HIV Epidemic Control among Sex Workers and their Clients

by

Nicola Mulberry

B.Sc., University of British Columbia, 2016

Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of
Master of Science

in the
Department of Mathematics
Faculty of Science

SIMON FRASER UNIVERSITY
Fall 2018

Approval

Name:	**Nicola Mulberry**
Degree:	**Master of Science (Applied and Computational Mathematics)**
Title:	**HIV Epidemic Control among Sex Workers and their Clients**
Examining Committee:	**Chair:** John Stockie Professor
	Ralf Wittenberg Senior Supervisor Associate Professor
	Alexander Rutherford Co-Supervisor Adjunct Professor
	Marni Mishna Supervisor Professor
	Caroline Colijn Internal Examiner Professor
Date Defended:	**December 6, 2018**

Abstract

Controlling the spread of HIV among hidden, high-risk populations such as sex workers and their clients is becoming increasingly important in the fight to end AIDS. In this thesis, we identify a number of sociological and structural factors which render general control strategies ineffective among these key populations, and instead call for focused testing and interventions. A bipartite network model of sexual contacts between female sex workers and male clients is motivated using historical data from a South African mining community. HIV transmission and progression is modelled as a stochastic process on the network, and the effect of various intervention strategies on HIV prevalence in the population is determined through numerical simulations. We find that preventative interventions are highly cost-effective when targeted at female sex workers. For aggressive reduction in HIV prevalence, however, the client population cannot be ignored and treatment of both populations is necessary.

Keywords: bipartite networks; HIV; epidemiology; control strategies; sex workers

Acknowledgements

The data from the Carletonville-Mosuthumpilo project was provided by Brian Williams. I would further like to thank Brian for insightful comments regarding this data and on the public health implications of this research.

The simulations performed in this thesis relied on the software NepidemiX developed by Lukas Ahrenberg, who was extremely helpful with any technical issues which arose.

This research was enabled in part by computational facilities provided by WestGrid (www.westgrid.ca) and Compute Canada Calcul Canada (www.computecanada.ca).

I am also grateful to Marni Mishna for insightful discussions and feedback, and to my supervisors Ralf Wittenberg and Sandy Rutherford for their guidance and support.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

The global HIV burden is largely confined to key populations such as men who have sex with men, injection drug users, and sex workers. Even in the generalized epidemic in sub-Saharan Africa, female sex workers (FSWs) are disproportionately affected [32]. In many communities—especially those in which sex work is criminalized—sex workers face widespread human rights violations and high levels of sexual violence, which may lead to an increase in HIV susceptibility [4]. This is coupled with potential barriers to access of treatment and prevention options for HIV [16]. Further contributions to the wider HIV epidemic come from the fact that clients of FSWs typically act as bridges to the larger community, with many having one or more regular sexual partners outside of transactional sex [36].

HIV transmission is associated with complex social interactions such as sexual relationships or needle sharing. Therefore, knowledge of the underlying social network is important to understanding disease spread and for designing control strategies. The use of networks to model epidemics is well-established [47]. Various authors have investigated immunization strategies on networks [15, 46, 12, 52, 11]. These works investigate the most effective distribution of vaccines within a single population. In particular, Chen and Lu [11] describe an efficient and feasible immunization strategy for FSWs.

While a vaccine is not currently available, HIV can be controlled with both treatment and preventative options. For infected (HIV-positive) individuals, antiretroviral therapy (ART) increases life expectancy and has been shown to significantly reduce transmission. ART is a combination antiretroviral drug which may lead to viral suppression with consistent use. Pre-exposure prophylaxis (PrEP) is another combination antiretroviral drug, but is taken by high-risk HIV-negative individuals to reduce the likelihood of infection. Condoms are an inexpensive prevention measure against HIV and other sexually transmitted infections. Simulation studies have proposed that condom distribution programs alone can significantly reduce HIV incidence among FSWs and clients [6]. However, FSWs frequently cite instances of forced unprotected sex, and it has been proposed that PrEP may be particularly beneficial in the FSW community [20]. Additional control strategies targeted at

commercial sex workers include creating safer work environments, improving access to care, community empowerment and decriminalization [7].

Previous HIV models have considered the effect of ART and, more recently, PrEP on HIV prevalence in sub-Saharan Africa and elsewhere. The review by Gomez *et al.* [23] explored deterministic and stochastic compartmental models to evaluate the potential impact of a large scale PrEP roll-out alongside existing HIV prevention programs, such as condoms and ART. This review determined that targeting PrEP to key populations would be a cost-effective strategy. Similarly, Pretorius *et al.* [49] concluded that PrEP targeted at 25–35 year old women would have a positive impact on the South African epidemic. Vissers *et al.* [55] developed a compartmental model which captures the high-risk populations of sex workers and their clients, as well as the low-risk general population in India and Africa. They found that PrEP would have a substantial impact in many African settings when targeted at both FSWs and clients, because condom use is frequently low in such settings.

Much of the sex work in sub-Saharan Africa is "informal", and could be described as transactional or survival sex work [53]. These terms encompass a variety of sex-for-goods or -money exchanges, and typically exclude women working as escorts or in brothels. Survival sex work is not unique to sub-Saharan Africa, and this term may encompass certain street-based sex workers [54]. Women who engage in survival sex work are at particular risk of exploitation and have limited ability to advocate for their health and welfare [61]. We will consider the sexual contacts described in this work to be of this informal nature. In this thesis, we consider the implications of the nature of survival sex work on the structure of the sexual contact network and hence on our model.

Outline of Thesis

There are three main components of our model: the disease process, which describes how HIV progresses in a host; the contact process, which models how HIV spreads through the population; and our numerical simulations. In the remainder of this introductory chapter, we discuss the relevant background information on HIV epidemiology and control. This information will be important when we discuss the model disease process, and for identifying the key assumptions made. This is followed by an introduction to social network analysis which leads to our work in Chapter 3, in which key characteristics of sex worker–client networks are identified that can be incorporated into our model of sexual contacts. In the last part of Chapter 1, we review the mathematical study of random graphs, which provides an analytical framework for describing and generating networks. This review directly leads to the discussion regarding simple dynamics on networks in Chapter 2. That chapter provides an example of analytical results for epidemics on networks, and furthermore allows us to test our numerical method for performing large-scale simulations. Using all of the above ideas and methodologies, we finally construct our model and evaluate various HIV control strategies in Chapter 4.

Original Contributions

The original contributions of this thesis are: (1) Our model for sex worker–client interactions, and (2) our conclusions regarding HIV control among these high-risk individuals. Our model for the sexual contact network is a bipartite network model. Such models have been used in previous studies on disease spread; for example, Bisanzio *et al.* [8] use a bipartite network to model vector-borne diseases, and Meyers *et al.* [39] use a directed bipartite network to investigate the spread of pneumonia from patients to hospital workers. In Gomez *et al.* [24], the spread of generic sexually transmitted infections—including HIV—on heterosexual contact networks is considered. The authors use mean-field methods to analyze SIR dynamics on scale-free bipartite networks, which is not a focus of this work. In this thesis, we use both sociological literature and our analysis of historical data to construct a model for sex worker–client interactions. To our knowledge, this is the first work which investigates the spread of a sexually transmitted disease specifically among sex workers and clients using a bipartite network.

Although some of the aforementioned modelling studies have examined the potential impact of PrEP or immunization strategies in sex worker populations, in the present work we study the distribution of combination HIV prevention between two distinct populations, namely the FSWs and their clients. Using our model, calibrated against historical data from a South African study, we investigate the potential effectiveness of both PrEP and ART for HIV control and prevention in a population of sex workers and their clients. Unlike previous simulation studies which only distinguish between "high-risk" individuals (both sex workers and their clients) and "low-risk" individuals (the general population) we will consider sex workers and clients separately and will not directly consider the impact on the general population. While it is intuitive to imagine that the most cost-effective strategy would be to target all resources towards the FSW population, our model shows that this is only true up to a certain point. For aggressive reduction in HIV prevalence, targeting the clients for treatment is also necessary.

1.1 HIV Epidemiology

The typical course of an HIV infection, as shown in Figure 1.1, consists of at least three distinct stages. Flu-like symptoms may start to appear in patients 3–6 weeks following infection. This acute stage is associated with a high viral load and high infectivity. This period is followed by a long latent stage (typically asymptomatic); if left untreated, the disease progresses to AIDS, on average, after 8–10 years.

Mathematical models of infectious diseases aim to describe the coupled dynamics of the transmission of the disease from the infectious to the susceptible population and the progression of the disease within the infected population. Deterministic compartmental models are classical epidemiological tools; disease states are identified and flux laws are

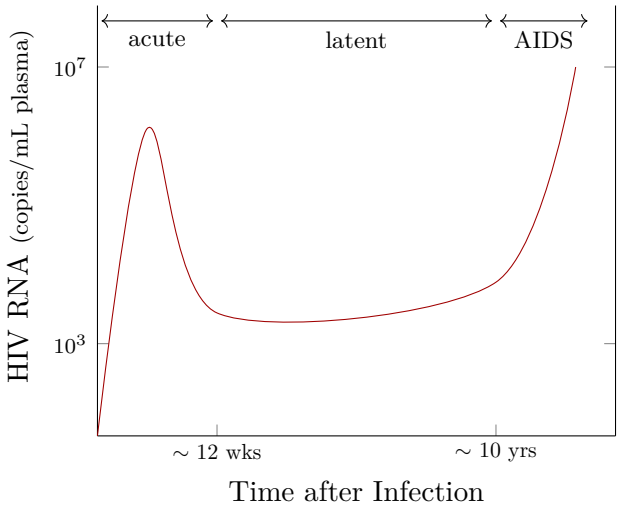


Figure 1.1. *Change in viral load, which is proportional to the infectivity of the patient, over the course of HIV infection. Adapted from [14].*

used to describe the flow of the population between states. By contrast, in many stochastic models, the movement of individuals between states is determined by constant probabilities per unit time. Both deterministic and stochastic models can be analytically tractable—as illustrated, for example, in Section 2.2.1—but realistic models often rely on either numerical solutions in the deterministic case or simulation results in the stochastic case.

A fundamental epidemiological quantity is the *basic reproduction number*, R_0. This quantity is defined as the expected number of secondary infections resulting from an index case in an otherwise susceptible population. If $R_0 > 1$, the disease persists and may lead to an epidemic. We can decompose the basic reproduction number into the following fundamental components [34]:

$$R_0 = c\beta D, \tag{1.1}$$

where

- c is the rate of infectious contacts;
- β is the probability of transmission per infectious contact; and
- D is the duration of infection.

An *infectious contact* is defined to be any contact which may result in infection. In the context of sexually transmitted HIV, an infectious contact is any sexual contact between an HIV-positive and an HIV-negative individual. Therefore, all of the above factors—rate of contact, probability of transmission, and duration—contribute to disease spread and can be targeted by control strategies.

Reducing the number of sexual contacts can clearly lower an individual's risk of contracting HIV; however, mathematical models have shown that the timing of partnerships may be equally important [42]. Keeping the number of sexual contacts fixed, a population

with more concurrent partnerships (as opposed to partnerships which are serially monogamous) leads to an increased epidemic and an increased growth rate in the early stages of an outbreak. To gain some intuition for this idea, consider an epidemic in its very early stages. If a susceptible individual i is in a monogamous partnership with a second individual j, the probability of j being already infected and subsequently transmitting to i is low. If j is in concurrent relationships, then the probability of i becoming infected is now proportional to both the probability that j is initially infected and the probability that j becomes infected via other partnerships. Most importantly for HIV, concurrency increases the number of partners exposed to the disease during the highly infectious acute stage (this is also when infected individuals are least likely to be on treatment). Whether or not concurrency is a driving factor in the generalized HIV epidemic in sub-Saharan Africa has been a subject of debate; recently, the model of Leung and Kretzschmar [33] has supported this theory.

In an attempt to quantify the increase in infectivity accompanying the change in viral load, Wawer *et al.* [59] study a cohort of serodiscordant (one HIV-negative and one HIV-positive partner) heterosexual couples. They observed an average rate of transmission of 0.0015 per sexual act within 6–15 months of infection. During the acute stage of infection, it is estimated that the probability of transmission per sexual contact is up to 26 times higher than that during the latent stage [29]. The transmission probability per unprotected coital act is considered to be a function of both the infectivity of the HIV-positive partner and the susceptibility of the HIV-negative partner. In heterosexual couples, it is sometimes assumed that women are more susceptible to sexually transmitted infections due to biological factors; the study of Wawer *et al.* [59], however, does not reveal any of this asymmetry.

Studies have shown that consistent condom usage is up to 95% effective at preventing HIV transmission [48]. It has been shown that a man's risk of contracting HIV is reduced by over 50% [26] if he is circumsised; however, unlike condoms, male circumcision has not conclusively been shown to protect against transmitting HIV in the case where the circumcised partner is HIV-positive [27]. In this thesis, we are primarily interested in biomedical control strategies. In particular, we will look at pre-exposure prophylaxis (PrEP) and antiretroviral therapy (ART). PrEP is designed to be taken daily by HIV-negative individuals, and can reduce the likelihood of infection through sexual contact by more than 70% [3]. ART increases the life expectancy of HIV-positive individuals and has been shown to reduce transmission by over 90% [2]. Therefore, ART also serves to prevent new infections and reduce HIV incidence in the population [40]. Life-long ART is necessary for continued suppression of viral load.

1.2 Complex Networks

In models of diseases which are primarily airborne and/or diffuse rapidly throughout a population, a homogeneous mixing assumption, such as that required for compartmental

models, may be suitable. However, it is inappropriate to assume that sexual transmission of HIV is equally likely to occur between any two members of a population. Therefore, in our model of disease transmission, transmission can only occur along a sexual contact network.

A network is a set of connections which represents some physical, technological or social system. Social networks in particular are used to understand the connections between individuals, or perhaps groups of individuals. A network is described mathematically by a graph, $G = (V, E)$, with vertex set V and edge set E. Vertices (or nodes) typically represent individuals or groups of individuals, and an edge between vertices indicates some sort of relationship between them. The edges can be either directed or undirected, as appropriate. The network imparts rules on the disease propagation; an infected vertex can transmit the disease only to adjacent (neighbouring) vertices, and even then only if this neighbour is susceptible to the disease. In this thesis, a network will refer to either an empirical or a generated set of connections and a graph will refer to the mathematical objects which can be used to describe networks. The sociological discipline of social network analysis is briefly reviewed in Section 1.2.1 and the probabilistic study of random graphs is introduced in Section 1.2.2.

1.2.1 Sexual Contact Networks

Social network analysis is the study of empirical networks, the goal being to characterize and find patterns in the connections between people or groups. A classical example is the friendship network, where a directed edge from ego to alter signifies that ego considers alter to be a friend. A sexual contact network is a particular social network in which the edges represent sexual contact (or the possibility of sexual contact). The study of a sexual contact network can reveal more than just the average number of sexual partners. It may identify key individuals with an unusually high number of partners, or subnetworks which are more highly connected. One could also answer questions such as whether individuals of the same race or age are more or less likely to be partners.

In [35], the authors find that the distributions of the number of sexual partners over an 18-month period among adult heterosexual men and women approximately follow a power law. In a more detailed study, Bearman *et al.* [5] investigate the nature of sexual contacts in an American high school. They characterize this network as having one large component—comprised of approximately half of all students—with a ring-like structure and few short cycles. The remainder of the network is mostly comprised of dyads (pairs) and triads (triangles). The authors note that the distribution of the data fits well to a power law, but find it unlikely that this represents the actual distribution of partnerships. Both of the above works discuss the implications of the observed network structure on disease spread and control. To our knowledge, no detailed social network analysis has been published for sex worker contact networks.

While there are many techniques which sociologists use to uncover social networks [10], much of the published data regarding sex worker relationships has relied heavily on questionnaire data. Unfortunately, such data suffer from biases and may potentially be unreliable. Ferguson *et al.* [22] compared female sex worker (FSW) responses from standard recall questions to diary entries. They determined that FSWs consistently over-estimated the average number of daily clients when asked on a questionnaire. When asked to estimate the number of clients over a longer time period, the researchers discovered that the respondents had difficulty separating the number of encounters from the number of different partners. This poses a significant challenge in the interpretation of questionnaire data for the purposes of constructing a sexual contact network.

1.2.2 Random Graphs

The mathematical description of a network uses the language of graph theory. Let G be an undirected graph with n vertices and m edges. An edge between vertices i and j is represented as the unordered pair (i, j), and the number of edges incident to a vertex is its *degree*. The degree is thus a local measure of connectedness. An important global property of a random graph is the *degree distribution*:

Definition 1. Degree Distribution
The *degree distribution* $p(k)$ gives the probability that a randomly chosen vertex has degree k.

The theory of random graphs originated with Erdős and Rényi in 1959 [21]. Their model supposed that G is initialized with n vertices, and edges are subsequently added between each pair of vertices with fixed probability q. The degree of each vertex is therefore a binomial random variable, with

$$p(k) = \binom{n-1}{k} q^k (1-q)^{n-1-k}.$$

Let $\lambda = nq$ be constant. Then in the limit as $n \to \infty$,

$$p(k) = \frac{\lambda^k e^{-\lambda}}{k!}.$$

Therefore, an Erdős-Rényi random graph would be a suitable model for a real-world network which exhibits a Poisson degree distribution, or in a case where the underlying generative mechanism is applicable: that is, if connections between nodes are thought to form randomly with some fixed probability.

The Configuration Model

A random graph model which is more flexible with respect to the degree distribution is the Configuration Model, which will be referred to often in this thesis. All of the networks in Chapter 2 and Chapter 4 have been generated using this algorithm. Introduced by Newman *et al.* [45], this algorithm generates a random graph which has a prescribed degree distribution. This model is especially useful in cases where the only known network property is its degree distribution.

The following algorithm generates a configuration model random graph (CMRG) of size n:

Algorithm 1. Generating a CMRG

1. Initialize the graph with n vertices and no edges.
2. Given the desired degree distribution $p(k)$, sample $p(k)$ to construct a degree sequence s_j representing the degree of each vertex.
3. For $j = 1, \ldots, n$, choose a random untouched vertex j and attach to it s_j *stubs* (or half-edges).
4. For each stub in the network, choose a partnering stub uniformly at random, and combine them to form an edge. Repeat until no stubs are left unmatched[1].

This process will generate a (not necessarily simple) random graph G having the desired degree distribution $p(k)$ in the limit as $n \to \infty$. More precisely, G is chosen uniformly from the set of all possible graphs on n vertices with the specified degree distribution.

Now suppose G is a CMRG with n vertices and m edges. Then the probability that any two vertices have an edge connecting them is simply

$$P(\text{edge between } i \text{ and } j) = \frac{\deg i \deg j}{2m-1} \simeq \frac{\deg i \deg j}{2m}. \tag{1.2}$$

Given this fact, many properties of the random graph are readily determined. For example, we can determine the excess degree distribution of a CMRG:

Definition 2. The Excess Degree Distribution
The *excess degree distribution* $q(k)$ gives the probability that a randomly chosen edge e leads to a vertex which has k other edges (not counting e). Here, k is the excess degree of this vertex.

[1]This is guaranteed if the number of stubs is even. In practice, a stub may need to be added at random.

Choose a random stub in the configuration network. This stub attaches to any other stub in the network with equal probability, $\frac{1}{2m-1} \approx \frac{1}{2m}$. So the probability that the stub attaches to a particular vertex of degree k is $\frac{k}{2m}$. Since there are $np(k)$ such vertices, the probability that any neighbour is of degree k is

$$k\frac{np(k)}{2m} = k\frac{p(k)}{\langle k\rangle}, \tag{1.3}$$

where $\langle k\rangle = 2m/n$ is the mean degree. That is, we are k times more likely to arrive at a vertex of degree k than a vertex of degree 1. Then the excess degree distribution $q(k)$ is given by

$$q(k) = \frac{(k+1)\cdot p(k+1)}{\langle k\rangle}. \tag{1.4}$$

Both self-loops and multi-edges are included in the generation process. However, the following theorem states that the density of self-loops and multi-edges is $\sim O(1/n)$.

Theorem 1. [19, Theorem 3.1.2]
Let $\langle k\rangle := \sum_k kp_k$ and $\langle k^2\rangle := \sum_k k^2 p_k$ both be finite. Then as $n \to \infty$, the expected number of self-loops is

$$\frac{\langle k^2\rangle - \langle k\rangle}{2\langle k\rangle},$$

and the expected number of multi-edges is

$$\left(\frac{1}{2}\frac{\langle k^2\rangle - \langle k\rangle}{\langle k\rangle}\right)^2.$$

Therefore, as $n \to \infty$, the density of multi-edges and self-loops approaches zero. These formulas are readily derived by noting that the probability of an edge between any two vertices is given by Equation (1.2).

A *component* of a graph is a maximal subset of vertices for which every vertex is reachable via some path from every other vertex in the subset. The following theorem describes the components of a CMRG. In particular, we are interested in the emergence of a so-called *giant* component, a component with $O(n)$ vertices as $n \to \infty$. All other components are referred to as *small* components. If a random graph has only small components, it is said to be *sub-critical*, otherwise it is said to be *super-critical* if there exists a giant component.

Theorem 2. [19, Theorem 3.2.2]
Suppose $\frac{1}{\langle k\rangle}\sum_k k(k-1)p(k) > 1$. Then there is a giant component of size $O(n)$ and no other clusters of size greater than $\beta \log n$ for some constant β.

The size of the giant component can be derived using branching process methods (see Section 2.1 for a brief review). In [19], the size of the small components, and the proof of the emergence of a giant component, are proved using an approximation to a random walk.

An important consequence of the above theorem is that, in the limit as $n \to \infty$, the small components of G are tree-like. In the giant component, however, there exist many cycles. This will be important in the analysis of the SIR process on a CMRG in Chapter 2.

Chapter 2

The SIR Model

The classical SIR model describes a disease process in which individuals cannot become reinfected. The population is divided into Susceptible (S), Infected (I) and Removed/Recovered (R) states. Susceptibles either remain in that state for the entirety of the experiment, or they become infected and then removed as illustrated in Figure 2.1. The disease dynamics cease once every individual is in either the Susceptible or Recovered state, with the latter representing the total number of individuals who were at any point infected.

Since HIV is a life-long disease, an SIR-type disease model may seem to be an appropriate first choice. For reasons which are discussed in more detail in Section 4.1.2, in our simulations we ultimately use a SIRS-type model, and so the analytical results presented in the following sections cannot be directly applied. However, this Chapter is relevant to our discussion since it provides examples of analytical approaches to the study of epidemics on networks, and more importantly, allows us to test our numerical simulations against theoretical results.

Analytical results exist for SIR dynamics on networks generated by the Configuration Model (described in Section 1.2.2) in the infinite size limit. Specifically, the studies presented here aim to quantify:

i. **The Epidemic Threshold.** A critical graph parameter. Above this value, the probability of sparking an epidemic is nonzero. Below the threshold, the disease spread is confined to a finite number of individuals.

ii. **The Epidemic Size.** The proportion of nodes which ever become infected in the regime above the epidemic threshold.

iii. **The Outbreak Size.** The number of nodes which ever become infected in the regime below the epidemic threshold.

Figure 2.1. *Basic schematic of the SIR process.*

iv. **The Probability of an Epidemic.** Below the epidemic threshold, the probability of sparking an epidemic is, by definition, zero. Above the epidemic threshold, the probability is nonzero but not necessarily one.

There has been additional research focused on time-dependent properties of SIR dynamics on random graphs [57, 17, 30], but this subject is not covered here.

2.1 Background

In order to follow the analyses presented in [44] and [31] which are reviewed in Sections 2.2.1 and 2.2.2 respectively, we provide some necessary background information below. First, we review probability generating functions, followed by a review of homogenous branching processes using the generating function formalism. Finally, we will cover the basics of percolation theory before applying this concept to disease spread on networks.

Probability Generating Functions

A probability generating function (pgf) is a notational structure containing the information of a distribution.

Definition 3. Probability Generating Function
Let Z be a discrete random variable with

$$p(k) \coloneqq P(Z = k), \quad k = 0, 1, 2, \ldots.$$

Then

$$G(x) \coloneqq E\left(x^Z\right) = \sum_{k=0}^{\infty} p(k)x^k$$

is the *probability generating function* of Z.

In particular,

$$G(1) = \sum_k p(k) = 1$$

for a properly normalized distribution, and the mean $\langle k \rangle$ can be recovered as

$$G'(1) = \sum_k kp(k) = \langle k \rangle.$$

So if $p(k)$ denotes the degree distribution of a random graph, we take $G_0(x)$ to be the corresponding generating function. In the case of a CMRG, we can further determine the generating function for the *excess degree distribution* $q(k)$ (1.4):

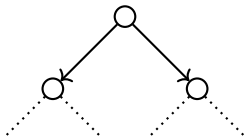
$$G_1(x) = \sum_k q(k)x^k = \frac{\sum_k (k+1)p(k+1)x^k}{\langle k \rangle} = \frac{G_0'(x)}{G_0'(1)}.$$

Branching Processes

A branching process is a probabilistic model of reproduction. At each generation, each organism gives rise to a random number of offspring independently. The number of offspring of each individual is distributed according to the *offspring distribution*,

$$p(k) \coloneqq P(\text{an organism has } k \text{ offspring}).$$

Consider a homogeneous branching process starting with a single organism:



Let X_i denote the number of offspring of the i^{th} individual, and let the number of offspring in the first generation be generated by $G(z)$,

$$G(z) \coloneqq E[z^{X_0}] = \sum_k p(k) z^k.$$

Since the process is homogeneous, by definition every subsequent generation will have the same offspring distribution as the first. If there are k offspring in the first generation, then the total number of offspring ever generated is

$$1 + X_1 + X_2 + \cdots + X_k.$$

So the probability generating function $H(z)$ for the total number of offspring is

$$H(z) = E\left(z^{1+\sum_{i=1}^{k} X_i}\right) = zG\left(H(z)\right). \tag{2.1}$$

Percolation Theory

Percolation theory—in particular, discrete percolation theory—was developed to understand the flow or spread of some medium on a random lattice. It is a model which exhibits thresholding behaviour, as with our study of the structure of random graphs. There is a parameter which takes on a *critical* value at which the general behaviour of the process changes. For example, the thresholding parameter in the study of the generation of Erdős-Rényi random graphs is the mean degree λ, with critical value 1. For $\lambda < 1$ the probability of a giant component is zero, while if $\lambda > 1$, a giant component exists with high probability.

Definition 4. Bond Percolation

In a *bond percolation* process, each edge—or bond—is occupied (kept) with probability p, or not (removed) with probability $1 - p$. For an example, see Figure 2.2.

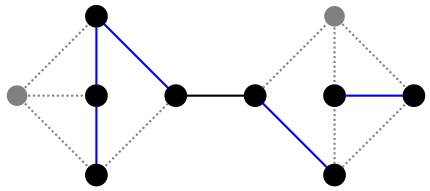


Figure 2.2. *Example of bond percolation with $p = \frac{6}{15}$. The occupied edges are in blue, and the dashed edges are edges from the original graph which were not occupied. Isolated vertices (in grey) are subsequently removed.*

Definition 5. Site Percolation

In a *site percolation* process, each vertex—or site—is occupied with probability p, or not with probability $1 - p$. If the vertex is unoccupied, all adjacent edges are additionally removed. For an example, see Figure 2.3.

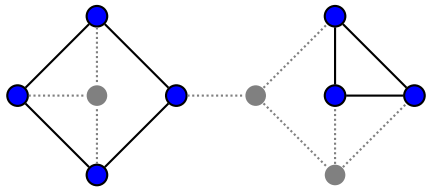


Figure 2.3. *Example of site percolation with $p = \frac{7}{10}$. Occupied vertices are in blue, and the grey vertices are those from the original graph which were not occupied. All edges incident to the removed vertices are not included.*

Here, the critical parameter is the bond (or site) occupation probability p and the critical phenomenon is the emergence of a giant component on the percolation network, after occupying or not the vertices (or edges) of the underlying network.

2.2 Analytical Results

2.2.1 Bond Percolation Approach

In the context of the SIR disease process, the super-critical regime of the percolation process is the epidemic regime, and in the sub-critical regime, the small components are outbreaks. It may not seem obvious at first which process—site or bond—is appropriate to model disease spread. Vertices either become infected and then removed or they remain susceptible, but the disease itself propagates along edges. The first analysis of this problem, due to Newman in 2002 [44], used a bond percolation model on an undirected CMRG; we describe his approach below. In the following, we let G_0 be the probability generating function (pgf) for

the degree distribution of the underlying CMRG, and G_1 be the pgf for the excess degree distribution:

$$G_0(x) = \sum_k p(k)x^k \qquad \text{and} \qquad G_1(x) = \sum_k q(k)x^k = \frac{G_0'(x)}{G_0'(1)}.$$

Suppose i is an infected vertex, and j is a susceptible vertex which is adjacent to i in the contact network. Let τ_i be the length of time vertex i remains in the infectious state, and let r_{ij} be the infectious contact rate between adjacent vertices i and j. Then the probability that i infects j over the time period τ_i is $T_{ij} = 1 - e^{-r_{ij}\tau_i}$.

If we suppose the r_{ij} and τ_i are independent and identically distributed (i.i.d.), then we define the *transmissibility* T of the disease to be the average over the T_{ij}:

$$T = 1 - \int_0^\infty \int_0^\infty e^{-r\tau} P(\tau)P(r)\, dr\, d\tau. \tag{2.2}$$

The connection between the percolation process and the disease process comes from letting T be the bond occupation probability, where in this case, occupation of an edge is synonymous with transmission of the disease along that edge. Starting from an initially infected vertex i (the index case) of degree k, assume that the vertices adjacent to i either become infected or not with probabilities T and $1 - T$, respectively, independent of each other[1]. Then the probability that j out of k incident edges become infected is

$$\binom{k}{j} T^j (1-T)^{k-j}.$$

Therefore, the pgf $G_0(x;T)$ for the distribution of occupied edges of a randomly chosen vertex is

$$G_0(x;T) = \sum_{j=0}^{\infty} \sum_{k=j}^{\infty} p(k) \binom{k}{j} T^j (1-T)^{k-j} x^j = G_0(1 - T + xT). \tag{2.3}$$

Similarly, the pgf for the distribution of occupied edges of a vertex arrived at by following a randomly chosen edge is

$$G_1(x;T) = G_1(1 - T + xT). \tag{2.4}$$

To continue this analysis, it is assumed that the disease process can be mapped to a branching process. A consequence of Theorem 2 is that the small components of a CMRG are tree-like, conditioned on the network having a bounded second moment. So if the CMRG is sub-critical, then equation (2.1) may be applied. The expected number of vertices which will ever become infected is now approximated by the expected number of offspring in the branching process. In the CMRG, the degree of a neighbour is distributed according to G_1,

[1]This assumption of independence is incorrect, as explained in the discussion surrounding equation (2.9).

and each subsequent neighbour is also distributed according to G_1. So by equation (2.1), the pgf H_1 for the size of the cluster reachable from a randomly chosen edge is

$$H_1(x;T) = xG_1(H_1(x;T);T). \tag{2.5}$$

The pgf H_0 for the size of the cluster reachable from a randomly chosen index case is similarly derived. In this case, the degree of the index case is distributed according to G_0; however, the sizes of the clusters reachable from the neighbours of this index case will be distributed according to H_1. Thus,

$$H_0(x;T) = xG_0(H_1(x;T);T). \tag{2.6}$$

From here, the expected outbreak size is calculated directly from H_0. The epidemic threshold T_c is defined to be the value of T when the expected outbreak size ceases to be finite, $\lim_{T\to T_c} H_0(x;T) = \infty$. In the super-critical regime, although the tree-like assumption is not valid in the giant component, H_0 is again the distribution of the small components, from which the size of the giant component can be inferred. Note that in this model, the expected fraction of vertices in the epidemic and the probability of a randomly chosen index case sparking an epidemic are the same: they are both the probability that the randomly chosen index case is located in the giant component of the percolation network. A summary of the results from [44] is shown in Table 2.1. As expected, the results depend only on the parameter T and on the degree distribution of the CMRG.

For example, consider a Markovian SIR process where β is the constant infection rate and γ is the constant recovery rate. Then, combining equation (2.2) with the results in Table 2.1, we find

$$T_c = \frac{\beta_c}{\beta_c + \gamma_c} = \frac{1}{G_1'(1)} = \frac{\langle k\rangle}{\langle k^2\rangle - \langle k\rangle},$$

and so

$$\frac{\beta_c}{\gamma_c} = \frac{\langle k\rangle}{\langle k^2\rangle - 2\langle k\rangle}. \tag{2.7}$$

Expected Outbreak Size	$\langle s\rangle$	$= H_0'(1;T) = 1 + \dfrac{TG_0'(1)}{1 - TG_1'(1)}$
Epidemic Threshold	T_c	$= \dfrac{1}{G_1'(1)}$
Expected Epidemic Size (as fraction)	S	$= 1 - H_0(1;T)$
Probability of Epidemic	E	$= 1 - H_0(1;T)$

Table 2.1. *Results from the bond percolation model derived in [44].*

Unfortunately, as pointed out in [19] and [31], there is a fundamental error with this model. The transmission events among neighbours are not independent. Consider the fol-

lowing counterexample from [19]: Let $r_{ij} = \beta$ and let τ_i be exponentially distributed (with mean 1). Then

$$1 - T = \int_0^\infty e^{-\beta t} e^{-t} dt = \frac{1}{\beta + 1} \tag{2.8}$$

is the probability that a neighbour of an infected vertex is not itself infected along that edge. However, the probability that two neighbours are not infected is

$$\int_0^\infty e^{-2\beta t} e^{-t} dt = \frac{1}{2\beta + 1} > \left(\frac{1}{\beta + 1}\right)^2. \tag{2.9}$$

Thus the T_{ij} are non-negatively correlated between adjacent vertices. In fact, T_{ij} are independent only if the *recovery time* is constant. Therefore, the assumption that we can take T to be a bond occupation probability is incorrect, and so the generating functions derived above are incorrect. However, as is shown in the work of Kenah and Robins [31] discussed in the following section, Newman's formulas in [44] for the epidemic threshold, expected outbreak size and expected epidemic size are correct. This is because only the first moment of H_0 is correct, as will be discussed at the end of the following section.

2.2.2 Epidemic Percolation Networks

Keeping with a percolation model approach, Kenah and Robins [31] developed a new framework for understanding the SIR process on CMRGs which corrected the mistake in Newman's bond percolation approach. Using the same notation as in the previous section, from the underlying CMRG, the authors generate a so-called *Epidemic Percolation Network* (EPN) as follows: For each pair of adjacent vertices i, j, either delete the edge (i, j), keep the undirected edge, or convert it to a directed edge with the following probabilities (recall that $e^{-r_{ij}\tau_i}$ is the probability that i does not infect j):

i. $P(i \to j) = (1 - e^{-r_{ij}\tau_i})e^{-r_{ji}\tau_j}$

ii. $P(i \leftarrow j) = e^{-r_{ij}\tau_i}(1 - e^{-r_{ji}\tau_j})$

iii. $P(i - j) = (1 - e^{-r_{ij}\tau_i})(1 - e^{-r_{ji}\tau_j})$

iv. $P(i, j \text{ disconnected}) = (e^{-r_{ij}\tau_i})(e^{-r_{ji}\tau_j})$

So the edge between i and j is left undirected with the probability that both transmission events occur; that is, if either of i or j becomes infected, then the disease will spread to the other. Otherwise, the disease will only spread according to the direction of the edge. This algorithm therefore constructs a semi-directed network.

To illustrate how this construction leads to independent transmission events among neighbours, consider the counterexample from the previous section. Assume a constant contact rate β, and consider an infected node i with two neighbours j and j'. There are four

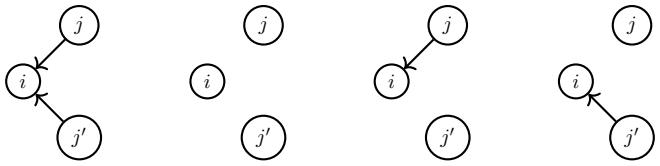


Figure 2.4. *Given an initially infected node i and two neighbours (in the underlying CMRG) j and j', these are the four possible configurations in the epidemic percolation network in which both j and j' escape infection.*

possible configurations of the three nodes in which the disease will not spread to either j nor j', as shown in Figure 2.4. Hence the probability that two neighbours of i escape infection in the EPN is

$$\begin{aligned} &P(i \leftarrow j)P(i \leftarrow j') + P(i, j \text{ disconnected})P(i, j' \text{ disconnected}) \\ &+ P(i \leftarrow j)P(i, j' \text{ disconnected}) + P(i, j \text{ disconnected})P(i \leftarrow j'), \end{aligned}$$

which gives a total probability of $e^{-2\beta\tau_i}$. Then if τ is exponentially distributed with mean 1, integrating over the distribution of τ gives

$$\int_0^\infty e^{-2\beta t} e^{-t} dt = \frac{1}{2\beta + 1},$$

which is now the correct probability that both neighbours escape infection (2.9). A proof that the T_{ij} in the EPN are independent is given in [31].

Since the EPN is a semi-directed graph, the interpretation of an outbreak and an epidemic need to be redefined. In order to do so, we first introduce the following definitions for the components of semi-directed graphs. A directed path is a path which obeys the direction of the edges; in a semi-directed network, an undirected edge can be followed in either direction.

Definition 6. Components of (Semi-) Directed Graphs

1. An in-component of a vertex i is the component which includes all vertices from which i is reachable via some directed path.

2. An out-component of a vertex i is the component which includes all vertices which are reachable from i via some directed path.

3. A strongly connected component of a vertex i is the component which includes all vertices from which i is reachable and which can be reached from i following directed paths. It is the intersection of the in- and out-components of i.

As with undirected graphs, there is a super-critical regime in which we expect to see the emergence of a giant component. Now there is a giant strongly connected component (G_{SCC}), with unique giant in- and giant out-components (G_{IN} and G_{OUT}, respectively). The

interpretation of Kenah and Robins [31] for the disease process on the EPN is summarized below:

Theorem 3. Interpretation of EPNs

1. The epidemic threshold corresponds to the emergence of a *giant strongly connected component* G_{SCC}.
2. If the EPN contains a G_{SCC}, then any index case in the *giant in-component* G_{IN} will lead to infection in the entire *giant out-component* (G_{OUT}).
3. The probability of an epidemic corresponds to the fraction of vertices in G_{IN}.
4. The final size of the epidemic corresponds to the fraction of vertices in G_{OUT}.

With these interpretations in mind, the authors now proceed in a similar manner to Newman [44]. The probability generating functions for the distribution of occupied edges of a given vertex are derived, and hence the expected size of the cluster reachable from a randomly chosen initial vertex (for full details, refer to [31]). Kenah and Robins find that their approach gives the same expected outbreak size, epidemic threshold and expected epidemic size as Newman's bond percolation model. However, they find that the bond percolation model overestimates the probability of an epidemic. These results are not too surprising. Even though Newman's generating function H_0 (2.6) is wrong, the first moment of H_0 is correct. The occupied edges are not *independently* binomially distributed, but the expectation of a sum of i.i.d. random variables is the same as the expectation of a sum of correlated i.d. random variables.

2.3 Simulation Results

We test the validity of our numerical approach described below by reproducing some of the analytical results of Section 2.2.1. In particular, many of the results in Table 2.1 have simple expressions which are readily evaluated. We perform discrete-time simulations with infectious contact rate β (constant in time and independent of the vertex pairs) and constant recovery rate γ. The simulations terminate if there are no longer any nodes in the infected state. The simulations are performed as follows:

Algorithm 2. Discrete Time SIR Process
Take a network with n nodes and an initial state space $\{s_1^0, s_2^0, \ldots, s_i^0, \ldots, s_n^0\}$, where s_i^j denotes the state of node i at time $t_j = j\Delta t$. Each s_j^i is one of S, I or R for all i, j. Then:

For $j = 0, 1, 2, \ldots$:

For $i = 1, 2, \ldots n$:

1. Compute a pseudo-random number p.

2. If $s_i^j = \text{S}$, then compute $\mathcal{N}(\text{I})$, the number of infectious neighbours of i. If $p < \beta\mathcal{N}(\text{I})\Delta t$, set $s_i^{j+1} = \text{I}$; else set $s_i^{j+1} = \text{S}$.
3. If $s_i^j = \text{I}$, then if $p < \gamma\Delta t$, set $s_i^{j+1} = \text{R}$; else set $s_i^{j+1} = \text{I}$.
4. If $s_i^j = \text{R}$, set $s_i^{j+1} = \text{R}$.

The random networks used for the following experiments were generated using the python package NetworkX [28]. The disease processes on each network were simulated using the python package NepidemiX [1]. Appendix B provides further details regarding the software used.

Since the recovery rate γ is constant, the recovery time τ must be exponentially distributed and so the transmissibility T becomes

$$T = 1 - \int_0^\infty \gamma e^{-(\beta+\gamma)\tau} d\tau = 1 - \frac{\gamma}{\beta+\gamma} = \frac{\beta}{\beta+\gamma}. \tag{2.10}$$

We first consider SIR dynamics on a unipartite configuration model network with Poisson degree distribution. To construct these networks, a Poisson degree sequence is randomly generated such that

$$p(k) = e^{-\lambda}\frac{\lambda^k}{k!}, \tag{2.11}$$

where the mean degree $\langle k \rangle = \lambda$. It follows that the generating function for the degree distribution is

$$G_0(z) = \sum_k p(k) z^k = e^{-\lambda} \sum_k \frac{(\lambda z)^k}{k!} = e^{\lambda(z-1)}, \tag{2.12}$$

and the generating function for the excess degree is

$$G_1(z) = \frac{G_0'(z)}{G_0'(1)} = \frac{\lambda e^{\lambda(z-1)}}{\lambda} = e^{\lambda(z-1)}. \tag{2.13}$$

In this case, the theoretical expressions for the epidemic threshold, T_c, and expected outbreak size, $\langle s \rangle$, are (using Table 2.1):

$$T_c = \frac{1}{\lambda} \tag{2.14}$$

and

$$\langle s \rangle = 1 + \frac{\lambda T}{1 - \lambda T} = 1 + \frac{\lambda \frac{\beta}{\beta+\gamma}}{1 - \lambda \frac{\beta}{\beta+\gamma}}. \tag{2.15}$$

The expected epidemic size S is determined numerically[2] by first finding u such that

$$u = e^{\lambda T(u-1)}, \tag{2.16}$$

and subsequently solving

$$S = 1 - e^{\lambda T(u-1)}. \tag{2.17}$$

Figure 2.5 shows the agreement between these theoretical results and simulation results. The clustering of data points around the predicted epidemic threshold shows that this corner can be accurately resolved. Below the epidemic threshold, the fraction of nodes which are infected is $O(1/n)$, as expected, and the standard error is small. This is expected since below the threshold, the probability of sparking an epidemic should be close to zero. For $T \in [0.6, 0.8]$, the average epidemic size of the simulated data is lower than the theoretical predictions. We expect this is due to the "finite-network" effect: the theoretical results were developed in the asymptotic size limit, while the simulated networks were taken with $n = 40,000$. The network size was chosen by considering the trade-off between computational time and accuracy.

The curves obtained from varying only the infectious contact rate β, and from only varying the recovery rate γ, are more or less identical. This confirms that only the ratio given in equation (2.10) need be considered. For all further simulations, only γ is varied with β fixed.

A second test was performed on CMRGs having a power-law distribution:

$$p(k) \propto k^{-\alpha}, \quad \alpha > 0,\ k \geq 1.$$

However, since the second moment

$$\langle k^2 \rangle = \sum_{k=1}^{\infty} k^{2-\alpha}$$

diverges for $\alpha \leq 3$, we instead used a power-law distribution with an exponential cutoff in order to obtain a degree distribution with a finite second moment for all values of α. This distribution has the form

$$p(k) = \frac{e^{-k/\kappa} k^{-\alpha}}{Li_{\alpha}(1/\kappa)}, \tag{2.18}$$

where the normalization constant is given in terms of the polylogarithm function,

$$Li_{\alpha}(z) = \sum_{k=1}^{\infty} \frac{z^k}{k^{\alpha}}.$$

[2]To solve equation (2.16), a root-finding algorithm based on the Powell hybrid method is used, called from python's scipy optimization library.

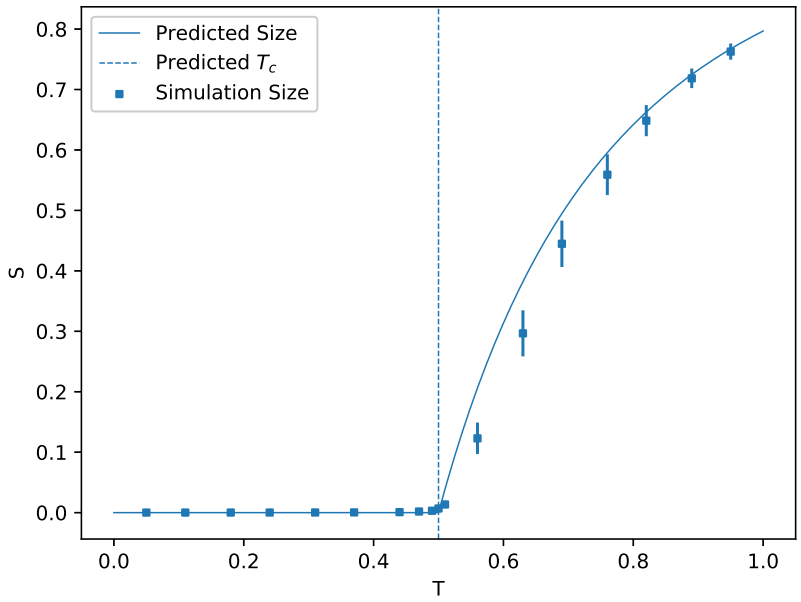


Figure 2.5. *Comparison of simulation and theoretical results for the expected epidemic size. The theoretical values S are found by solving equation* (2.17) *and equation* (2.16)*. The epidemic threshold T_c is determined by equation* (2.14)*. Each data point is an average of 20 runs repeated over 40 random CMRGs. Both γ and β are varied. Each network has Poisson degree distribution with mean $\lambda = 2$* (2.11) *and is generated using the Configuration Model. Each network has 40,000 vertices, and simulations used a time step of $\Delta t = 0.1$. Standard error bars are included.*

The expected epidemic threshold, mean outbreak size and mean epidemic size do not have closed form solutions, and are thus calculated numerically. The results are shown in Figure 2.6, and show similar agreement as before.

2.4 Bipartite Networks

A *bipartite network* is one in which the nodes can be partitioned into two distinct sets, with edges running only from nodes of one type to nodes of the other type as shown in Figure 2.7. Such a network is an appropriate choice to model heterosexual contacts. To generate a bipartite network having prescribed degree distributions for each node set, a slight modification of the configuration model from Section 1.2.2 is necessary.

Suppose we have a bipartite network with m male and n female nodes. Let $p(k)$ be the degree distribution for the male nodes with mean degree λ, and let $f(k)$ be the degree distribution for the female nodes with mean degree μ. Similar to before, a degree sequence is generated from each distribution and the stubs are attached correspondingly. The total number of male stubs and the total number of female stubs must be equal, requiring

$$\lambda m = \mu n.$$

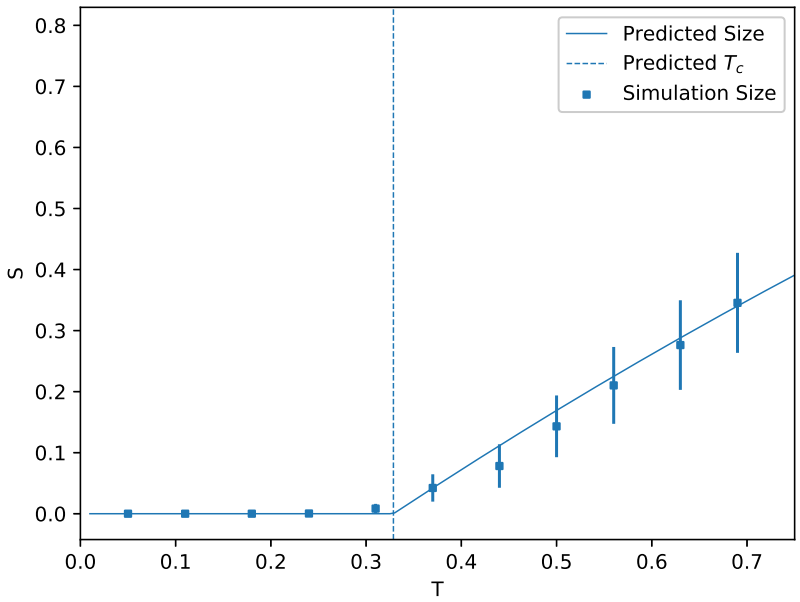


Figure 2.6. *Comparison of simulation and theoretical results for the expected epidemic size. Each data point is an average of 10 runs repeated over 40 random CMRGs. The degree distribution of each network follows a power law with exponential cutoff with exponent* $\alpha = 2$ *and exponential cutoff* $\kappa = 10$ (2.18) *and is generated using the Configuration Model. Each network has 60,000 vertices, and simulations used a time step of* $\Delta t = 0.1$. *Standard error bars are included.*

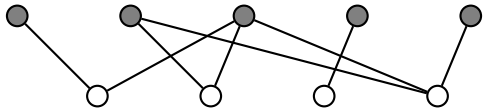


Figure 2.7. *An example of a bipartite graph.*

Under this constraint, the stubs can be matched uniformly at random as before. Of course, since the generation of both degree sequences is pseudo-random, the actual numbers of stubs for each population are frequently unequal. Therefore, after generating the degree sequences, the number of stubs are counted and if there is a mismatch, a node is chosen at random from the set which has excess stubs and its degree is reduced by one. This process continues until a matching can occur.

2.4.1 Bond Percolation Approach

In Newman's 2002 paper [44], he argues that SIR dynamics on bipartite networks can be described by a three-stage branching process. Let

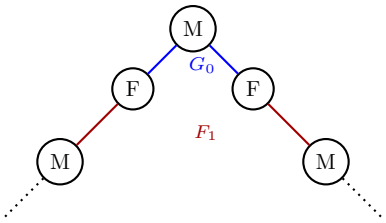


Figure 2.8. *Branching process describing transmission on a bipartite network. G_0 is the generating function for the male degrees; F_1 is the generating function for the female excess degrees.*

G_0 be the generating function for the degree of male nodes,

F_0 be the generating function for the degree of female nodes,

G_1 be the generating function for the excess degree of male nodes, and

F_1 be the generating function for the excess degree of female nodes.

The transmissibility T is defined as in equation (2.2), except that we now distinguish between male-to-female transmission T_{mf} and female-to-male transmission T_{fm}. Similar to Section 2.2.1, $G_0(x;T_{mf})$ generates the number of occupied edges of a randomly chosen male node, and $G_1(x;T_{mf})$ generates the number of occupied edges of a male node arrived at by following a randomly chosen edge. This approximating branching process is shown in Figure 2.8.

Starting at a randomly chosen male node, the generating function $\hat{G}_0$ for the number of infected males after two steps is generated by

$$\hat{G}_0(x;T_{mf},T_{fm}) = G_0(F_1(x;T_{fm});T_{mf}).$$

The generating function $\hat{G}_1$ for the number of infected males starting at a randomly chosen edge after two steps is similarly

$$\hat{G}_1(x;T_{mf},T_{fm}) = G_1(F_1(x;T_{fm});T_{mf}).$$

So, as a function of both transmissibilities, the generating function H_1 for the size of the cluster of infected males reachable from a randomly chosen edge is

$$H_1(x;T_{mf},T_{fm}) = x\hat{G}_1(H_1(x;T_{mf},T_{fm});T_{mf},T_{fm}).$$

The generating function H_0 for the size of the cluster of infected males reachable from a randomly chosen male node is then

$$H_0(x;T_{mf},T_{fm}) = x\hat{G}_1(H_0(x;T_{mf},T_{fm});T_{mf},T_{fm}).$$

The average outbreak size for males starting at a randomly chosen male node is thus

$$\langle s \rangle^M = H_0'(1),$$

as before. The average outbreak size for females is similarly derived and not necessarily equal to the average male outbreak size. The epidemic threshold turns out to be symmetric in F and G:

$$T_{mf}T_{fm} = \frac{1}{\hat{F}_1'(1)\hat{G}_1'(1)} = \frac{\langle k \rangle \langle j \rangle}{\big(\langle k^2 \rangle - \langle k \rangle\big)\big(\langle j^2 \rangle - \langle j \rangle\big).} \tag{2.19}$$

As with the unipartite case, while the calculation of the epidemic threshold is correct, this bond percolation approach cannot accurately describe the full disease dynamics. The reasoning for this is the same as before: while the generating function H_0 is wrong, the first moment of H_0 is correct. This is because the occupied edges are not *independently* binomially distributed, but the expectation of a sum of i.i.d. random variables is the same as the expectation of a sum of correlated i.d. random variables.

Intuitively, disease transmission on a bipartite network can be eliminated by preventing transmission in one direction only (either male to female or female to male). Therefore, in the context of immunization strategies, one might expect that only one subpopulation of the bipartite network needs to be targeted in order to prevent an epidemic. This idea is supported analytically by equation (2.19) which is symmetric in both female and male degree distributions. As we will see in Chapter 4, the ability of such an immunization strategy to suppress an epidemic is reduced when a more realistic model is considered.

2.4.2 Simulation Results

The analytical predictions from Section 2.4.1 are tested via stochastic simulations on large networks using the same methods as described in Section 2.3. We wish to generate a bipartite network on m male nodes and n female nodes are such that both populations have Poisson degree distributions. We first generate a Poisson degree sequence of size m with some input mean $\hat{\lambda}$. We subsequently generate a Poisson degree sequence of size n with mean $\hat{\mu} = \frac{\hat{\lambda} m}{n}$. The trial networks are then generated using the bipartite CMRG algorithm; as noted above, since in practice the number of stubs may not match exactly, the actual mean degrees λ and μ may change slightly after matching. The resulting bipartite networks are such that

$$p(k) \sim Poisson(\lambda) \quad \text{and} \quad f(k) \sim Poisson(\mu),$$

where $p(k)$ is the male degree distribution, $f(k)$ is the female degree distribution, and $\mu n = \lambda m$.

Following equation (2.19), the expected critical threshold is then

$$T_{mf}T_{fm} = \frac{1}{F_1'(1)G_1'(1)} = \frac{1}{\mu\lambda}.$$

If there is a constant rate β_{mf} at which males infect neighbouring females, a constant rate β_{fm} that females infect neighbouring males, and every node recovers at a constant rate γ, then

$$T_{mf} = \frac{\beta_{mf}}{\beta_{mf}+\gamma} \quad \text{and} \quad T_{fm} = \frac{\beta_{fm}}{\beta_{fm}+\gamma}.$$

Above the epidemic threshold, the expected size of the epidemic in females is calculated by first finding u such that

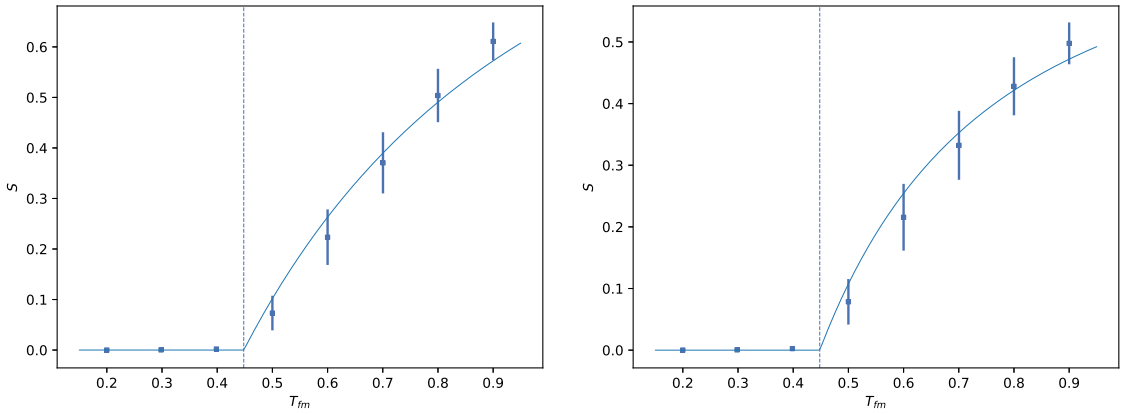
$$u = F_1(G_1(u;T_{mf});T_{fm}) = \exp\left[\mu T_{fm}\left(\exp\left(\lambda T_{mf}(u-1)\right)-1\right)\right],$$

and subsequently solving

$$S = 1 - F_0(G_1(u;T_{mf});T_{fm}) = 1 - \exp\left[\mu T_{fm}\left(\exp\left(\lambda T_{mf}(u-1)\right)-1\right)\right] = 1-u.$$

The expected size of the epidemic in males is similarly calculated.

These results are compared to simulations in Figure 2.9. In these simulations, we observe that that the size of the smaller population—as opposed to the total size of the network—impacts the quality of the simulation results.



(a) *Males* **(b)** *Females*

Figure 2.9. *Mean epidemic size in a bipartite CMRG with $\lambda \approx 2$ and m=90,000, $\mu \approx 6$ and n=30,000. Standard error bars are included. T_{mf} is fixed at 0.2 and T_{fm} is varied. Sample means (squares) compared to theoretical results (curve). Each data point is the mean of 1200 runs distributed over 80 different random networks.*

Chapter 3

Data Analysis

The characterization of sex worker–client interactions is a major consideration in this work. Previous research has largely focused on female sex worker (FSW) communities [36], since clients typically do not form an identifiable group (particularly in large urban settings) and complete social network data appears to be scarce. We analyze historical data collected on sex worker–client interactions in a South African mining community as part of the Carletonville-Mosuthumpilo Project [60] which took part from 1998–2000, and compare statistics of these interactions to those reported in other studies conducted in similar populations.

3.1 The Carletonville-Mosuthumpilo Project

The town of Carletonville, in Gauteng, South Africa, was the location of the largest gold mining complex in world around the year 2000, when this data was collected. The region was home to approximately 75,000 miners, most of whom were migrant workers from rural South Africa. Commercial sex work was prevalent in the community, and particularly in squatter settlements—so-called "hotspots"—surrounding the mine complexes, which housed an estimated 400–500 women [9]. Much of this sex work was survival sex work. HIV prevalence in 2000 was found to be 29% among miners, and almost 70% among FSWs in hotspots [60], while systematic treatment programs had yet to be introduced.

The Carletonville-Mosuthumpilo Project [60] was a community-led intervention program aimed at spreading awareness of and limiting the HIV epidemic in Carletonville and in its neighbouring township, Khutsong. The study interviewed approximately 800 miners, 1500 residents of Khutsong and 100 women living in hotspots, and we compiled the questionnaire data into summary data. In our analysis of the data, we take the FSWs to be those women who responded in the questionnaire that their usual occupation was sex work[1]. Approximately 7% of the women surveyed identified as FSWs, for a total sample size of 71.

[1]No men surveyed identified as sex workers.

About 8% of the men surveyed indicated that they had ever visited a FSW, giving a total client sample size of 109. This proportion was slightly higher among miners at roughly 9%, and miners constitute about 64% of the clients. We then estimate that there are roughly 6000–7000 clients, and in her book, Campbell estimates that there are 400–500 FSWs living in hotspots. These numbers suggest a client to FSW ratio of at least 10:1. Further data from the Carletonville study is shown in Appendix A.

The subtleties regarding the nature of the sex work in this region, and possible issues regarding the interpretation of what is and is not considered to be sex work, are described in detail in Campbell's book [9]. In particular, due to the nature of transactional sex work, the line between a "boyfriend" and a "client" may become blurred. When analyzing sex worker responses from the questionnaire data, we consider any non-regular partner to be a client since this is the only data which is available. The men surveyed were asked specifically about visits to sex workers, and so these responses are used. It is not always clear that our interpretation of this data matches the intended meaning.

3.2 Population Characteristics

The Carletonville-Mosuthumpilo Project questionnaire data provides insight into the demographics and behaviour of both clients and FSWs. Table 3.1 compares some summary metrics from this project with studies of similar communities. The statistics reported in this table indicate strong similarities between the Carletonville-Mosuthumpilo data and other studies of sex work in rural communities [56] and along commercial trucking routes [22, 38, 50].

We observe that the FSWs have a much higher migration rate than their clients, shown by the data from the Carletonville-Mosuthumpilo project on the mean time spent in the Carletonville area by FSWs and clients, respectively. This is consistent with the Ramjee and Gouws [50] study (in which the truckers were asked how long they had been at their current job). Although the specific factors contributing to the migration rate of FSWs are uncertain, it is plausible that the challenges faced by survival sex workers would motivate their departure once circumstances permit. We hypothesize that a significantly higher migration rate among FSWs compared with the client subpopulation is a typical characteristic of many survival sex worker–client sexual networks [53].

Table 3.1 further gives us information on HIV prevalence among FSWs and their clients. We are especially interested in the relative difference in disease prevalence between the two populations. Since the 29% prevalence from Carletonville is for *all* miners, one might expect the client-only prevalence to be higher than this. The studies from KwaZulu-Natal [50] and Uganda [38] show client prevalence of up to 54–100% of the FSW prevalence.

Study	Mean age	HIV prevalence	Mean number different partners past month	Mean number different partners past year	Mean number encounters past month	Mean time in location or on job
Carletonville [60]						
FSWs	30 (21–51)	70%	x	22 (1–300)	9.7 (0–62)	4 yr
Clients (miners)	35 (17–55)	29%[1]	x	1.94 (0–10)	2.7 (0–15)	14 yr
Rural Nyanza [56]						
FSWs	26 (17–37)	75%	2.5 (0–12)	x	6.6[2]	x
Trans-Africa Hwy [22, 41]						
FSWs	27 (16–46)	50–80%	13.6 (1–79)	129	25 (1–88)	x
Clients (truckers)	x	27%[1]	x	2.8	x	x
KwaZulu-Natal [50]						
FSWs	25 (15–48)	56%	x	x	x	2.5 yr
Clients (truckers)	37 (18–71)	56%	x	x	x	8.4 yr
Uganda [38]						
FSWs	x	33%	19.1	74.5	x	x
Clients (truckers)	x	18%	7.4	44.7	x	x

[1] This is the estimated prevalence among all sampled miners/truckers respectively, not just those who self-identified as clients.

[2] This statistic is the number of sexual acts reported over the past 14 days.

Table 3.1. *Comparison of summary statistics across five studies of FSWs and their clients in sub-Saharan Africa. The range is also given (where data exists). An x denotes no information.*

3.3 Population Structure

Perhaps the most significant yet unsurprising feature of such populations is that the FSWs on average have a greater number of different sex partners than do clients. The ratio of the number of different partners varies significantly, with FSWs reporting between roughly double to almost fifty times the number of different partners than do the clients. In the Carletonville-Mosuthumpilo data, FSWs report on average approximately ten times as many different sexual partners over the past year than clients report. The data from Ferguson *et al.* [22] suggest that the majority of clients are likely strangers. This is echoed in interviews with FSWs in Campbell [9].

In our network model, the degree of a node is interpreted to be the number of distinct sexual partners; the above observation thus implies that the FSWs have a greater mean degree. While it is widely accepted that the distribution of sexual contacts approximately follows a power law [35, 51], one cannot assume that the characteristics of sexual interactions between sex workers and clients is the same as that modeled by typical sexual networks. We find that the degree distribution for clients from the Carletonville data is well-characterized by a power law distribution with an exponent of almost 2 as shown in Figure 3.2. The degree distribution of the FSWs, however, is quite different: we see a peak near degree 4 as shown in Figure 3.1, and the data is fitted well by a Poisson distribution for degrees less than about 10. However, there are a number of women reporting high numbers of sexual partners who are not captured by this distribution. These responses, which are grouped at round numbers, are indicative of recall bias, making the FSW responses difficult to interpret. Therefore, we do not propose that Figure 3.1 is a fit to the Carletonville-Mosuthumpilo data; we simply observe that, unlike the client responses (and the other men and women responses shown in

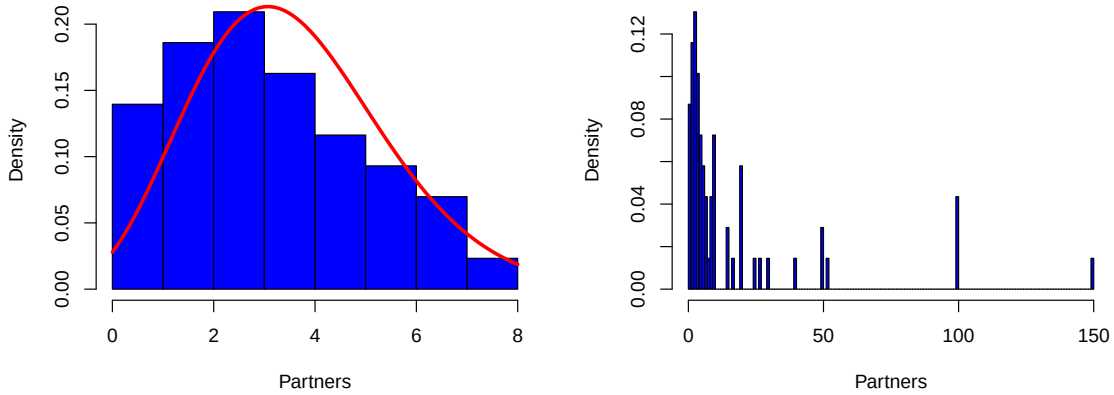


Figure 3.1. *(Left) Histogram of the number of different non-regular partners reported by FSWs in the past year where responses with more than 8 partners are omitted. A Poisson distribution fit to the data is shown in red. (Right) Full distribution of responses excluding one response of 300 partners.*

Figure A.2) there is a peak away from zero and that a Poisson distribution is a reasonable fit for this truncated data set.

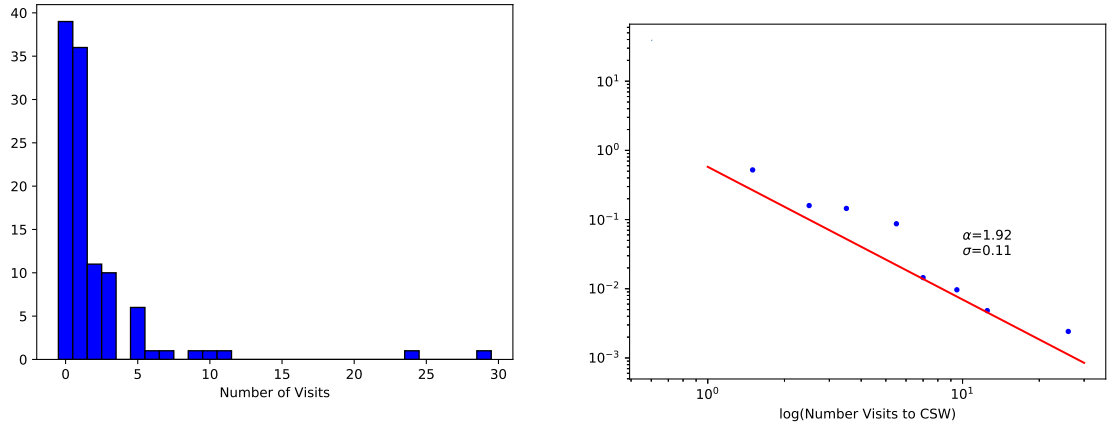


Figure 3.2. *Client responses for the number of visits to a sex worker over the past year. To the right, a power law fit with exponent α and error σ excluding the zero-th bin. Based on a Maximum Likelihood Estimation as implemented by [13]. The MLE method has been demonstrated to be superior than graphical methods (e.g. fitting a line in log-space).*

Chapter 4

Model and Results

4.1 Mathematical Model

There are two main components of our model; the disease process which describes how the disease progresses within an individual, and the contact process which defines how the disease spreads among the population. The bipartite network on which HIV transmission takes place is described in Section 4.1.1. Following transmission, the disease progresses independently and stochastically within each node according to the model specified in Section 4.1.2. The resulting model is calibrated as described in Section 4.2.

4.1.1 Network Model

To construct the contact network, the data from the Carletonville-Mosuthumpilo project are used to infer degree distributions for each population. We first perform a fit to the client distribution as follows. We assume that the distribution follows a power law, and take the exponent to be approximately 2 as suggested in Figure 3.2. Then, an exponential cutoff is chosen so that the maximum degree does not exceed $k = 30$. This gives the client distribution

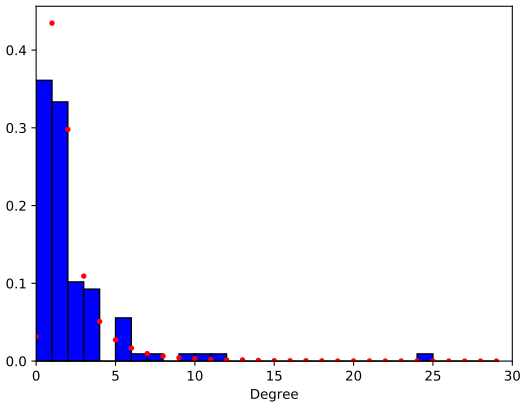
$$p_c(k) = \frac{e^{-k/6}k^{-2}}{Li_2(e^{-1/6})} \tag{4.1}$$

with mean degree λ, where the normalization constant $Li_2(e^{-1/6})$ is given in the form of the polylogarithm function. If we infer the FSW distribution by looking at the responses for the number on non-regular partners over the past year, however, the result is harder to characterize (Figure 4.1b). Instead, if we assume that each sex worker is equally likely to be chosen as a partner, then the corresponding degree distribution must be a Poisson distribution. Let

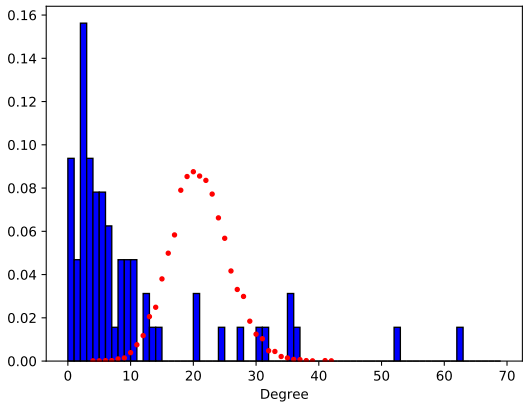
$$\mu = \frac{m}{n}\lambda.$$

Then

$$p_f(k) = \binom{m}{k}\left(\frac{\mu}{n}\right)^k\left(1-\frac{\lambda}{n}\right)^{m-k}.$$



(a) *Client responses (blue) for the number of visits to a FSW over the past year. In red, an example of a client distribution after generating a model bipartite network. The input client degree distribution is* p_c (4.1).



(b) *FSW responses (blue) for the number of non-regular partners over the past year. In red, an example of a FSW degree distributing after generating a model bipartite network. The input FSW degree distribution is* p_f (4.2).

Figure 4.1

If we further take the limit as $m \to \infty$ with $\frac{m}{n}$ constant, the female degree distribution becomes

$$p_f(k) = \frac{\mu^k e^{-\mu}}{k!}. \tag{4.2}$$

After generating each degree sequence, stochastic bipartite networks of m clients and n FSWs can then be generated using the Configuration Model [44], as described in Section 2.4.

As can be seen in Figure 4.1b, our model female degree distribution is not a good fit to the data from the Carletonville-Mosuthumpilo project. Our goal is not to fit the data or to use the empirical degree sequences. Instead, our model distribution is primarily justified by qualitative research on survival sex worker–client interactions, which suggests that random selection is a significant underlying social dynamic. As described in [61], survival sex workers often face situations where their safety or welfare is at risk, and may show a lack of agency in choosing their sexual partners. These ideas are further echoed in Campbell [9].

We did choose fit the client distribution to the data (Figure 4.1a) since the clients are less prone to recall biases, and to fix the client-to-FSW population ratios at 10:1. Then, since we require

$$\frac{\mu}{\lambda} = \frac{m}{n},$$

the mean degree for the FSW distribution must be 10λ. We further noted in Section 3.3 that a Poisson distribution with a mean of approximately 3 was a reasonable fit for the truncated data set shown in Figure 3.1. Since we expect the mean degree to be much higher considering the data summarized in Table 3.1, what we have done here in our network model is shifted the distribution from Figure 3.1 to match the reported mean degree of about 20.

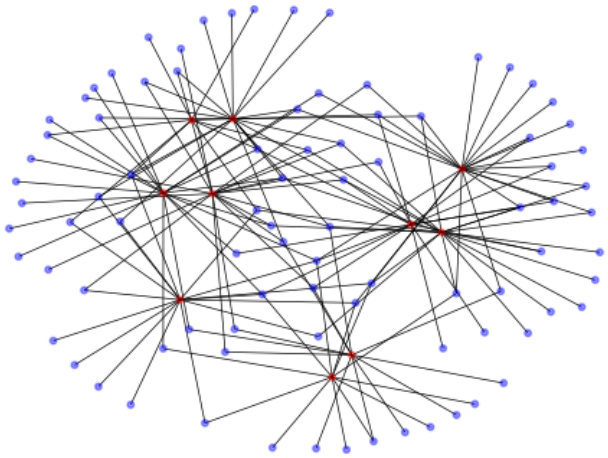


Figure 4.2. *An example of a model bipartite network with input degree distributions* $p_c(k)$ (4.1) *and* $p_f(k)$ (4.2)*. Client nodes are in blue and FSW nodes are in red, with a client population size of 100 and a FSW population size of 10. Isolate client nodes are not visualized.*

4.1.2 HIV Model

In Chapter 2, we discussed the toy SIR model on configuration model random graphs, and the corresponding derivation of analytical results such as the expected number of individuals who will become infected over time. It was further noted that HIV could be modelled as an SIR-type disease. However, in Section 3.2 we identified migration as being an important characteristic of our population. Therefore, in our model we will allow a node to become reinfected and interpret this as a migration event. The full disease dynamics are described below.

Each node of the bipartite random network is in one of five states: immune to HIV, susceptible to HIV infection, acute stage HIV infection, latent stage HIV infection, or on treatment with ART. These states are summarised in Table 4.1. Nodes in the susceptible state may become infected through a contact process on edges which connect to a nearest neighbour in either the acute state or the latent state. The contact process is Markovian, with a fixed probability β of disease transmission per unit time on each edge. Neighbours in the acute state are weighted more heavily due to the higher infectivity of this state.

Upon initial infection, nodes are in the acute state. Nodes transition from the acute state to the latent state with a fixed probability per unit time. Nodes in the latent state are randomly placed on treatment with a fixed probability per unit time. We make the simplifying assumption that nodes in the acute state cannot transition directly to the treated state. This is because the window period for HIV testing, combined with the time to achieve viral suppression after ART initiation, means a direct transition from the acute state to viral suppression under ART would be rare. The transition to the treated state in the model combines a positive HIV test, engagement in care, initiation of ART, and viral suppression.

Therefore, we will consider the treatment rate in the model to be a combination of the rates for each of these processes.

State	Interpretation
V	Vaccinated, or an immune state (cannot become infected)
S	Susceptible (may become infected)
A	Acute (highly infectious, short duration)
L	Latent (moderately infectious, long duration)
T	on Treatment (cannot become reinfected, noninfectious)

Table 4.1. *Possible node states in the disease model.*

Each node may be "removed" from the population with a probability per unit time that depends on the state of the node. This captures migration, natural death and death due to AIDS in the model. Under the static network assumption, the removed node is then replaced with an equivalent node (all edges remaining intact), which has a given probability α of being initialised in the latent stage, and probability $1-\alpha$ of being initialised as susceptible. The value of α is given by the background HIV prevalence in the general population, which is assumed to be fixed. For simplicity, we assume that nodes are not initialised in the relatively short acute stage. The state diagram for the client nodes is shown in Figure 4.3. The state diagram for FSWs is analogous, except that certain parameters governing state transitions take different values.

In our analysis, we interpret the "immune" state V as being a node which is currently taking PrEP. However, from an epidemiological standpoint it could equally well represent a vaccinated state. The primary difference lies in cost, because PrEP requires ongoing treatment to maintain immunity. We assume, for simplicity, that both PrEP and ART are 100% effective in preventing infection and transmission of HIV.

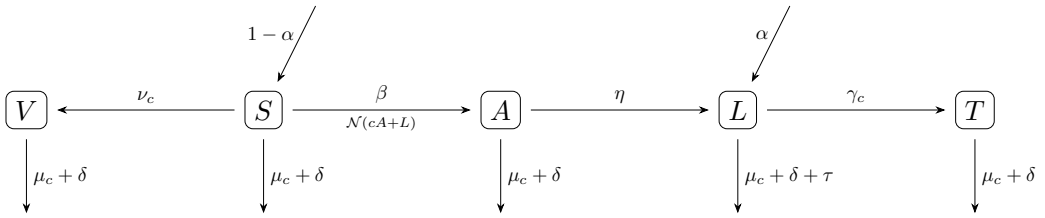


Figure 4.3. *Schematic of model progression in client nodes. The model structure for FSW nodes is identical, but the values of some parameters differ between client and FSW nodes (indicated by subscripts c and f, respectively). Parameters are explained and estimated in Table 4.2. The notation $\mathcal{N}(cA+L)$ refers to the total number of infectious nearest neighbours, combining those in the acute state (weighted by parameter c) and in the latent state.*

Parameter	Interpretation	Estimated Value
c	Relative infectivity of acute to latent individuals	$c = 10$ [59]
β	Infectious contact rate	0.06 (Section 4.2)
η^{-1}	Expected time spent in acute stage	10.5 weeks [14]
γ_c	Client treatment rate	Variable (Section 4.3.2)
γ_f	FSW treatment rate	Variable (Section 4.3.2)
μ_c	Client migration rate	0.077 (Table 3.1)
μ_f	FSW migration rate	0.25 (Table 3.1)
δ	Natural death rate	0.016 [18]
τ^{-1}	Expected time in latent stage if left untreated	10 yrs [14]
α	Background HIV prevalence	0.2 [60]
ν_c	Client "immunization" rate	Variable (Section 4.3.3)
ν_f	FSW "immunization" rate	Variable (Section 4.3.3)

Table 4.2. *Model parameters and estimated values. Rates are given per year.*

4.1.3 Simplifications and Assumptions

Both the network model and the disease model contain a number of assumptions. Transmission occurs with probability proportional to the number of a susceptible node's nearest neighbours in the acute and latent stages. The underlying assumption here is that patients who progress to the AIDS stage are not sexually active, and thus do not contribute to transmission events. The increase in viral load during the onset of AIDS is significant (Figure 1.1), however, and transmission events during this stage may contribute to the overall epidemic. Nodes are not moved onto treatment from the acute state. The length of time required to achieve complete (or near-complete) viral suppression on anti-retroviral therapy (ART) may be up to 6 months [43]. This far exceeds the expected time spent in the acute state. Nodes return to the model with intact edges (i.e. the static network assumption). This assumption is made for computational simplicity, and because there are no data which tracks the evolution of the social network over time. All partnerships are treated as concurrent; this is another consequence of having static degree distributions. Unfortunately, the data from the Carletonville-Mosuthumpilo project provide no indication on the timing of the reported sexual encounters. We can therefore interpret the infectious contact parameter β as both a function of the probability of transmission per sex act, and on the probability of there having been a sexual encounter during the time step.

4.2 Calibration

Most of the parameters in Table 4.2 have been estimated using public health sources or the data from the Carletonville-Mosuthumpilo project. It remains to estimate the infectious

contact parameter, β. To do this, we first assume that there is no asymmetry between the male-to-female and female-to-male contact rates. This assumption has been shown to be acceptable in sub-Saharan African populations [59]. Next, we calibrate the model under zero treatment rates as shown in Figure 4.4. A value of $\beta = 0.06$ results in an equilibrium disease prevalence which closely matches the recorded epidemic among the FSW population (see Table 3.1).

When $\beta = 0.06$, the corresponding client prevalence is 44% at equilibrium, which is more than 60% of the FSW prevalence, which appears reasonable given the discussion in Section 3.2. Unfortunately, if only 9% of miners are clients and the remaining miners have a 20% HIV prevalence, then these parameters would underestimate the total miner prevalence which was recorded to be 29% (Table 3.1). It is possible that the Carletonville-Mosuthumpilo study under-reported the number of miners who were clients, or perhaps that our background HIV prevalence α is too small.

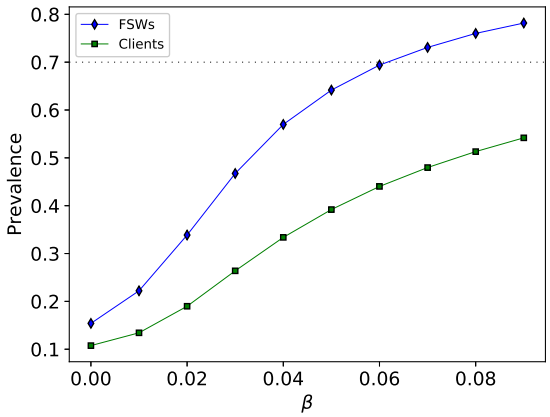


Figure 4.4. *Equilibrium HIV prevalence among each subpopulation under zero treatment. The dashed line marks the recorded FSW prevalence.*

4.3 Simulation Results

The model developed in the previous section is used to investigate possible control strategies. Each simulation is initialized with the equilibrium HIV prevalences found for zero treatment (see Figure 4.4), namely 69% and 44% for the FSW and client populations, respectively. We evaluate our interventions in two ways. First, we consider the total untreated HIV prevalence after 20 years, where total untreated HIV prevalence is defined to be the proportion of the total population (both clients and FSWs) in the acute stage HIV and latent stage HIV states at $t = 20$. HIV-positive individuals in the treated state are not included in our computation of HIV prevalence, because that would mask the impact of treatment scale-up on the epidemic.

Additionally, we look at the relative cost for each PrEP and treatment strategy to achieve a given HIV prevalence after 20 years. To compare the cost of PrEP and ART, we assume that the cost of ART is double that of PrEP. This is a conservative estimate, as the cost of ART may be up to 4 times that of PrEP [49]. The cost of PrEP is taken to be one unit per individual on PrEP per simulation time step, and so the cost of ART is two units per individual on ART per simulation time step. We do not consider targeted PrEP or ART strategies within each subpopulation. That is, when applying such strategies to FSWs and/or clients, nodes are placed uniformly at random on PrEP or ART according to a stationary Poisson process at the relevant rate.

For each experiment, we produce two graphs: one showing the effectiveness of each strategy as a function of either treatment rate or PrEP coverage (after 20 years), and a second showing the cost to achieve this prevalence after 20 years. Therefore, each point on the second graph can be mapped back to the first, and vice versa. For all graphs shown in this chapter, each marker denotes a simulated result (averaged over many random networks and multiple runs per network). The lines are included for visualization purposes, and are piecewise linear interpolants of the simulation data. These lines do not indicate a theoretical result.

4.3.1 Performing Simulations

As before, the random networks were generated using the software NetworkX [28] and our simulations were performed using NepidemiX [1], a software package for simulating stochastic processes on networks. Some minor modifications to NepidemiX were needed in order to initialize prevalence in each node set separately (more details are given in Appendix B).

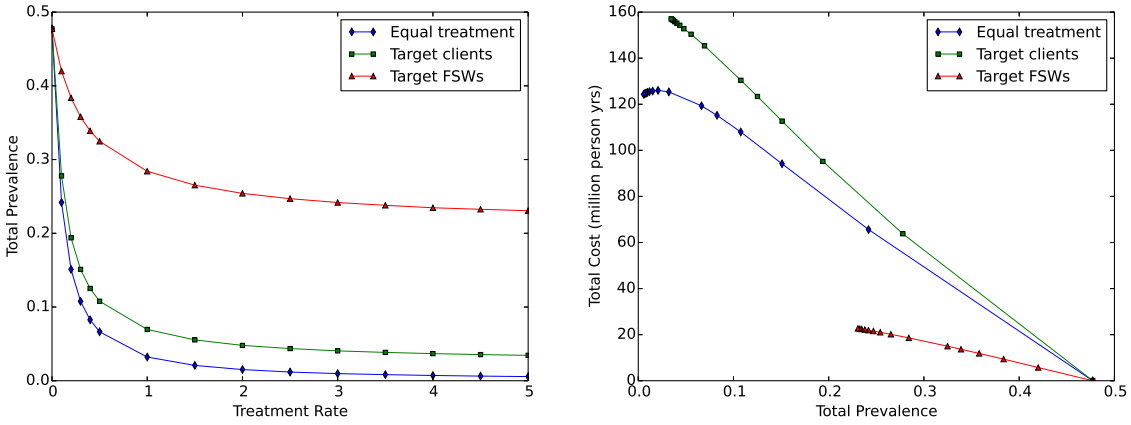
We average over many trials (at least 10) on many random networks (at least 40) for a range of parameter values (for a total of at least 400 samples for each parameter). In order to obtain relatively stable network statistics, our model population needs to be appropriately large (following the discussion in Section 2.4). In considering the accuracy of our numerical results in Section 2.4.2, we use node sets of sizes 90,000 and 9,000 for clients and FSWs, respectively, and a time step of $\Delta t = 0.1$/yr. A simulation of an HIV epidemic on a network of this size for 200 time steps currently requires about 10 minutes on a standard machine. Each experiment was carried out on the Compute Canada cluster Cedar and typically took about three days to run. This runtime depends strongly on the degree to which the computations can be parallelized, which is dictated by the resources available.

4.3.2 Treatment-Based Interventions

We first investigate the effect of treatment-only interventions, without implementing PrEP ($\nu_f = \nu_c = 0$). Clients and FSWs with latent stage HIV are placed on ART at rates γ_c and γ_f, respectively. Three treatment-based intervention strategies are evaluated: (1) Treat both populations with an equal treatment rate; (2) Target treatment to the FSW population;

(3) Target treatment to the client population. A comparison of these three interventions is shown in Figure 4.5.

The impact of scaling up treatment rates from 0 to 5/yr is shown in Figure 4.5a. A treatment rate of 5/yr corresponds to an expected time to treatment of approximately two and a half months. The graph in Figure 4.5a shows that the most significant reductions in total prevalence occur when the treatment rate is increased from 0 to 1/yr, which corresponds to a mean time to treatment of one year after infection. Further increasing the treatment rate only has a minimal impact on reducing prevalence. The persistence of HIV in the model, even for high treatment rates, occurs because of new infections from individuals in the acute stage HIV state and through migration into the model of HIV-positive individuals. The latter effect also explains why treatment of only FSWs results in the highest prevalence: the migration rate for FSWs is approximately 3.3 times greater than the migration rate for clients in this model.



(a) *Effect of treatment on total HIV prevalence after 20 yrs.*

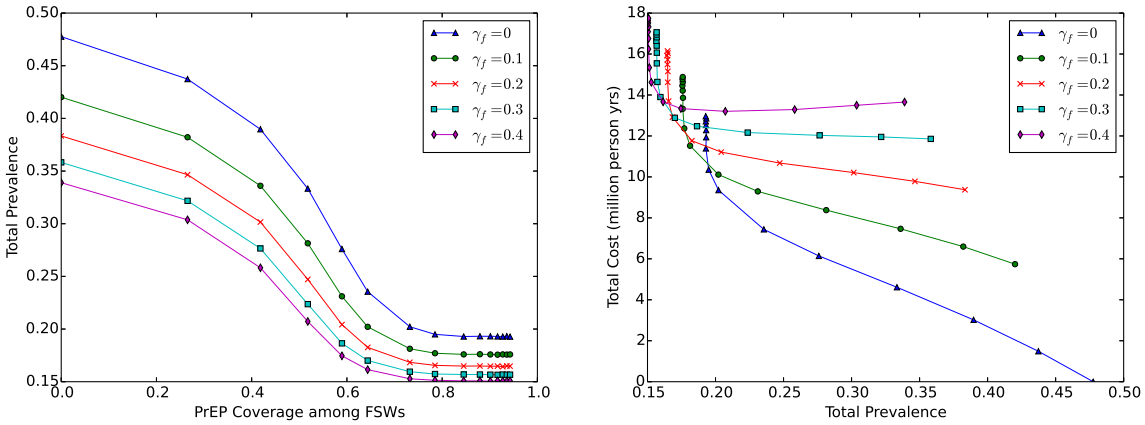
(b) *Total cost over 20 yrs to achieve a given total prevalence.*

Figure 4.5. *Comparison of three treatment strategies over a 20-yr period.*

Figure 4.5b shows the total cost over 20 years required to achieve a given HIV prevalence. The lowest-cost strategy is to target treatment to only the FSW subpopulation. This is primarily due to the FSW population size being one-tenth of the client population size. However, this strategy can only reduce total prevalence down to approximately 23%. To reduce the total prevalence below this threshold, the lowest-cost strategy is to treat both subpopulations.

4.3.3 Combined Interventions

We now consider interventions which combine both PrEP and ART. Susceptible nodes transition to the PrEP state V at the constant rates ν_c for clients and ν_f for FSWs. Nodes in the latent stage HIV state transition to the treated state at the constant treatment rates



(a) *Effect of the combined strategy on total HIV prevalence after 20 yrs.*

(b) *Total cost over 20 yrs to achieve a given total prevalence.*

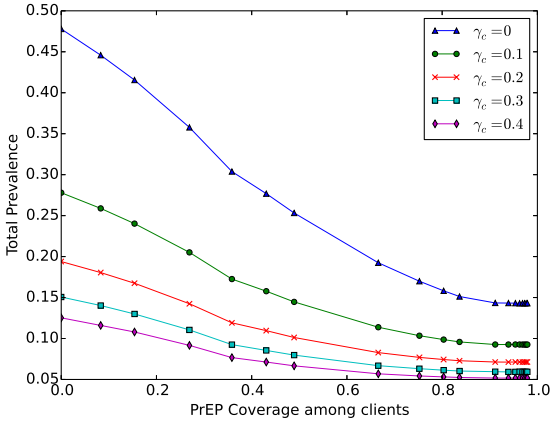
Figure 4.6. *Comparison of combined treatment and prevention strategies over a 20-yr period, when both treatment and PrEP are targeted towards FSWs.*

γ_c for clients and γ_f for FSWs. PrEP coverage is defined to be the proportion of HIV-negative individuals on PrEP after 20 yrs. The following combined strategies are evaluated: (1) treat and administer PrEP to FSWs only; (2) treat FSWs and administer PrEP to clients; (3) treat clients and administer PrEP to FSWs; and (4) treat and administer PrEP to clients only. We additionally investigate the effect of treating both populations while administering PrEP to FSWs only.

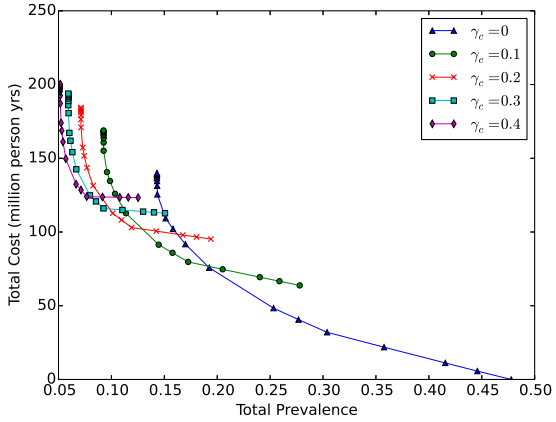
When both ART and PrEP are targeted at the FSW population, as shown in Figure 4.6, there is a threshold PrEP coverage of approximately 80%, beyond which there is not a significant additional decrease in HIV prevalence. Figure 4.6b shows that focusing only on PrEP is the lowest-cost option for reducing total prevalence to approximately 20%. PrEP should be combined with treatment to reduce prevalence to below 20% in this population.

Figure 4.7 shows the impact of targeting both treatment and PrEP to the client subpopulation alone. The response of the total prevalence to increasing PrEP coverage in the client subpopulation is more gradual than for the FSW subpopulation so that higher coverage is required to realise the lowest prevalence. Figure 4.7b shows a larger window in which combining treatment with PrEP is the lowest-cost solution.

The next two interventions considered target both populations. Figure 4.8 shows that if we consider treating only the FSW population, it is never cost-effective to prioritize PrEP to clients. Figure 4.9 shows the opposite case where we treat only clients and provide PrEP to FSWs. Here, we see that for any given total prevalence target, it is cost-beneficial to re-allocate testing and treatment resources from the clients toward increasing PrEP coverage in the FSWs. However, if the targeted total prevalence is below approximately 20%, then a testing and treatment programme for the clients is necessary.

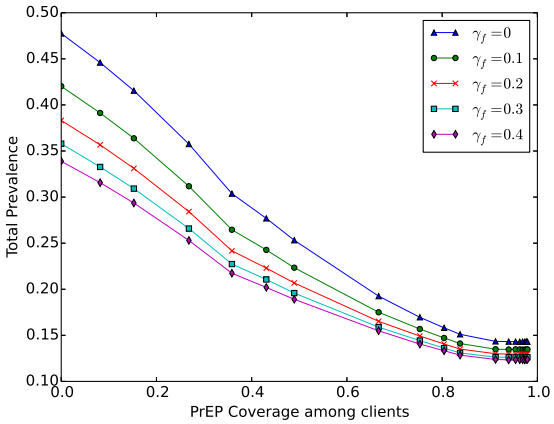


(a) *Effect of the combined strategy on total HIV prevalence after 20 yrs.*

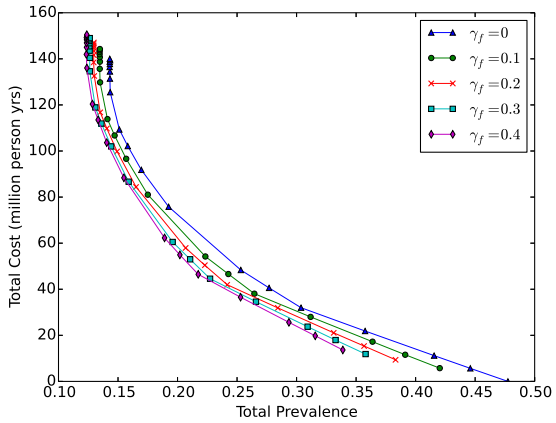


(b) *Total cost over 20 yrs to achieve a given total prevalence.*

Figure 4.7. *Comparison of combined treatment and prevention strategies over a 20-yr period, when PrEP and treatment are both targeted towards clients.*

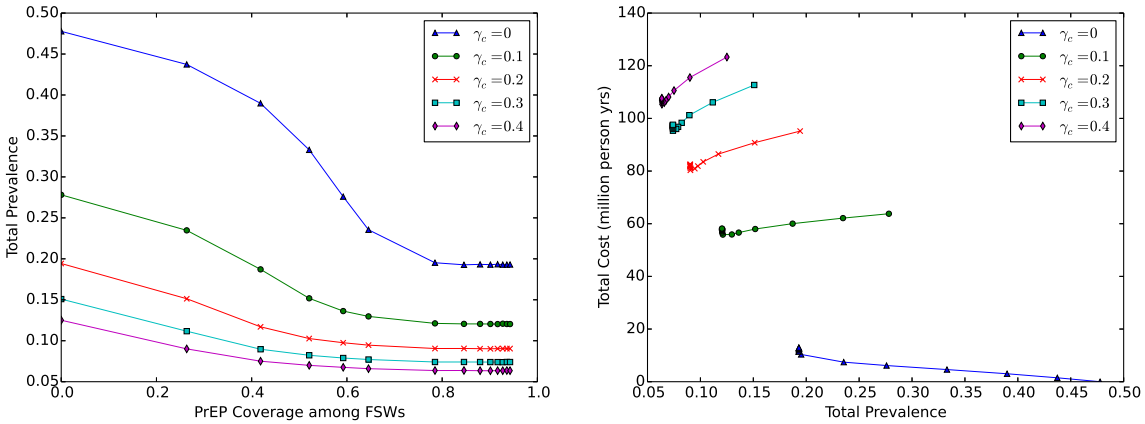


(a) *Effect of the combined strategy on total HIV prevalence after 20 yrs.*



(b) *Total cost over 20 yrs to achieve a given total prevalence.*

Figure 4.8. *Comparison of combined treatment and prevention strategies over a 20-yr period, when PrEP is targeted towards clients and treatment is targeted towards FSWs.*



(a) *Effect of the combined strategy on total HIV prevalence after 20 yrs.*

(b) *Total cost over 20 yrs to achieve a given total prevalence.*

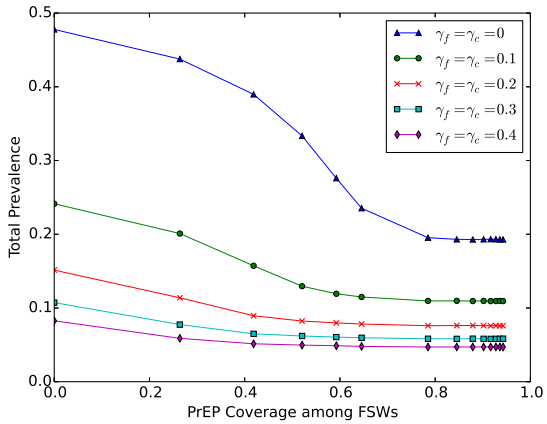
Figure 4.9. *Comparison of combined treatment and prevention strategies over a 20-yr period, when PrEP is targeted towards FSWs and treatment is targeted towards clients.*

The final intervention considered is arguably the most realistic from a public health policy perspective. In this case, testing and treatment programmes are implemented for both FSWs and clients; however, PrEP is targeted to only FSWs. The results for this scenario are shown in Figure 4.10. As with the scenario in which treatment is only targeted to clients, we find that for a given prevalence target it remains cost-advantageous to prioritize PrEP for FSWs. Figure 4.10b shows that the cost advantage of prioritizing PrEP is even slightly greater when treatment is applied to both subpopulations. However, it remains the case that a treatment programme is required to achieve total HIV prevalence below approximately 20%.

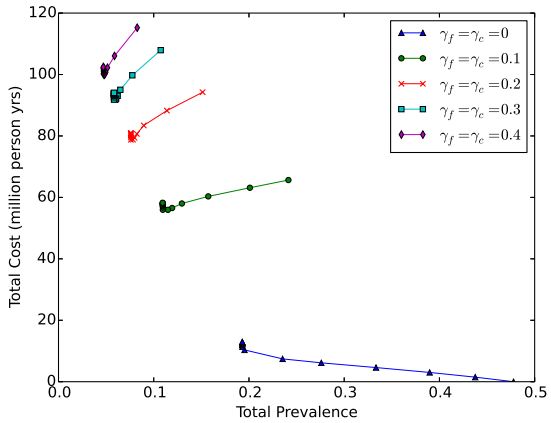
4.3.4 Sensitivity Analysis

In our discussion in the previous section, we hypothesize that the most significant reason that a PrEP program would not be a cost-effective strategy when targeted at the clients is due to the power-law degree distribution, not just the relative difference in the mean degrees between clients and FSWs. A power-law distribution exhibits high density around low degrees, which means that a randomized strategy would be unlikely to reach all of the high degree nodes. It would be much more efficient to target the FSW population, who exhibit a Poisson degree distribution in which nodes are densely distributed about the mean degree. However, it is unclear how much of this effect is due to the degree distribution and how much is due to the mean degrees.

Therefore, we wish to repeat some of the experiments of the previous section using a Poisson distribution for the clients and a truncated power-law distribution for the FSWs, while keeping the mean degrees the same (on average) as those of the previous section.



(a) *Effect of the combined strategy on total HIV prevalence after 20 yrs.*



(b) *Total cost over 20 yrs to achieve a given total prevalence.*

Figure 4.10. *Comparison of combined treatment and prevention strategies over a 20-yr period, when PrEP is targeted towards FSWs and treatment is provided equally to both clients and FSWs.*

When we calibrate the model to determine the appropriate value of β, we find very different behaviour than that shown in Figure 4.4. In particular, for $\beta > 0$, we see that the client HIV prevalence is greater than the FSW prevalence. Therefore, this does not appear to be a reasonable model. In order to compare our results, therefore, we need to look at different combinations of degree distributions; this is an area for future work.

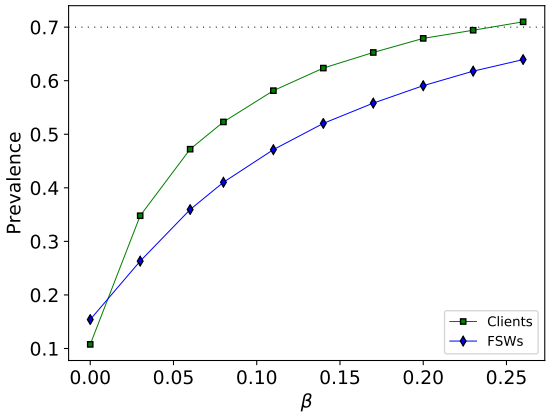


Figure 4.11. *Equilibrium HIV prevalence among each subpopulation under zero treatment in the modified model where clients have a Poisson distribution and FSWs have a power-law distribution. The dashed line marks the recorded FSW prevalence.*

We also wish to investigate the effect of lowering the migration rate on our results from Section 4.3.3. In Figure 4.12, the minimum achievable HIV prevalence is only slightly reduced from that shown in Figure 4.6. Therefore, a change in migration rate would not

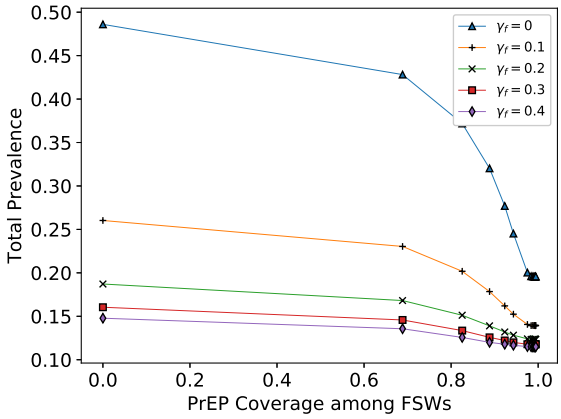


Figure 4.12. *Effect of the combined strategy on total HIV prevalence after 20 yrs, when FSW migration rate is set to zero.*

qualitatively affect our results. This furthermore illustrates that the importance in this population of acute stage transmissions.

Of course, a full sensitivity analysis of all parameters in the model would be required in order to determine the confidence of our qualitative results. Due to the computationally intensive nature of these simulations, we need to prioritize those parameters which we believe may impact our results the most. Foremost, as mentioned above, we wish to conduct a thorough sensitivity analysis on the two degree distributions. Secondly, the impact of network size on our results may be significant; however, if it does not, it may be possible to decrease simulation time. Lastly, we could investigate the impact of changing the infectious contact parameter β, perhaps allowing for asymmetric transmission.

Chapter 5

Discussion and Conclusions

The sexual network structure of survival female sex worker–client populations has several features which distinguish it from other sexual networks. We have examined the potential importance of these features in designing control strategies for an HIV epidemic. The sexual network is modelled as a bipartite random network in which the FSW subpopulation is significantly smaller than the client subpopulation. Other key features of our model are the higher migration rate for female sex workers and the nature of the degree distribution for the sex workers, arising from their limited agency in sexual negotiations. Each of these characteristics affect the impact of the control strategies which were tested.

Intuitively, one would expect that the most effective strategy for control of an HIV epidemic on a bipartite random network would be to target both treatment and prevention strategies to the smaller subpopulation—in our case the FSWs. Certainly for an SIR-type disease process, there would exist a threshold vaccine coverage above which the epidemic would be suppressed (as we discussed in Section 2.4).

This intuition is only partially supported by our model. We find that, up to a certain point, the lowest-cost strategy for achieving a given prevalence at any treatment rate is to prioritise PrEP in the FSW population. The caveat is that targeting only the FSW subpopulation cannot reduce the overall HIV prevalence to near-zero levels. The lowest-cost strategy for further reductions in prevalence is to combine PrEP in the FSW subpopulation with expansion of testing and treatment in both the FSW and client subpopulations.

Our observation that targeting interventions towards FSWs can not completely suppress disease dynamics appears to be primarily driven by acute infection events. The high mean degree of our model FSWs relative to their clients means that once a given female node is infected, she is expected to have many concurrent contacts with clients. Since the probability of transmission is much higher during this initial (untreated) stage, this may lead to a large number of infected clients. Therefore, when these clients are not moved onto treatment, a significant proportion of the total population remains HIV-positive.

Both the mean degree and the distribution of the number of sexual partners in the FSW subpopulation play a role in the effectiveness of PrEP. In the model, we assumed that this

degree distribution is a Poisson distribution and that this subpopulation has a significantly larger mean degree than the clients. Hence, there is a large clustering of FSW nodes around the mean degree: almost 90% of the female nodes in our generated networks have degree between 15 and 30. Since individuals are drawn randomly from each subpopulation for immunization, the likelihood is high that a female node chosen to be put onto PrEP has high degree, and therefore has a large influence in the network.

We also find that PrEP strategies which include or target the client population are generally not cost-effective strategies. In some cases, they may be significantly less so than treatment alone. The argument for why this is the case follows from the discussion above: under random selection, the likelihood is low that a chosen client node would have a high influence in the network. This is not only due to the low mean degree of client nodes relative to the FSWs, but also due to the large variance in degrees. In this model, we have included only clients who were active in the network. In reality, a client-based PrEP strategy would likely be even less cost-effective, because of the difficulty in identifying individuals who are actively involved with female sex workers.

5.1 Summary

We have presented a stochastic model which describes sex worker–client interactions in a high HIV prevalence setting. Using a random bipartite network model, we have investigated the relative cost-effectiveness of different HIV control strategies involving the administration of pre-exposure prophylaxis (PrEP, for prevention) and/or antiretroviral therapy (ART, for treatment). These resources are targeted to either clients or survival female sex workers (FSWs) variably, although within each subpopulation, we only assume that individuals are chosen randomly for either PrEP or ART.

The most effective strategy for reducing HIV prevalence in a survival female sex worker–client population is a PrEP program targeted at FSWs. As PrEP coverage of the FSWs is scaled up, the projected HIV prevalence after 20 years decreases dramatically until a threshold coverage is reached. Further expansion of PrEP coverage above the threshold level does not lead to significant further reduction in HIV prevalence. We find that a PrEP program targeted at the clients is not a cost-effective approach for reducing HIV prevalence.

A PrEP program alone for the female sex workers cannot effectively eliminate HIV as a public health threat in the sex worker–client population. The high number of concurrent partnerships typically associated with survival sex workers, combined with a high migration rate, limits the ability of a PrEP strategy to drive HIV prevalence close to zero. An aggressive reduction in prevalence requires a combination strategy of PrEP for the female sex workers and expanded testing and treatment for both the sex workers and the clients. In general, intervention strategies should be prioritized to the sex workers, but testing and treatment of the client population should not be neglected.

Our model is motivated by and calibrated to data from the Carletonville-Mosuthumpilo project—a South African study from 2000. However, we believe that our results are applicable to general survival sex worker populations. Key characteristics of our models include the relative differences in migration rates and mean number of partnerships (a consequence of the ratio of population sizes) and on the degree distributions of both populations. Here, the clients on average have a lower migration rate and have significantly fewer partners than the female sex workers; we find evidence from different data sources to support these ideas. We then assume that clients choose their partners randomly. While this generative mechanism seems plausible for a survival sex worker network, we acknowledge that the evidence to support the female degree distribution is lacking. The remaining parameters which were estimated from or calibrated to the Carletonville data are not expected to qualitatively affect our conclusions.

5.2 Future Work

Further improvements upon our model may include distinguishing between casual and regular contacts, and taking into account risk compensation or adherence. Whether or not risk compensation would be an issue in the context of PrEP is a subject of debate. In Eakle *et al.* [20], the female sex workers interviewed acknowledge the continued importance of condoms as a preventative measure, and that PrEP should be used as a back-up measure. A previous study on post-exposure prophylaxis (PEP) showed no increase in high-risk behaviour [37].

The conclusion that a prevention-based strategy would be ineffective when targeted at the client population is likely sensitive to the client degree distribution. The power-law degree distribution was inferred using questionnaire data from the Carletonville-Mosuthumpilo Project. Furthermore, as discussed above, while random selection seems like a good starting point to model the degree distribution of survival female sex workers, we do not currently have the data to support this distribution. Self-reported questionnaire data has been shown to be an unreliable indication of sexual contacts [25] and so we suggest the following areas of future work: the first is to collect and analyze reliable social network data among sex workers and their clients. Secondly, one could expand upon our sensitivity analysis with regards to the model network structure. Finally, the Poisson distribution considered for the female degree distribution assumed that clients chose sex workers randomly. We acknowledge that this is a simplistic generative mechanism, and that more complex social dynamics could be explored by using alternative models for random graph generation.

One could furthermore address the issue of concurrency. In our static network model, all possible partnerships of a given node are considered to be concurrent, which is not always a realistic assumption. We attempt to skirt this issue by lowering our transmission probability per sexual act, β. However, by using a dynamic network—one in which edges can be broken and/or rearranged—one would be able to consider serial partnerships. Analytical

results, similar to our discussion in Chapter 2, exist for simple disease processes on dynamic networks [58], however additional model complexity such as migration would likely prohibit the derivation of analytical results. Additionally, a mixture of static and dynamic edges could be used to model regular and non-regular partnerships, respectively. Unfortunately, such a model would be difficult to calibrate and validate, since, to our knowledge, research and published data showing the temporal evolution of social networks is scarce.

Bibliography

[1] L. Ahrenberg. NepidemiX. Technical report, Simon Fraser University, 2016.

[2] S. Attia, M. Egger, M. Müller, M. Zwahlen, and N. Low. Sexual transmission of HIV according to viral load and antiretroviral therapy: systematic review and meta-analysis. *AIDS*, 23:1397–1404, 2009.

[3] J. M. Baeten, D. Donnell, P. Ndase, N. R. Mugo, J. D. Campbell, J. Wangisi, J. W. Tappero, E. A. Bukusi, C. R. Cohen, E. Katabira, et al. Antiretroviral prophylaxis for HIV prevention in heterosexual men and women. *New England Journal of Medicine*, 367:399–410, 2012.

[4] S. Baral, C. Beyrer, K. Muessig, T. Poteat, A. L. Wirtz, M. R. Decker, S. G. Sherman, and D. Kerrigan. Burden of HIV among female sex workers in low-income and middle-income countries: a systematic review and meta-analysis. *The Lancet Infectious Diseases*, 12:538–549, 2012.

[5] P. Bearman, J. Moody, and K. Stovel. Chains of affection: The structure of adolescent romantic and sexual networks. *American Journal of Sociology*, 110:44–91, 2004.

[6] L.-G. Bekker, L. Johnson, F. Cowan, C. Overs, D. Besada, S. Hillier, and W. Cates. Combination HIV prevention for female sex workers: what is the evidence? *The Lancet*, 385:72–87, 2015.

[7] C. Beyrer, A.-L. Crago, L.-G. Bekker, J. Butler, K. Shannon, D. Kerrigan, M. R. Decker, S. D. Baral, T. Poteat, A. L. Wirtz, et al. An action agenda for HIV and sex workers. *The Lancet*, 385:287–301, 2015.

[8] D. Bisanzio, L. Bertolotti, L. Tomassone, G. Amore, C. Ragagli, A. Mannelli, M. Giacobini, and P. Provero. Modeling the spread of vector-borne diseases on bipartite networks. *PLoS One*, 5(11):e13796, 2010.

[9] C. Campbell. *Letting Them Die: Why HIV/AIDS Intervention Programmes Fail*. Indiana University Press, 2003.

[10] P. J. Carrington, J. Scott, and S. Wasserman. *Models and Methods in Social Network Analysis*. Cambridge University Press, 2005.

[11] S. Chen and X. Lu. An immunization strategy for hidden populations. *Scientific Reports*, 7:2045–2322, 2017.

[12] Y. Chen, G. Paul, S. Havlin, F. Liljeros, and H. E. Stanley. Finding a better immunization strategy. *Physical Review Letters*, 101:058701, 2008.

[13] A. Clauset, C. R. Shalizi, and M. E. J. Newman. Power-law distributions in empirical data. *SIAM Review*, 51:661–703, 2009.

[14] J. M. Coffin, S. H. Hughes, et al. *Retroviruses*. Cold Spring Harbor Laboratory Press, 1997.

[15] R. Cohen, S. Havlin, and D. ben-Avraham. Efficient immunization strategies for computer networks and populations. *Physical Review Letters*, 91:247901, 2003.

[16] M. R. Decker, A.-L. Crago, S. K. Chu, S. G. Sherman, M. S. Seshu, K. Buthelezi, M. Dhaliwal, and C. Beyrer. Human rights violations against sex workers: burden and effect on HIV. *The Lancet*, 385:186–199, 2015.

[17] L. Decreusefond, J.-S. Dhersin, P. Moyal, V. C. Tran, et al. Large graph limit for an SIR process in a random network with hetereogeneous connectivity. *The Annals of Applied Probability*, 22:541–575, 2012.

[18] R. Dorrington, D. Bourne, D. Bradshaw, et al. *The impact of HIV/AIDS on adult mortality in South Africa*. Medical Research Council Cape Town, 2001.

[19] R. Durrett. *Random Graph Dynamics*. Cambridge Series in Statistical and Probabilistic Mathematics, 2007.

[20] R. Eakle, A. Bourne, J. Mbogua, N. Mutanha, and H. Rees. Exploring acceptability of oral PrEP prior to implementation among female sex workers in South Africa. *Journal of the International AIDS Society*, 21, 2018.

[21] P. Erdős and A. Rényi. On random graphs. *Publicationes Mathematicae*, 6:290–297, 1959.

[22] A. G. Ferguson, C. N. Morris, and C. W. Kariuki. Using diaries to measure parameters of transactional sex: an example from the trans-Africa highway in Kenya. *Culture, Health & Sexuality*, 8:175–185, 2006.

[23] G. B. Gomez, A. Borquez, K. K. Case, A. Wheelock, A. Vassall, and C. Hankins. The cost and impact of scaling up pre-exposure prophylaxis for HIV prevention: A systematic review of cost-effectiveness modelling studies. *PLOS Medicine*, 10:1–16, 2013.

[24] J. Gómez-Gardeñes, V. Latora, Y. Moreno, and E. Profumo. Spreading of sexually transmitted diseases in heterosexual populations. *Proceedings of the National Academy of Sciences*, 105(5):1399–1404, 2008.

[25] C. A. Graham, J. A. Catania, R. Brand, T. Duong, and J. A. Canchola. Recalling sexual behavior: A methodological analysis of memory recall bias via interview using the diary as the gold standard. *The Journal of Sex Research*, 40:325–332, 2003.

[26] R. H. Gray, G. Kigozi, D. Serwadda, F. Makumbi, S. Watya, F. Nalugoda, N. Kiwanuka, L. H. Moulton, M. A. Chaudhary, M. Z. Chen, et al. Male circumcision for HIV prevention in men in Rakai, Uganda: a randomised trial. *The Lancet*, 369:657–666, 2007.

[27] J. M. Grund, T. S. Bryant, I. Jackson, et al. Association between male circumcision and women's biomedical health outcomes: a systematic review. *The Lancet Global Health*, 5:e1113–e1122, 2017.

[28] A. A. Hagberg, D. A. Schult, and P. J. Swart. Exploring network structure, dynamics, and function using NetworkX. In *Proceedings of the 7th Python in Science Conference (SciPy2008)*, pages 11–15, Pasadena, CA USA, 2008.

[29] T. D. Hollingsworth, R. M. Anderson, and C. Fraser. HIV-1 transmission, by stage of infection. *The Journal of Infectious Diseases*, 198:687–693, 2008.

[30] S. Janson, M. Luczak, and P. Windridge. Law of large numbers for the SIR epidemic on a random graph with given degrees. *Random Structures and Algorithms*, 45, 2014.

[31] E. Kenah and J. M. Robins. Second look at the spread of epidemics on networks. *Physical Review E*, 76:036113, 2007.

[32] D. Kerrigan, A. Wirtz, I. Semini, N. N'Jie, A. Stanciole, J. Butler, R. Oelrichs, and C. Beyrer. *The Global HIV Epidemics among Sex Workers*. The World Bank, 2012.

[33] K. Y. Leung and M. Kretzschmar. Concurrency can drive an HIV epidemic by moving R0 across the epidemic threshold. *AIDS*, 29, 2015.

[34] F. Liljeros, C. R. Edling, and L. A. N. Amaral. Sexual networks: implications for the transmission of sexually transmitted infections. *Microbes and Infection*, 5:189–196, 2003.

[35] F. Liljeros, C. R. Edling, L. A. N. Amaral, H. E. Stanley, and Y. Åberg. The web of human sexual contacts. *Nature*, 411:907–908, 2001.

[36] C. M. Lowndes, M. Alary, C. A. Gnintoungbé, E. Bédard, L. Mukenge, N. Geraldo, P. Jossou, E. Lafia, F. Bernier, É. Baganizi, et al. Management of sexually transmitted diseases and HIV prevention in men at high risk: targeting clients and non-paying sexual partners of female sex workers in Benin. *AIDS*, 14:2523–2534, 2000.

[37] J. N. Martin, M. E. Roland, T. B. Neilands, M. R. Krone, J. D. Bamberger, R. P. Kohn, M. A. Chesney, K. Franses, J. O. Kahn, T. J. Coates, et al. Use of postexposure prophylaxis against HIV infection following sexual exposure does not lead to increases in high-risk behavior. *AIDS*, 18:787–792, 2004.

[38] J. K. B. Matovu and B. N. Ssebadduka. Sexual risk behaviours, condom use and sexually transmitted infection treatment-seeking behaviours among female sex workers and truck drivers in Uganda. *International Journal of STD & AIDS*, 23:267–273, 2012.

[39] L. A. Meyers, M. Newman, and B. Pourbohloul. Predicting epidemics on directed contact networks. *Journal of theoretical biology*, 240(3):400–418, 2006.

[40] J. S. Montaner, R. Hogg, E. Wood, T. Kerr, M. Tyndall, A. R. Levy, and P. R. Harrigan. The case for expanding access to highly active antiretroviral therapy to curb the growth of the HIV epidemic. *The Lancet*, 368:531–536, 2006.

[41] C. N. Morris and A. G. Ferguson. Estimation of the sexual transmission of HIV in Kenya and Uganda on the trans-Africa highway: the continuing role for prevention in high risk groups. *Sexually Transmitted Infections*, 82:368–371, 2006.

[42] M. Morris and M. Kretzschmar. Concurrent partnerships and the spread of HIV. *AIDS*, 11:641–648, 1997.

[43] A. Mujugira, C. Celum, R. W. Coombs, J. D. Campbell, P. Ndase, A. Ronald, E. Were, E. A. Bukusi, N. Mugo, J. Kiarie, et al. HIV transmission risk persists during the first 6 months of antiretroviral therapy. *JAIDS*, 72:579–584, 2016.

[44] M. E. J. Newman. Spread of epidemic disease on networks. *Physical Review E*, 66, 2002.

[45] M. E. J. Newman, S. H. Strogatz, and D. J. Watts. Random graphs with arbitrary degree distributions and their applications. *Physical Review E*, 64:026118, 2001.

[46] F. Nian and X. Wang. Efficient immunization strategies on complex networks. *Journal of Theoretical Biology*, 264:77–83, 2010.

[47] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani. Epidemic processes in complex networks. *Reviews of Modern Physics*, 87:925–979, 2015.

[48] S. D. Pinkerton and P. R. Abramson. Effectiveness of condoms in preventing HIV transmission. *Social Science and Medicine*, 44:1303–1312, 1997.

[49] C. Pretorius, J. Stover, L. Bollinger, N. Bacaër, and B. Williams. Evaluating the cost-effectiveness of pre-exposure prophylaxis (PrEP) and its impact on HIV-1 transmission in South Africa. *PLOS ONE*, 5:1–10, 2010.

[50] G. Ramjee and E. Gouws. Prevalence of HIV among truck drivers visiting sex workers in KwaZulu-Natal, South Africa. *Sexually Transmitted Diseases*, 29:44–49, 2002.

[51] A. Schneeberger, C. H. Mercer, S. A. Gregson, N. M. Ferguson, C. A. Nyamukapa, R. M. Anderson, A. M. Johnson, and G. P. Garnett. Scale-free networks and sexually transmitted diseases: a description of observed patterns of sexual contacts in Britain and Zimbabwe. *Sexually Transmitted Diseases*, 31:380–387, 2004.

[52] C. M. Schneider, T. Mihaljev, S. Havlin, and H. J. Herrmann. Suppressing epidemics with a limited amount of immunization units. *Physical Review E*, 84:061911, 2011.

[53] F. Scorgie, M. F. Chersich, I. Ntaganira, A. Gerbase, F. Lule, and Y.-R. Lo. Socio-demographic characteristics and behavioral risk factors of female sex workers in sub-saharan Africa: a systematic review. *AIDS and Behavior*, 16:920–933, 2012.

[54] K. Shannon, T. Kerr, S. Allinott, J. Chettiar, J. Shoveller, and M. W. Tyndall. Social and structural violence and power relations in mitigating HIV risk of drug-using women in survival sex work. *Social Science & Medicine*, 66:911–921, 2008.

[55] D. C. Vissers, H. A. Voeten, N. J. Nagelkerke, et al. The impact of pre-exposure prophylaxis (PrEP) on HIV epidemics in Africa and India: a simulation study. *PLOS ONE*, 3:1–7, 2008.

[56] H. A. Voeten, O. B. Egesah, C. M. Varkevisser, and J. Habbema. Female sex workers and unsafe sex in urban and rural Nyanza, Kenya: regular partners may contribute more to HIV transmission than clients. *Tropical Medicine & International Health*, 12:174–182, 2007.

[57] E. Volz. SIR dynamics in random networks with heterogeneous connectivity. *Journal of Mathematical Biology*, 56:293–310, 2008.

[58] E. Volz and L. A. Meyers. Susceptible–Infected–Rcovered epidemics in dynamic contact networks. *Proceedings of the Royal Society of London B: Biological Sciences*, 274:2925–2934, 2007.

[59] M. J. Wawer, R. H. Gray, N. K. Sewankambo, D. Serwadda, X. Li, O. Laeyendecker, N. Kiwanuka, G. Kigozi, M. Kiddugavu, T. Lutalo, et al. Rates of HIV-1 transmission per coital act, by stage of HIV-1 infection, in Rakai, Uganda. *The Journal of Infectious Diseases*, 191:1403–1409, 2005.

[60] B. Williams, C. MacPhail, C. Campbell, D. Taljaard, E. Gouws, S. Moema, Z. Mzaidume, and B. Rasego. The Carletonville-Mothusimpilo project: Limiting transmission of HIV through community-based interventions. *South African Journal of Science*, 96:351–359, 2000.

[61] J. M. Wojcicki and J. Malala. Condom use, power and HIV/AIDS risk: sex-workers bargain for survival in Hillbrow/Joubert Park/Berea, Johannesburg. *Social Science & Medicine*, 53:99–121, 2001.

Appendix A

Additional Data from Carletonville Questionnaire

The Carletonville-Mosuthumpilo project [60] collected questionnaire data from approximately 800 miners living and working in Carletonville, 1500 residents of the neighbouring township of Khutsong and 100 women living in hotspots surrounding the mine complexes was collected. Along with standard demographic questions, all respondents were asked a number of questions regarding their sexual relationships.

Q214. What is you usual occupation?

- ☐ Work for a mining company
- ☐ Domestic work outside your home
- ☐ Selling things
- ☐ Unemployed
- ☐ Manual labour
- ☐ Professional
- ☐ Sex work
- ☐ Other

Q242. Did you ever have sexual intercourse with a sex worker?

- ☐ Yes
- ☐ No

Q243. How many times did you have sex with a sex worker in the last 12 months?

Q244. Last time (having intercourse with a sex worker) did you use a condom?

- ☐ Yes
- ☐ No

Q310. How many current regular partners do you have?

Q315. How many times did you have sex with your regular partner in the last week? (Recorded for up to two regular partners).

Q401. How many different people have you had sexual intercourse with in the last 12 months (other than a regular partner)? This includes mistresses, girlfriends, casual partners, prostitutes, or somebody you met in a bar or at a special occasion and partners you are committed to but don't live with.

Q403. How many times did you have sex with these partners (non-regular) in the last MONTH?

Q404. Thinking about the times that you had sex in the last MONTH, with these partners (non-regular) how many times did you use a condom?

Q414. Think of the last non-regular partners that you had sex with in the last 12 months. How many times in the last month did you have sex with this person?

The *clients* are taken to be those who responded "Yes" to Q242. This gives a total of 109. We have identified the female sex workers to be those who responded "Sex work" to Q214, regarding current occupation. In total, this gives 71 female sex workers. We are also interested in how the distribution of the number of non-regular partners is different between men and women engaged in sex work and all other men and women respondents. *Other men* refers to all men who indicated "No" to Q242. *Other women* refers to all women who indicated an occupation other than sex work in Q214.

A.1 Demographics

Some basic information regarding age and birthplaces of both clients and female sex workers is shown below.

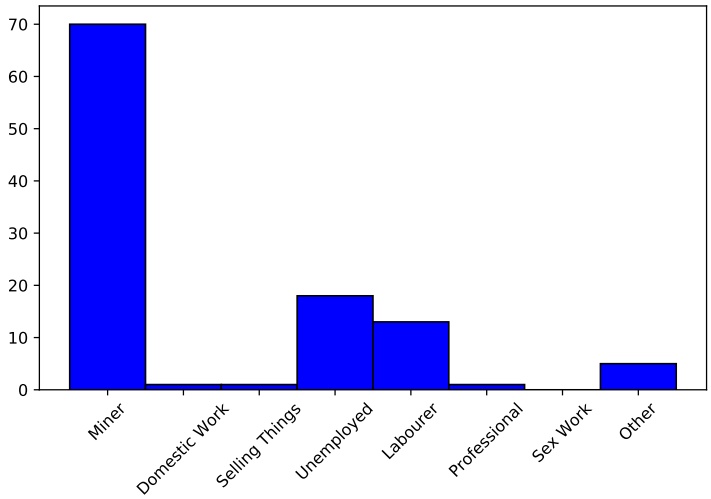
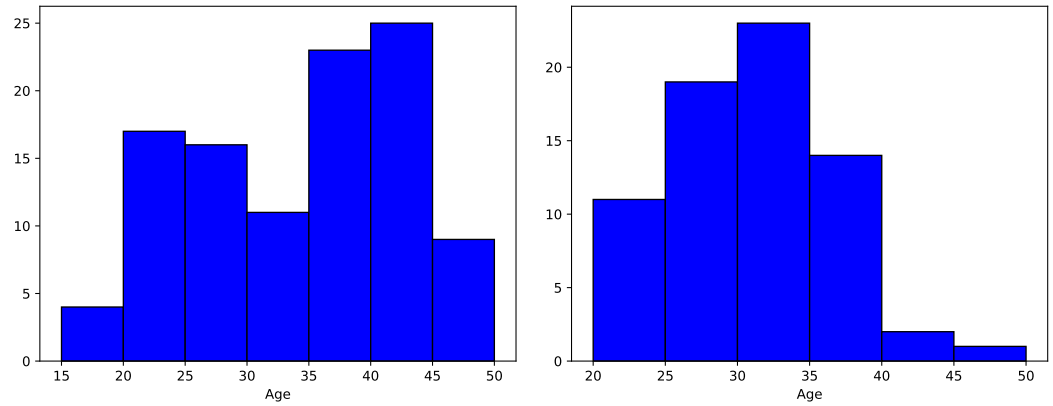


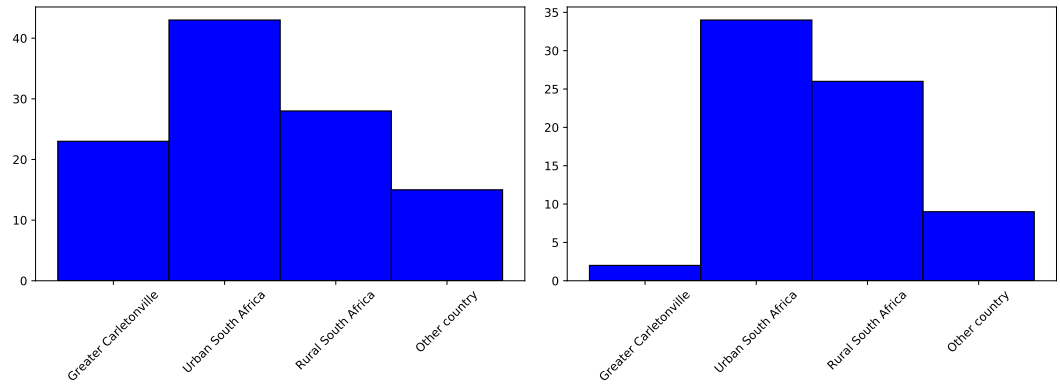
Figure A.1. *Responses to Q214, regarding current occupation, among men who self-identified as clients. Men who work for a mining company constitute 64% of all clients.*



(a) *Recorded ages of men who self-identified as clients. The mean recorded age is 35 years old, with range 17–55 years old.*

(b) *Recorded ages of women who self-identified as sex workers. The mean recorded age is 30 years old, with range 21–51 years old.*

Figure A.2. *Distribution of ages among clients and sex workers.*



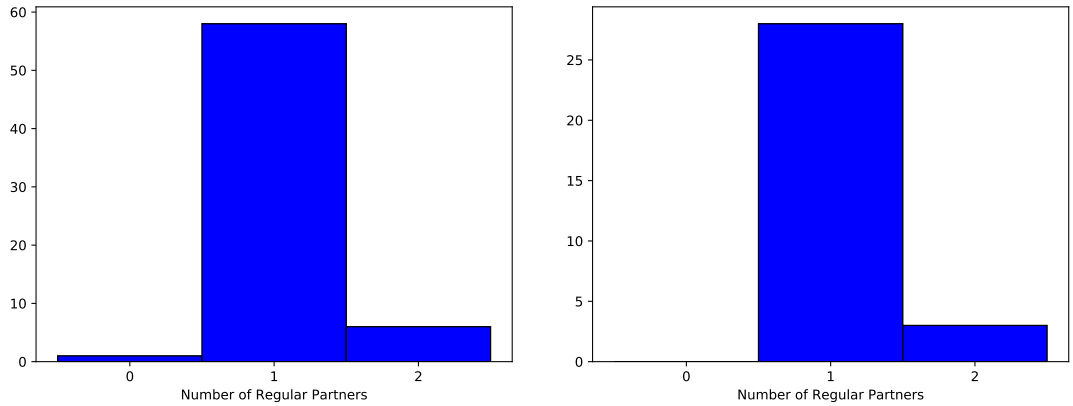
(a) *Birthplaces among men who self-identified as clients.*

(b) *Birthplaces among women who self-identified as sex workers.*

Figure A.3. *Comparison of recorded birthplaces.*

A.2 Sexual Behaviour

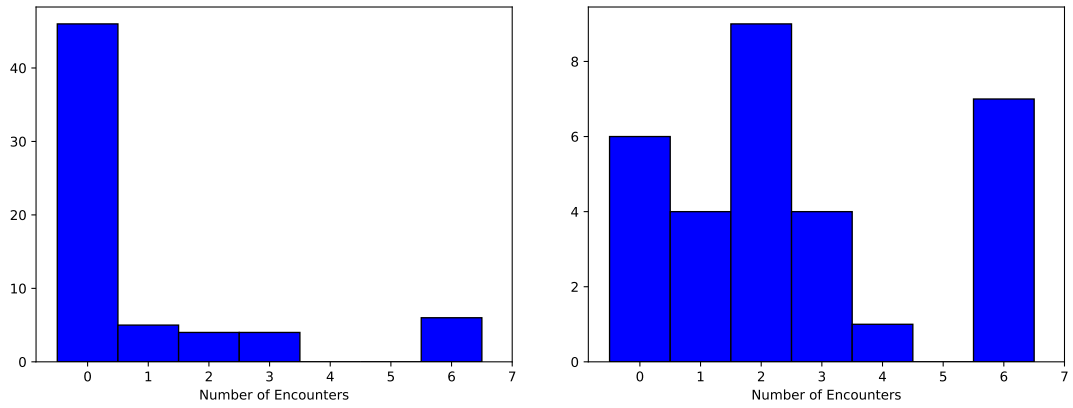
A major focus of Chapter 3 was on the nature of the distribution of non-regular partners among both clients and sex workers. How do these distributions compare to other men and women in the Carletonville area? How frequent is sexual contact among non-regular partners? Do those who engage in sex work have frequent contact with regular (spouse-like) partners?



(a) *Among men who self-identified as clients. There were 44 non-responses, corresponding to about 40% of all clients.*

(b) *Among women who self-identified as sex workers. There were 40 non-responses, corresponding to about 56% of all female sex workers.*

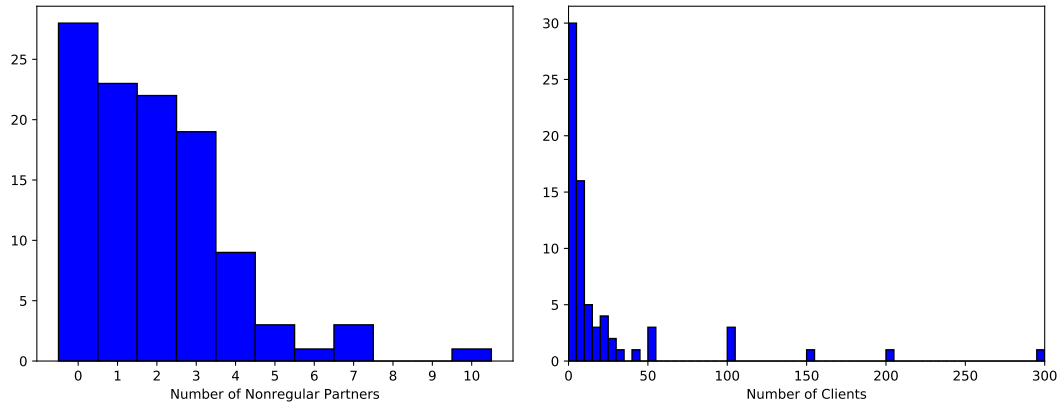
Figure A.4. *Responses to Q310, regarding the number of current regular partners. Participants who had never been married did not respond.*



(a) *Among men who self-identified as clients.*

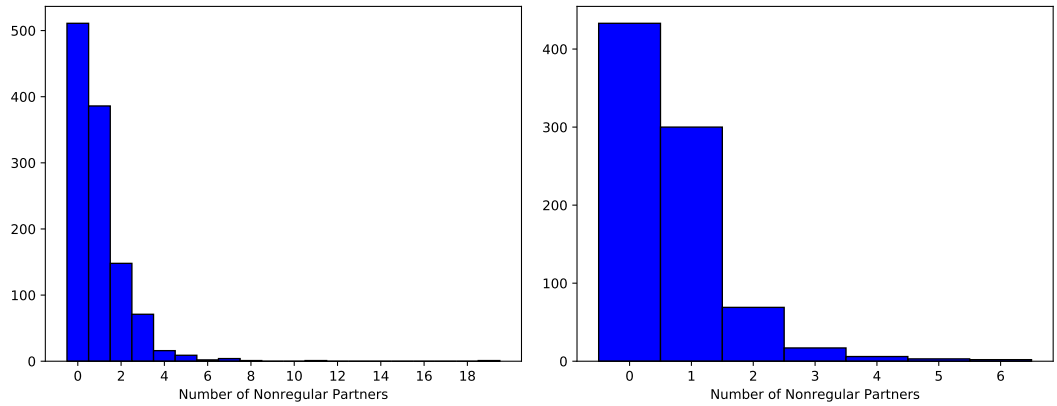
(b) *Among women who self-identified as sex workers.*

Figure A.5. *Responses to Q315, regarding the number of sexual encounters with one regular partner over the past week. Those who had indicated that they had never had a regular partner did not respond.*



(a) *Among men who self-identified as clients.*

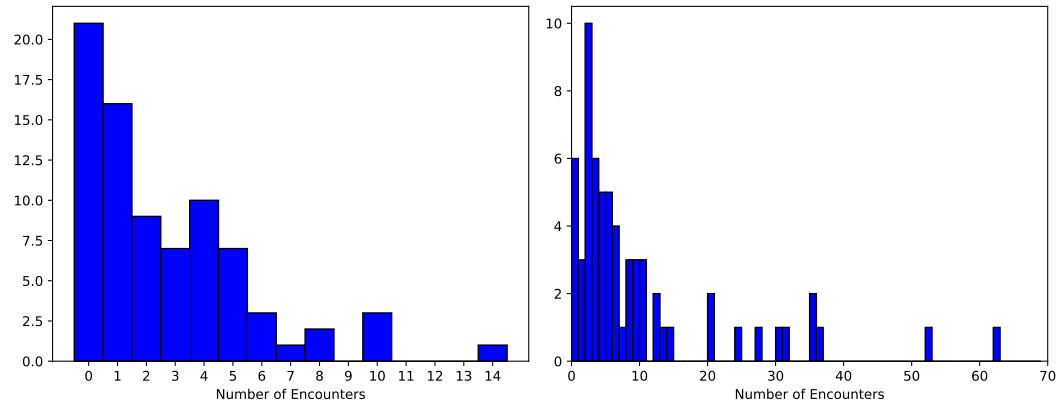
(b) *Among women who self-identified as sex workers.*



(c) *Among all other men.*

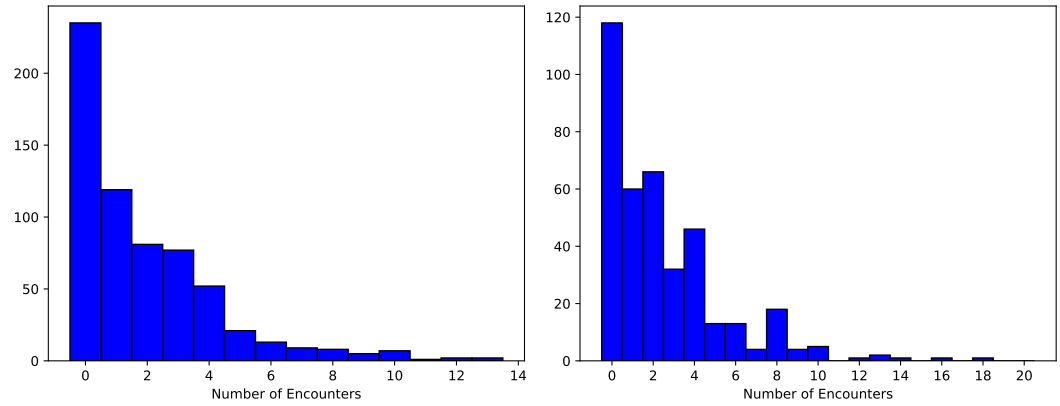
(d) *Among all other women.*

Figure A.6. *Responses to Q401, regarding the number of different non-regular partners in the last 12 months. Note that Figure 3.1 gives a higher resolution of the responses to Q401 among female sex workers, in which we can clearly see a peak away from zero.*



(a) *Among men who self-identified as clients. There were no respondents who indicated that they did not know.*

(b) *Among women who self-identified as sex workers. We exclude respondents who indicated that they did not know, corresponding to about 10% of all responses.*



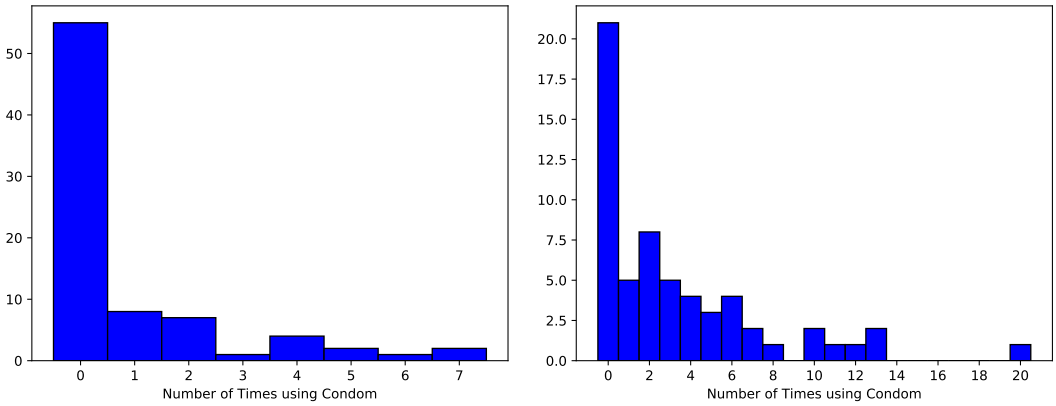
(c) *Among all other men. We exclude respondents who indicated that they did not know, corresponding to about half of all responses.*

(d) *Among all other women. We exclude respondents who indicated that they did not know, corresponding to about 60% of all responses.*

Figure A.7. *Responses to Q403, regarding the total number of sexual encounters with all non-regular partner over the past year.*

A.3 Condom Usage

One might also be interested in condom usage among those who participate in sex work.



(a) *Among men who self-identified as clients.*

(b) *Among women who self-identified as sex workers. Excluding all responses indicating "do not know", corresponding to 5 responses.*

Figure A.8. *Responses to Q404, regarding condom use with non-regular partners over the past month.*

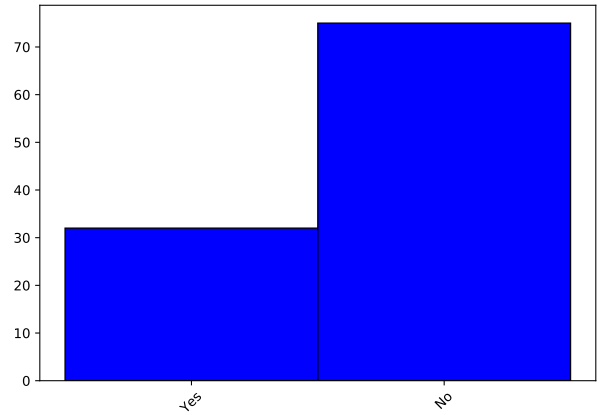


Figure A.9. *Responses to Q244, regarding condom use with sex workers, among men who self-identified as clients.*

Appendix B

Software

The simulated disease processes on each network were performed using the python package NepidemiX [1], the latest version of which can be found here:

`https://github.com/impact-hiv/NepidemiX`

and more information on the project can be found here:

`http://nepidemix.irmacs.sfu.ca/`

The version that was used for this thesis was forked in July 2017. The modifications made include: adding bipartite network functionality, adding unipartite and bipartite network generators (using NetworkX [28]) and adding optional stopping criteria. This version can be found here:

`https://github.com/nmulberry/NepidemiX`

Dependencies: Python 2.7, NetworkX 1.11 (this version of NepidemiX is incompatible with Python 3.x and NetworkX 2.x).