavioural interventions for HIV positive prevention in developing countries: a systematic review and meta-analysis. *Bulletin of the World Health Organization*, 88(8):615–623, 2010.

[67] Christine J De Rosa and Gary Marks. Preventive counseling of HIV-positive men and self-disclosure of serostatus to sex partners: New opportunities for prevention. *Health Psychology*, 17(3):224, 1998.

[68] Romulus Breban, Raffaele Vardavas, and Sally Blower. Mean-field analysis of an inductive reasoning game: application to influenza vaccination. *Physical review E*, 76(3):031127, 2007.

[69] RM Anderson. Transmission dynamics and control of infectious disease agents. In *Population biology of infectious diseases*, pages 149–176. Springer, 1982.

[70] James A Yorke, Herbert W Hethcote, and Annett Nold. Dynamics and control of the transmission of gonorrhea. *Sexually transmitted diseases*, 5(2):51–56, 1978.

[71] Ron S Gold and Michael J Skinner. Judging a book by its cover: gay men's use of perceptible characteristics to infer antibody status. *International journal of STD & AIDS*, 7(1):39–43, 1996.

[72] Stefan Z Wiktor, Robert J Biggar, Mads Melbye, Peter Ebbesen, Ginga Colclough, Richard DiGioia, William C Sanchez, Ronald J Grossman, and James J Goedert. Effect of knowledge of human immunodeficiency virus infection status on sexual activity among homosexual men. *JAIDS Journal of Acquired Immune Deficiency Syndromes*, 3(1):62–68, 1990.

[73] Karen M Page and Martin A Nowak. Unifying evolutionary dynamics. *Journal of theoretical biology*, 219(1):93–98, 2002.

[74] JM Heffernan, RJ Smith, and LM Wahl. Perspectives on the basic reproductive ratio. *Journal of the Royal Society Interface*, 2(4):281–293, 2005.

[75] World Health Organization et al. Ebola virus disease fact sheet, 2016. URL `http://www.who.int/mediacentre/factsheets/fs103/en/`.

[76] Centers for Disease Control and Prevention. Ebola (ebola virus disease), 2016. URL `http://www.cdc.gov/vhf/ebola/`.

[77] KA Alexander, CE Sanderson, M Marathe, BL Lewis, CM Rivers, J Shaman, JM Drake, E Lofgren, VM Dato, MC Eisenberg, et al. What factors might have led to the emergence of ebola in west africa. *PLOS Neglected Tropical Diseases*, 2014.

[78] Silvia Dallatomasina, Rosa Crestani, James Sylvester Squire, Hilde Declerk, Grazia Marta Caleo, Anja Wolz, Kathryn Stinson, Gabriela Patten, Raphael Brechard, Osman Bamba-Moi Gbabai, et al. Ebola outbreak in rural west africa: epidemiology, clinical features and outcomes. *Tropical Medicine & International Health*, 20(4):448–454, 2015.

[79] Joseph A Lewnard, Martial L Ndeffo Mbah, Jorge A Alfaro-Murillo, Frederick L Altice, Luke Bawo, Tolbert G Nyenswah, and Alison P Galvani. Dynamics and control of ebola virus transmission in montserrado, liberia: a mathematical modelling analysis. *The Lancet Infectious Diseases*, 14(12):1189–1195, 2014.

[80] Caitlin Rivers. Ebola: models do more than forecast. *Nature*, 515(7528):492–492, 2014.

[81] Paul Richards, Joseph Amara, Mariane C Ferme, Prince Kamara, Esther Mokuwa, Amara Idara Sheriff, Roland Suluku, and Maarten Voors. Social pathways for ebola virus disease in rural sierra leone, and some implications for containment. *PLoS Negl Trop Dis*, 9 (4):e0003567, 2015.

[82] Joseph Prescott, Trenton Bushmaker, Robert Fischer, Kerri Miazgowicz, Seth Judson, and Vincent J Munster. Postmortem stability of ebola virus. *Emerging infectious diseases*, 21(5): 856, 2015.

[83] Kerton R Victory, Fátima Coronado, Sâa O Ifono, Therese Soropogui, and Benjamin A Dahl. Ebola transmission linked to a single traditional funeral ceremony– kissidougou, guinea, december, 2014–january 2015. *MMWR. Morbidity and mortality weekly report*, 64(14):386–388, 2015.

[84] Carrie F Nielsen, Sarah Kidd, AR Sillah, Edward Davis, Jonathan Mermin, Peter H Kilmarx, et al. Improving burial practices and cemetery management during an ebola virus disease epidemic—sierra leone, 2014. *MMWR Morb Mortal Wkly Rep*, 64(1):20–27, 2015.

[85] World Health Organization. Field situation: How to conduct safe and dignified burial of a patient who has died from suspected or confirmed ebola virus disease. Technical report, World Health Organization, 2014.

[86] WHO Ebola Response Team. West african ebola epidemic after one year—slowing but not yet under control. *N Engl J Med*, 2015(372):584–587, 2015.

[87] Abhishek Pandey, Katherine E Atkins, Jan Medlock, Natasha Wenzel, Jeffrey P Townsend, James E Childs, Tolbert G Nyenswah, Martial L Ndeffo-Mbah, and Alison P Galvani. Strategies for containing ebola in west africa. *Science*, 346(6212):991–995, 2014.

[88] Paolo Francesconi, Zabulon Yoti, Silvia Declich, Paul Awil Onek, Massimo Fabiani, Joseph Olango, Roberta Andraghetti, Pierre E Rollin, Cyprian Opira, Donato Greco, et al. Ebola hemorrhagic fever transmission and risk factors of contacts, uganda. *Emerging infectious diseases*, 9(11):1430–1437, 2003.

[89] Scott F Dowell, Rose Mukunu, Thomas G Ksiazek, Ali S Khan, Pierre E Rollin, and CJ Peters. Transmission of ebola hemorrhagic fever: a study of risk factors in family members, kikwit, democratic republic of the congo, 1995. *Journal of Infectious Diseases*, 179(Supplement 1): S87–S91, 1999.

[90] Roger Ndambi, Philippe Akamituna, Marie-Jo Bonnet, Anicet Mazaya Tukadila, Jean-Jacques Muyembe-Tamfum, and Robert Colebunders. Epidemiologic and clinical aspects of the ebola virus epidemic in mosango, democratic republic of the congo, 1995. *Journal of Infectious Diseases*, 179(Supplement 1):S8–S10, 1999.

[91] Central Intelligence Agency. The world factbook, 2016. URL `https://www.cia.gov/library/publications/resources/the-world-factbook/`.

[92] World Health Organization. Origins of the 2014 ebola epidemic, 2015. URL `http://www.who.int/csr/disease/ebola/one-year-report/virus-origin/en/`.

[93] World Health Organization. Ebola situation reports, 2016. URL `http://apps.who.int/ebola/ebola-situation-reports`.

[94] Aaron A King, Matthieu Domenech de Cellès, Felicia MG Magpantay, and Pejman Rohani. Avoidable errors in the modelling of outbreaks of emerging pathogens, with special reference to ebola. In *Proc. R. Soc. B*, volume 282, page 20150347. The Royal Society, 2015.

[95] DM Hamby. A review of techniques for parameter sensitivity analysis of environmental models. *Environmental monitoring and assessment*, 32(2):135–154, 1994.

[96] Abba B Gumel and Suzanne Lenhart. *Modeling paradigms and analysis of disease transmission models*, volume 75. American Mathematical Soc., 2010.

[97] Christian L Althaus. Estimating the reproduction number of ebola virus (ebov) during the 2014 outbreak in west africa. *PLOS Currents Outbreaks*, 2014.

[98] World Health Organization. Millennium development goal drinking water target met, 2012. URL `http://www.who.int/mediacentre/news/releases/2012/drinking_water_20120306/en/`.

[99] Mark D Sobsey, Christine E Stauber, Lisa M Casanova, Joseph M Brown, and Mark A Elliott. Point of use household drinking water filtration: a practical, effective solution for providing sustained access to safe drinking water in the developing world. *Environmental science & technology*, 42(12):4261–4267, 2008.

[100] Benjamin F Arnold and John M Colford. Treating water with chlorine at point-of-use to improve water quality and reduce child diarrhea in developing countries: a systematic review and meta-analysis. *The American journal of tropical medicine and hygiene*, 76(2):354–364, 2007.

[101] Everett M Rogers. *Diffusion of innovations*. Simon and Schuster, 2010.

[102] Ghislaine Rosa and Thomas Clasen. Estimating the scope of household water treatment in low-and medium-income countries. *The American journal of tropical medicine and hygiene*, 82(2):289–300, 2010.

[103] Kyle S Enger, Kara L Nelson, Thomas Clasen, Joan B Rose, and Joseph NS Eisenberg. Linking quantitative microbial risk assessment and epidemiological data: informing safe drinking water trials in developing countries. *Environmental science & technology*, 46(9):5160–5167, 2012.

[104] Kyle S Enger, Kara L Nelson, Joan B Rose, and Joseph NS Eisenberg. The joint effects of efficacy and compliance: A study of household water treatment effectiveness against childhood diarrhea. *Water research*, 47(3):1181–1190, 2013.

[105] Joe Brown and Thomas Clasen. High adherence is necessary to realize health gains from water quality interventions. *PLoS One*, 7(5):e36735, 2012.

[106] AJ Fischer, A Threlfall, S Meah, R Cookson, H Rutter, and MP Kelly. The appraisal of public health interventions: an overview. *Journal of Public Health*, page fdt076, 2013.

[107] Barbara J McNeil and Stephen G Pauker. Decision analysis for public health: principles and illustrations. *Annual Review of Public Health*, 5(1):135–161, 1984.

[108] Michael J Messner, Cynthia L Chappell, and Pablo C Okhuysen. Risk assessment for cryptosporidium: a hierarchical bayesian analysis of human dose response data. *Water Research*, 35(16):3934–3940, 2001.

[109] Herbert L DuPont, Samuel B Formal, Richard B Hornick, Merrill J Snyder, Joseph P Libonati, Daniel G Sheahan, Eugene H LaBrec, and John P Kalas. Pathogenesis of escherichia coli diarrhea. *New England Journal of Medicine*, 285(1):1–9, 1971.

[110] Richard L Ward, David I Bernstein, Elizabeth C Young, James R Sherwood, Douglas R Knowlton, and Gilbert M Schiff. Human rotavirus studies in volunteers: determination of infectious dose and serological response to infection. *Journal of Infectious Diseases*, 154(5): 871–880, 1986.

[111] MA Elliott, CE Stauber, F Koksal, FA DiGiano, and MD Sobsey. Reductions of e. coli, echovirus type 12 and bacteriophages in an intermittently operated household-scale slow sand filter. *Water research*, 42(10):2662–2670, 2008.

[112] Vinka A Oyanedel-Craver and James A Smith. Sustainable colloidal-silver-impregnated ceramic filter for point-of-use water treatment. *Environmental science & technology*, 42(3): 927–933, 2007.

[113] Joseph NS Eisenberg, Edmund YW Seto, John M Colford, Adam Olivieri, Robert C Spear, et al. An analysis of the milwaukee cryptosporidiosis outbreak based on a dynamic model of the infection process. *Epidemiology*, 9(3):255–263, 1998.

[114] Teresa Estrada-Garcia, Catalina Lopez-Saucedo, Rocio Thompson-Bonilla, Maricela Abonce, Daniel Lopez-Hernandez, Jose Ignacio Santos, Jorge L Rosado, Herbert L DuPont, and Kurt Z Long. Association of diarrheagenic escherichia coli pathotypes with infection and diarrhea among mexican children and association of atypical enteropathogenic e. coli with acute diarrhea. *Journal of clinical microbiology*, 47(1):93–98, 2009.

[115] Marc Gurwith, Wanda Wenman, Dorothy Hinde, Sheila Feltham, and Harry Greenberg. A prospective study of rotavirus infection in infants and young children. *Journal of infectious diseases*, 144(3):218–224, 1981.

[116] M Sobsey and J Brown. Evaluating household water treatment options: Health-based targets and microbiological performance specifications. *World Health Organization: Geneva, Switzerland*, 2011.

[117] AH Havelaar and JM Melse. Quantifying public health risk in the who guidelines for drinking-water quality: A burden of disease approach. 2003.

[118] MM Fiore, K Minnings, LD Fiore, et al. Assessment of biosand filter performance in rural communities in southern coastal nicaragua: an evaluation of 199 households. *Rural Remote Health*, 10(3):1483, 2010.

[119] Sundeep K Gupta, MS Islam, Richard Johnston, Pavani Kalluri Ram, and Stephen P Luby. The chulli water purifier: acceptability and effectiveness of an innovative strategy for household water treatment in bangladesh. *The American journal of tropical medicine and hygiene*, 78(6):979–984, 2008.

[120] Anwar Huq, Mohammed Yunus, Syed Salahuddin Sohel, Abbas Bhuiya, Michael Emch, Stephen P Luby, Estelle Russek-Cohen, G Balakrish Nair, R Bradley Sack, and Rita R Colwell. Simple sari cloth filtration of water is sustainable and continues to protect villagers from cholera in matlab, bangladesh. *MBio*, 1(1):e00034–10, 2010.

[121] Rita R Colwell, Anwar Huq, M Sirajul Islam, KMA Aziz, Mohammed Yunus, N Huda Khan, A Mahmud, R Bradley Sack, G Balakrish Nair, J Chakraborty, et al. Reduction of cholera in bangladeshi villages by simple filtration. *Proceedings of the National Academy of Sciences*, 100(3):1051–1055, 2003.

[122] Sheng Li, Joseph NS Eisenberg, Ian H Spicknall, and James S Koopman. Dynamics and control of infections transmitted from person to person through the environment. *American journal of epidemiology*, 170(2):257–265, 2009.

[123] Joseph H Tien and David JD Earn. Multiple transmission pathways and disease dynamics in a waterborne pathogen model. *Bulletin of mathematical biology*, 72(6):1506–1533, 2010.

[124] Sonia Altizer, Andrew Dobson, Parviez Hosseini, Peter Hudson, Mercedes Pascual, and Pejman Rohani. Seasonality and the dynamics of infectious diseases. *Ecology letters*, 9(4): 467–484, 2006.

[125] CE Cornelius 3rd. Seasonality of gonorrhea in the united states. *HSMHA health reports*, 86 (2):157, 1971.

[126] Neil S Wenger, Francoise S Kusseling, Keith Beck, and Martin F Shapiro. Sexual behavior of individuals infected with the human immunodeficiency virus: the need for intervention. *Archives of Internal Medicine*, 154(16):1849–1854, 1994.

[127] Michael AL Hayashi and Marisa C Eisenberg. Effects of adaptive protective behavior on the dynamics of sexually transmitted infections. *Journal of theoretical biology*, 388:119–130, 2016.