Fall 2014

Effective and efficient algorithms for simulating sexually transmitted diseases

Sean Lucio Tolentino
University of Iowa

EFFECTIVE AND EFFICIENT ALGORITHMS FOR SIMULATING SEXUALLY TRANSMITTED DISEASES

by

Sean Lucio Tolentino

A thesis submitted in partial fulfillment
of the requirements for the Doctor of
Philosophy degree in Computer Science
in the Graduate College of
The University of Iowa

December 2014

Thesis Supervisor: Professor Alberto Maria Segre

Graduate College
The University Of Iowa
Iowa City, Iowa

CERTIFICATE OF APPROVAL

_____

PH.D. THESIS

_____

This is to certify that Ph.D. thesis of

Sean Lucio Tolentino

has been approved by the Examining Committee for the thesis requirement for the Doctor of Philosophy degree in Computer Science at the December 2014 graduation.

Thesis Committee: _____

Alberto Maria Segre, Thesis Supervisor

_____

Philip Polgreen

_____

Ted Herman

_____

Sriram Pemmaraju

_____

Laurence Fuortes

All models are wrong, but some are useful.

George Edward Pelham Box

ACKNOWLEDGEMENTS

Thank you first and foremost to Emma Jacobs, who was a tremendous source of encouragement and support. And to her family, specifically Larry Jacobs and Julie Schumacher, who believed in me.

I'm also extremely grateful to my adviser, Alberto Maria Segre, who was fundamental to the direction and strength of the thesis. Wim Delva, my supervisor in South Africa, was monumental in making sure models were well situated in reality and questions were prescient and relevant. This work would not be possible without them.

Also thanks to the Sheryl Semler and Catherine Till who helped in so many little ways. I really appreciated it over these years. Thanks to the Graduate College and the Dean's Graduate Fellowship which made starting and finishing possible.

ABSTRACT

Sexually transmitted diseases affect millions of lives every year. In order to most effectively use prevention resources epidemiologists deploy models to understand how the disease spreads through the population and which intervention methods will be most effective at reducing disease perpetuation. Increasingly agent-based models are being used to simulate population heterogeneity and fine-grain sociological effects that are difficult to capture with traditional compartmental and statistical models. A key challenge is using a sufficiently large number of agents to produce robust and reliable results while also running in a reasonable amount of time.

In this thesis we show the effectiveness of agent-based modeling in planning coordinated responses to a sexually transmitted disease epidemic and present efficient algorithms for running these models in parallel and in a distributed setting. The model is able to account for population heterogeneity like age preference, concurrent partnership, and coital dilution, and the implementation scales well to large population sizes to produce robust results in a reasonable amount of time. The work helps epidemiologists and public health officials plan a targeted and well-informed response to a variety of epidemic scenarios.

PUBLIC ABSTRACT

Sexually transmitted diseases affect millions of lives every year. In order to most effectively use prevention resources epidemiologists deploy models to understand how the disease spreads through the population and which intervention methods will be most effective at reducing disease perpetuation. Increasingly agent-based models are being used to simulate population heterogeneity and fine-grain sociological effects that are difficult to capture with traditional compartmental and statistical models. A key challenge is using a sufficiently large number of agents to produce robust and reliable results while also running in a reasonable amount of time.

In this thesis we show the effectiveness of agent-based modeling in planning coordinated responses to a sexually transmitted disease epidemic and present efficient algorithms for running these models in parallel and in a distributed setting. The model is able to account for population heterogeneity like age preference, concurrent partnership, and coital dilution, and the implementation scales well to large population sizes to produce robust results in a reasonable amount of time. The work helps epidemiologists and public health officials plan a targeted and well-informed response to a variety of epidemic scenarios.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I INTRODUCTION

When the World Health Organization announced the eradication of the infectious disease smallpox in 1977, there was a sense that humans would eventually win the war against the microscopic organisms that have plagued our existence. Public health efforts have now turned to other eradication plans of infectious disease. In particular, sexually transmitted diseases (STDs) are similar to smallpox in so far as they are both easily preventable—though through precautionary measures not vaccines. It is for this reason that hope for eradication of these types of diseases is fathomable, if not entirely possible within our generation.

Reducing disease burden and eventually eradicating it will require developing tools for understanding the disease and the processes through which it is perpetuated. Mathematical and compartmental models have been used for the past 50 years with much success, but it is becoming increasingly clear that there are many fine-grain processes underling STD epidemics that these models have difficulty capturing. For this reason epidemiologists and public health officials are turning to agent-based models to understand how sexually transmitted diseases are diffusing through populations.

Making population-based models of disease is difficult though, as we show in the preceding sections. Other sciences can conduct experiments in a highly controlled laboratory setting on a system governed by fundamental laws of nature. Here we are forced to gleam information through observation of a system that is governed by rules that are constantly changing and highly heterogeneous. Heterogeneity poses a formidable challenge: how an individual forms and dissolves relationships is specific to an individual and is nearly impossible to fully quantify (what is a person attracted to? How social is a person?). Agent-based models can

account for this heterogeneity by endowing agents with individual characteristics and qualities that reflect reality.

However, even when fully accounted for in an agent-based model, so much heterogeneity can produce highly variable results. The number of stochastic interactions and outcomes is a chaotic system with outcomes that are probabilistically distributed rather than a single exact constant. Narrowing the distribution and producing robust output requires that we use a sufficiently large number of agents. This effectively creates multiple copies of a particular "kind" of individual, and the significance any one agent and its actions are diluted by the "law of large numbers".

Increasing the number of agents in your model isn't always simple though: The nature of network simulations is that a linear increase in population size quadratically increases the amount of time required to run a simulation. This may be fine for extremely simple models – e.g., the number of agents needed for a model in which agents only form relationships based on potential partners sex won't dictate an unreasonable amount of computation time – but for this low level of heterogeneity a simple compartmental model is perhaps better suited. This quadratic relationship can be unfortunately untenable for models with even modest amounts of heterogeneity in agents. In the next section we review the epidemiology of Syphilis and HIV, and present sources of heterogeneity. The rest of this chapter reviews previous modeling approaches for accounting for these effects.

Chapter 2 describes a mathematical formulation for simulating a heterogeneous and dynamic sexual network. We show how the formulation can effectively simulate intra-host biological processes, many different age-mixing patterns, and reproduce demographic processes that have occurred over the past 30 years. The mathematical formulation is a basis for the agent-

based models described in thesis and is presented to showcase the large amount of heterogeneity that such a formulation can model.

Chapter 3 shows the usefulness of the mathematical formulation with a simplified version used to investigate combination HIV prevention. We present a simulation-based method that uses machine learning and search heuristics to efficiently allocate disease prevention resources and effectively reduce disease prevalence. The work helps governments on fixed budgets decide which intervention to implement (e.g. condom distribution, male circumcision campaign, increased access to anti-retroviral therapy), where to implement it, and how much to spend on it. The simulation results suggest that a combination of prevention methods implemented in a non-trivial way can avert more infections and reduce prevalence more than any single intervention in isolation.

In chapter 4 we return to investigating the quadratic relationship between heterogeneity and computational run time. The mathematical formulation indeed can handle simulating highly heterogeneous populations, but its initial implementation does not scale well to a large numbers of agents. For this reason in chapter 4 we present a parallelized algorithm for simulating dynamic sexual networks. We again use a simplified version of the mathematical formulation, but we show that large populations of highly heterogeneous populations can be efficiently simulated.

Chapter 5 is a further parallelization of the simulation. We exploit the natural geographic partition of sexual networks to distribute the computation onto multiple nodes of a cluster. The partition allows us to simulate even larger population sizes and hence more heterogeneity, as well as enables modeling of geographic processes such as migration and mobility.

In all this we show that effectively and efficiently simulating sexual transmitted diseases is possible. While the nature of the system precludes these models from being scientifically validated we show that they can be close enough to reality to be useful.

1.1 Syphilis and HIV Epidemiology

The human immunodeficiency virus (HIV) epidemic in Africa has not been overstated: there are an estimated 33.3 million individuals living with what has become known as one of the worst infectious diseases affecting mankind [4, 5]. In 2010, there were 1.8 million AIDs related deaths—contrast this with seasonal influenza which kills on the order of hundreds of thousands [1]. South Africa represents less than 1% of the world population, but carries about 35% of the worlds' HIV burden with the adult prevalence estimated at 29% [2]. Compare this prevalence to South Africa's Northern African neighbors Kenya, Tanzania, and Uganda which have rates of 6.3%, 5.6%, and 6.5% respectively as seen in [3]. Over the past three decades there have been many studies both implementing and analyzing specific interventions to combat HIV. Since each prevention method has a different financial cost of implementation, as well as varied community acceptance, the most effective intervention strategy likely requires a multi-level and multi-component approach. That is to say that the most effective way to ultimately eradicate HIV is to implement interventions in combination such that the combination of interventions has the optimal effect.

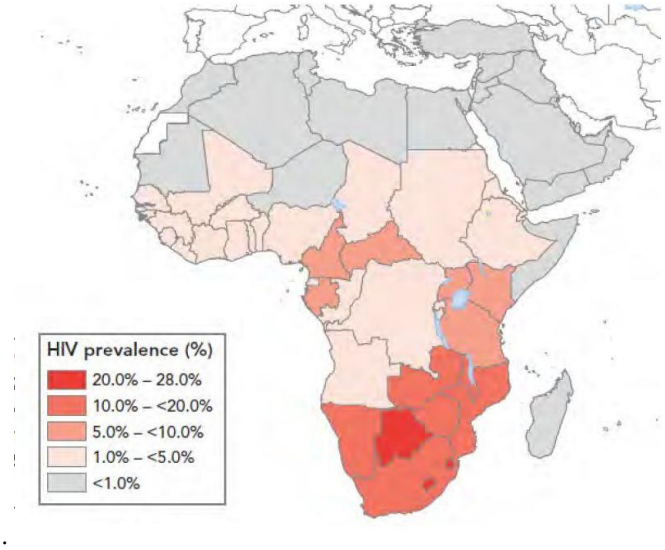

Figure 1: HIV prevalence is significantly higher in Southern Africa [8].

The global HIV/AIDS epidemic is a difficult problem to solve, but it is not impossible. The three difficulties faced are the problems of finding the right model for understanding disease diffusion, inferring parameters for the model, and computational limitations of finding an optimal allocation of resources. To find the best combination of prevention methods, we will consider the parameters of the disease and how it is spread; the sociological and political reasons for why things are the way they are; the geographic progression of the disease; and the societal determinants that may fall beyond the typical scope of disease eradication strategies. Due to the intricate interweaving of interventions and their interactions, a combination of methods is likely to be the most effective. Methods for finding this combination are discussed in future sections.

Syphilis, another sexually transmitted disease, is common in most parts of the world; those who suffer from it are plagued with rash and boils. If left untreated the disease can eventually lead to death [4]. Its derivative, congenital syphilis, is the disease that is transmitted from a syphilis-infected mother to her child during pregnancy. Syphilitic pregnant women are likely to infect their unborn children with congenital syphilis who then have an increased

likelihood of stillbirth or becoming victim to major birth defects such enlarged liver and spleen, rash, fever, extreme blistering, rhinorrhea, and oedema of the face [5]. Though not as prevalent as HIV/AIDS, the sobering fact of syphilis is that it is curable with a single dose of penicillin and can be eradicated with the right plan of action. However, a significant gap exists between the medical ability to cure syphilis, and the geographic and behavioral information necessary to contain syphilis: though we know how to treat the disease, we do not know how to control its spread. Agent-based simulations that consider different disease transmission parameters may provide insight into how the disease is perpetuated.

1.1.1 Disease Parameters

STD intervention methods can be grouped based on the specific exogenous attribute of the disease that the intervention aims to interrupt: either the infectivity or connectivity of individuals. For example, condoms attempt to reduce infectivity by reducing the amount of bodily fluids that come in contact between sero-discordant sexual partners, thereby reducing the overall probability of transmission in a single sexual act. A mass media campaign that encourages sexually active individuals to limit their number of sexual partners reduces overall connectivity – decreasing the overall number of possible transmissions. Campaigns that encourage serosorting, individuals engaging in unprotected sex only with others of the same infection status, similarly decrease the number of possible transmissions. Understanding these two variables of disease spread help us understand health interventions, their limitations, and how they might work in conjunction for an optimal combination prevention strategy.

Additionally, specific to HIV and of great importance, are the endogenous attributes of HIV that make typical “screen, treat, and release” methods implausible. These are attributes cannot be interrupted through public health interventions very easily.

1.1.1.1 Endogenous Factors

HIV is surprisingly difficult to transmit. In studies of monogamous couples in which one partner was HIV-positive, the transmission rate of HIV was is about 0.001 per sexual act [6]. That is to say that, if an individual has unprotected sex with someone who is HIV positive, the probability of becoming infected him or herself is less than 1 in 100. It is perhaps shocking then how, in the 50 years since the first known case of HIV in world, the virus was able to spread to nearly every country and reach a point of 33.3 million infected individuals in the world [3]. The counter-intuitive worldwide epidemic can be attributed to a few factors that distinguish it from other opportunistic infections. These intrinsic disease characteristics provide insight into HIV's global spread.

The first characteristic of importance is the virus's rapid rate of mutation: 1 in every 10,000 duplications is a mutation. This is as compared to a typical cell in which 1 in every 1,000,000,000 is a mutation [7]. This fast rate of change makes it difficult for scientists studying the disease to create a cure or even a vaccine because it quickly adapts to potential treatments and develops resistance. Additionally, the disease has a very high replication rate which means a typical HIV patient has a completely new viral load every two to three days [8].

The second characteristic of importance is the slow rate at which the disease kills an infected individual; it might be as many as six years before symptoms begin to appear [7], and nine to ten years before death [9]. The long window without symptoms equates to more exposures and thus increased transmission. Before HIV was even fully understood and recognized as transmitted through exchange of bodily fluids, the disease had many years to spread via prostitution and truck routes throughout all of Africa and the world [10]. This makes ART a double edged sword such that it prolongs patients' lives, but allows more opportunities for

infecting new persons. A “successful” pathogen balances host survival against transmission—this is why Ebola, which kills most infected individuals within a week, is not at pandemic levels [11].

The third important characteristic is its ability to hide. The virus reproduces itself by attacking healthy cells and using host cells’ replication abilities. This does not happen immediately however, as the virus may remain dormant within the host and only later begin reproduction [9]. This means that treatment like antiretroviral drugs may remove all HIV cells, but leave those that are dormant. Current efforts to cure the disease are aimed at finding methods for “waking up” these dormant cells so that they too may be attacked [12, 13].

Conversely, syphilis is transmitted between sexual partners in 30-50% cases of exposure [14]. It is, however, less of a public health threat than HIV for several reasons. First, unlike HIV, syphilis has not developed resistance to treatment through mutation; 50 years after first treating the disease with penicillin there is no evidence of penicillin resistant strains of syphilis [15]. Additionally, it does not “hide” as HIV does—a penicillin shot completely cures syphilis.

Second, individuals are typically only infectious during the primary and secondary stages of the disease which shows infections with lesions. The primary stage usually occurs within the first 90 days of infection and is identified by a large lesion or *chancre*. The secondary stage is indicated by similar rashes and ulcer as well as flu-like symptoms. If left untreated, infected individuals enter the latent and tertiary stages of the disease which leave him or her asymptomatic and highly unlikely to transmit syphilis. Additionally while experiencing lesions or rashes indicative of the disease, individuals may self-select out of dangerous sex patterns perhaps out of self-preservation.

1.1.1.2 Infectivity

There is a striking difference between prevalence rates of syphilis and HIV in sub-Saharan Africa and the Western world. Environmental factors, also known as *exogenous factors*, are more nuanced and advance the disease in a dramatic, albeit subtle, way. To begin, the probability of transmission per sexual act (PTSA) mentioned earlier is not static number. It varies with the viral load of the infected, the mode of transmission (heterosexual, homosexual, injection drug user), the presence or absence of other sexually transmitted diseases, etc. The contrasting social and cultural attributes of countries affect the disease PTSA, be it positively or negatively, ultimately making the disease spread more or less likely.

One of the most deterministic attributes about the infectivity of an HIV positive individual is his or her viral load [6, 16]. The viral load is a measure of the amount of the virus in an individual's bloodstream. When viral load is high (measured in copies of the virus per milliliter of blood), the infected individual is significantly more infectious. Table 1 below illustrates the increased risk of infecting virus free sexual partners with increased viral load. Viral load is reduced by ART, but can be expensive for infected individuals living in Sub-Saharan Africa.

Table 1: Risk of infection increases with viral load.

Viral Load	**Unadjusted Relative Risk**	**95% Confidence Interval**
0–3,000	1	1
3,000-14,500	3.56	(1.07–11.81)
14,500–76,000	7.18	(2.30–22.38)
>76,000	9.62	(3.00–30.84)

In the early 1980's when HIV was still not well understood, the widespread prevalence of HIV in the homosexual population led to the misconception that it was primarily a homosexual disease [7]. While it is now commonly accepted that both homosexual and heterosexual alike are susceptible to infection, the difference in its rapid spread through the homosexual community

may be attributable to the dissimilar mode of sex: while a vagina is biologically built for intercourse, the anus is not. Penile-anal sex, along with the common practice of fisting (inserting the entire hand / forearm into the partners rectum) very often leads to rectal tears and anal fissures. These openings increase the virus' ability to enter the body and ultimately infect an individual [7]. Co-infection with other sexually transmitted diseases (STDs) can lead to increased infectivity of the virus in a similar ways. Lesions manifested with syphilis allow the HIV virus to more easily enter and infect a new individual [6, 17]. Other non-ulcerative STDs such as gonorrhea and chlamydia "increase HIV shedding in the genital tract, probably by recruiting HIV infected inflammatory cells as part of the normal host response" [17].

While not a sexually transmitted disease, Tuberculosis (TB) complicates HIV elimination plans with the large co-infection rate. TB is the most common opportunistic infection of HIV positive individuals, with about 73% of TB infected individuals testing positive for HIV in South Africa [18]. This significant correlation has seen a call for a more collaborative approach between HIV and TB care providers [19]. The "silo approach," which is characterized by separate diagnoses, care, and treatment can be integrated through joint planning of surveillance and screening for other diseases at admission. Collaborative efforts in TB, HIV, and syphilis would lead to a significant decline in the overall mortality rates of these diseases [17].

While classified as a sexually transmitted disease, HIV is spread through exchange of bodily fluids and therefore does not necessarily require sexual contact. Thus, is not surprising that HIV has also had a marked impact on drug users that reuse non-sterile needles [7]. Additionally, in South Africa, mother-to-child transmission is second only to heterosexual sex as a mode of transmission [18]. It is estimated that 15-35% of HIV positive mothers will pass on the disease to their unborn child either during delivery or in utero [20].

1.1.1.3 Connectivity

While the probability of infection in a single sexual act is relatively low, the probability of ever getting the infection increases considerably when considering other variables. For example, a large number of sexual partners increase the likelihood of infection by increasing the probability of having sex with an HIV infected individual. Increased frequency of sexual intercourse also increases the likelihood by giving the HIV virus more opportunities for infection. Societies that are more accepting toward prostitution may see an increased prevalence rate due to the high number of sexual partners that sex workers have. This effectively creates "hubs" of infection.

Post-apartheid South Africa is still struggling with socio-economic disparities despite having one of the most functioning economies in Africa. Poverty and low-quality education (particularly about HIV) are among the results of such economic disparities. The housing crisis causes low-income communities to live in very close proximity to each other which increases the disease ability to spread (as compared to communities that are geographically dispersed).

1.1.1.4 Societal Determinants

While poverty reduction is often thought of as falling under a different human rights umbrella, reducing the number of people in extreme poverty may have many health implications. In a world without poverty any individual that is diagnosed with HIV would be able to afford ARV treatment, either through health insurance or out-of-pocket. Though not necessarily the case, a world without poverty would likely be a more educated world in which all citizens were knowledgeable of the risks and consequences of unsafe sex. These effects would undoubtedly have a positive outcome in minimizing HIV incidence. However, there are other more subtle interactions going on between poverty and the disease that make a compelling case for poverty reduction as a means of HIV prevention and control.

The fact that HIV has been so much more severe in Africa is easily seen from the Figure 2 below. Around 80% of the world's population lives in the developing world, and 95% of those infected with HIV live in the developing world. Part of the increased severity may be due to widespread malnutrition and parasitosis, results of pervasive poverty [21]. These reduce an individual's overall immunity, and consequently increase the likelihood of infection.

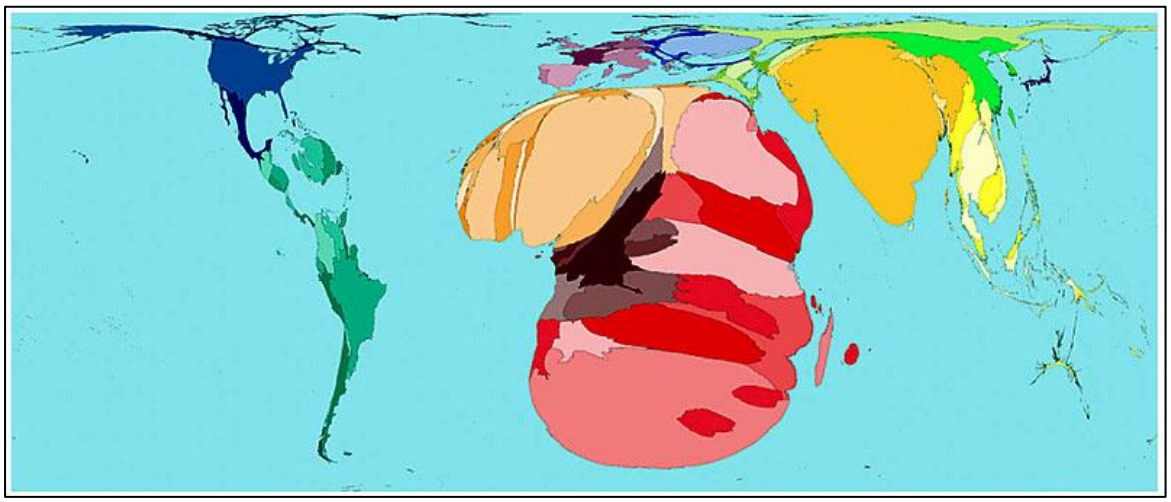

Figure 2: HIV prevalence in the world. Africa holds a significant burden of the disease.

As mentioned before, a lack of education leads to risky sexual behavior because of a misunderstanding about the disease. However, it also has the effect that the uneducated are less flexible in terms of working environments and conditions. In the case of South Africa, rural men migrate to the bigger cities in search of work. Many of them work long hours in the country's coal, gold, diamond, platinum, and chromium (used for stainless steel) mines. The mine's artificial environment weakens workers immune system and makes them more susceptible to HIV and TB infection [22, 23]. Additionally, the stressful work in the mining industry drives many men to alcohol, which is associated with less rational decisions and more risky sexual behaviors [24, 25].

Moreover, poor women, desperate for money, may turn to sex work as a means to feeding themselves. This unfortunate truth increases their risk of infection through the increased number

of sexual partners. Additionally, with the abundance of sex workers and limited regulation, there is a competitive incentive not to use condoms – there is likely another prostitute that will take away business because she is willing to "take the wrapper off the sweet" [26].

The fact that prostitution is culturally acceptable can be traced to the general consensus that young boys need an outlet for their sexual nature. However, young Muslim girls should save their virginity for their future husband, and so the concept of prostitution as a necessary evil develops [26].

Understanding societal determinants that may increase or sustain a high incidence rate of HIV is akin to understanding the weavings of a complex tapestry; there are many interacting layers, each exacerbating another and all contributing to the end result. The concept of poverty is itself a multi-faceted issue with many implications and many challenges to remedy; it is just one of many societal determinants adding to the problem. While eradicating poverty completely would surely not eliminate HIV transmission, it is not inconceivable that reducing poverty may have a substantial effect. Models that attempt to eradicate disease must then incorporate societal factors in some capacity.

1.1.2 Determining Prevalence

The syphilis epidemic has led to some research into the prevalence of the disease in specific areas [19], as well as surveillance by individual countries' ministries of health [27]. From these different sources, an educated guess can be made as to the disease's prevalence. There are an estimated 12 million new cases of syphilis in the world each year, a quarter of which occur in Africa [28]. Figure 3 below is a cartogram depicting the number of deaths due to syphilis in the world. As can easily be seen, a large portion of the deaths occur in Africa (approximately 30,000 in 2004). Infection rates in major African cities of Zambia and Cameroon were reported at 10%

and 6% in both genders [28], and ongoing tests in Madagascar suggest an infection rate of 30% [32].

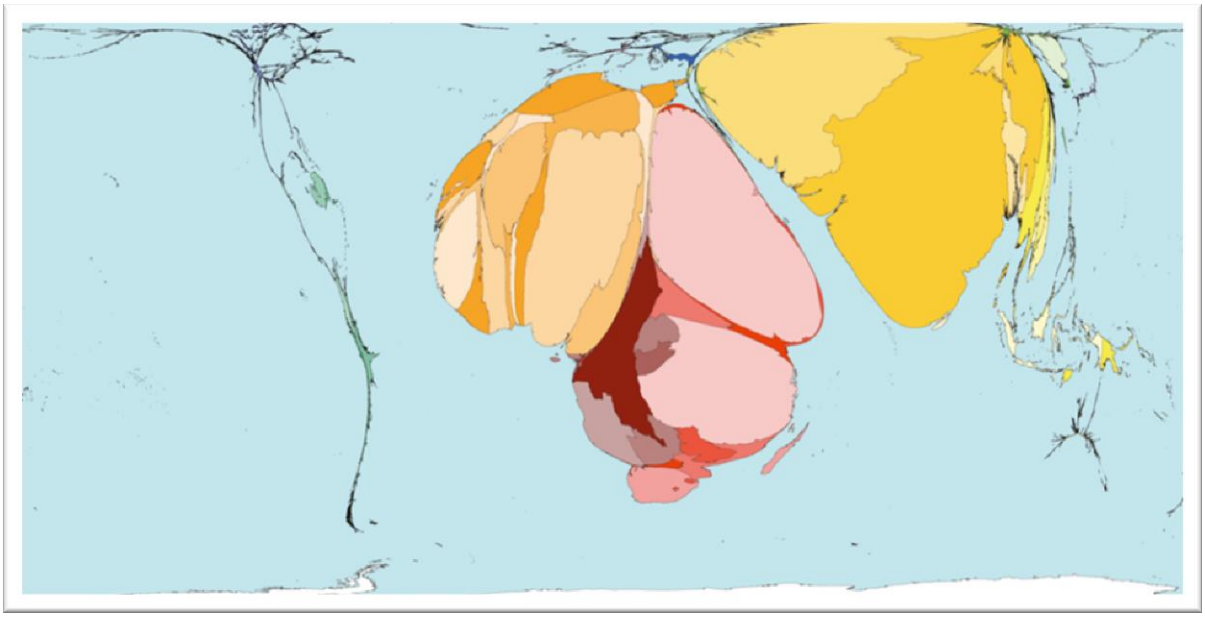

Figure 3: Cartogram of deaths due to syphilis in the entire world. Each color represents a region of the world: red is Southeastern Africa, orange is Northern Africa, and yellow is greater India and Far East. Africa holds a staggering amount of the burden of deaths due to syphilis. Deaths due to syphilis are mainly concentrated in Africa and South Asia.

Syphilis infection rates in pregnant women in Africa as a whole have been estimated to be between 3 and 15%. Of those, 30% of the untreated cases result in stillbirth and in another 30% the child will be born with congenital syphilis [27]. Half of infants born with congenital syphilis die within their first year of life. Though simply correlation, this may account for at least some of Africa's high infant mortality rates: 175.90 deaths per 1,000 live births in Angola; 81.04 deaths in Malawi; 66.0 deaths in Zambia. Compare this to 6.06 in the United States and 2.78 in Japan [29].

In the US, the Center for Disease Control (CDC) regularly produces publically available maps of syphilis and other diseases. Though it is possible to draw rough estimates of prevalence from specific case studies, and despite the large amount of research in testing and treating the disease [30], there seem to be no maps of syphilis in Africa [31]. This is a significant handicap

for epidemiologists attempting to prevent the diffusion of syphilis. We are thus left with open-ended questions and no channel through which to find answers: which geographic areas should be the focus of treatment and prevention, where to place treatment facilities, who are the most in need? Does geography play a larger role than demography? This frustrating lack of information can be attributed to logistical problems with screening tests, country-mandated data collection, as well as a lack of a unified aggregator.

Many African countries attempt to control syphilis prevalence through screening programs implemented at antenatal care clinics in the country[27]. While this is obviously a well-intentioned first step, it falls significantly short of a consistent source of data for many reasons. The two widely used tests for screening, Venereal Disease Research Laboratory (VDRL) and Rapid Plasma Reagin (RPR), have major flaws when used in developing nations. First, they require significant infrastructure to perform (necessitating a centrifuge, hot-water bath, and refrigeration, all of which require electricity which is unreliable in some areas) in addition to the training of individuals to interpret results [32]. Second, the time required to perform the complicated algorithm of testing can take 30-40 minutes, possibly resulting in a positive diagnosis for someone who has already left the clinic and may never return [32].

These two obstacles together lead to a disappointingly small percentage of pregnant women being screened for syphilis [32], where unscreened women become untreated women. In addition to the possibility of further spread, one-third will pass the disease onto their child in the form of congenital syphilis [33]**.** Moreover, many women do not attend a clinic during pregnancy in the first place. Thus, despite being national policy, an optimistic estimate of the screening rate is 38% for pregnant women [27].

The attractive alternative to VDRL and RPR are Rapid Syphilis Tests (RST), of which there are 20 commercially available versions. Though all differ slightly, their main benefits are that they are self-contained—there is no need for refrigeration or other machinery (let alone electricity) and require less training to administer, as well as producing results in 15 minutes—enough time for women to be treated in the same visit. Though the benefits seem overwhelmingly positive, the political reasoning for government's reluctance to implement their use is cost: one RST costs as much as $1.00, where an RPR costs as little as $0.15 per unit [32]. Compare this to a shot of penicillin which can cost from $50-$100. However the true cost relative to disability-adjusted life years depends mainly on how well equipped the country is. Antenatal clinics in Mwanza, Tanzania, for example, are much better equipped than most Tanzanian clinics, and so using RPRs may be more cost-effective in that community [34].

Though price is a major concern for Ministries of Health that are providing funding for screening, there are other difficulties with the tests. In the Gambia, approximately 75% of the population lives in rural areas [29]. Though the prevalence is significantly less than urban areas, syphilis infection in rural areas is estimated at 3% [30]. In a rural setting the procedures for the more complicated RPR tests become even more difficult; 100^{o}+ F temperatures reduce the number of antibodies, dusty environments distort blood samples, and poor light make reading instruments challenging. Together they all decrease the reliability of a positive/ negative diagnosis of syphilis. The RST tests tend to be subjectively easier to perform and do not suffer from the same environmental pitfalls of RPR. All this being considered, both tests show disappointing sensitivity to the disease; RPR was able to correctly identify 77.5% of positive cases, and RST 75.0%. This means many false negatives, and consequently under treatment of syphilis for those who need it, and many false positives resulting in unnecessary and expensive treatments. This is attributed to the relatively low prevalence in the rural areas—when a villager

receives a positive test there is only a 32.6% or 40.0% chance of actually having syphilis for PRP and RST, respectively [30].

The World Health Organization (WHO) has made an attempt to collect prevalence data for syphilis by means of the human immunodeficiency virus (HIV) surveillance programs [33]. Unfortunately, this is inherently flawed by the number of hands through which the data must pass first. Since prevalence of HIV can be an indicator of a country's developmental progress, there is a tendency for country officials to lie about numbers in order for their country to be perceived in a positive light [43]. Additionally, there may be economic incentive to underreport disease prevalence—money from tourism is good for the economy as a whole, and good for the government who receives a significant portion of the money through taxation [10].

Though it is difficult to create maps of diffusion of syphilis, it is possible to gain an understanding of its spread from the immense amount of literature on HIV. Since HIV attacks the immune system, those infected with it will be more susceptible to other diseases such as syphilis [35]. However the reverse is also true: it is widely accepted that there is a larger risk of contracting HIV because of a syphilis infection [17]. Studies of homosexual and heterosexual individuals consistently find an estimated 2.3 to 8.6 increased likelihood in the risk of transmission [17]. For this reason it is possible to make the simplifying assumption that syphilis spreads similar to HIV geographically. Even so, it is important to note that high prevalence of HIV in a community does not automatically imply high prevalence of syphilis [19].

1.1.3 Geographic Spread

The most notable cause of the spread of HIV is long distance truck drivers making shipments across national borders [36, 37]. This assertion is well-supported; 80% of bar girls working at truck stops along major highways are infected with HIV; various studies of truck

drivers show that anywhere from 30-80% are infected [10]. Most noteworthy is the trucking route from Djibouti, where HIV comes from many places via its heavily trafficked Red Sea port, toward southern Africa as can be seen in Figure 4. A docked sailor might visit one of the local prostitutes, who in turn is visited by a truck driver heading south to the Ethiopian capital of Addis Ababa. Perhaps not surprisingly, 50-60% of prostitutes in Djibouti are infected with HIV. Part of the reason for this continuing trend is the cultural acceptance of prostitution in their society, coupled with the Church's condemnation of condoms [10].

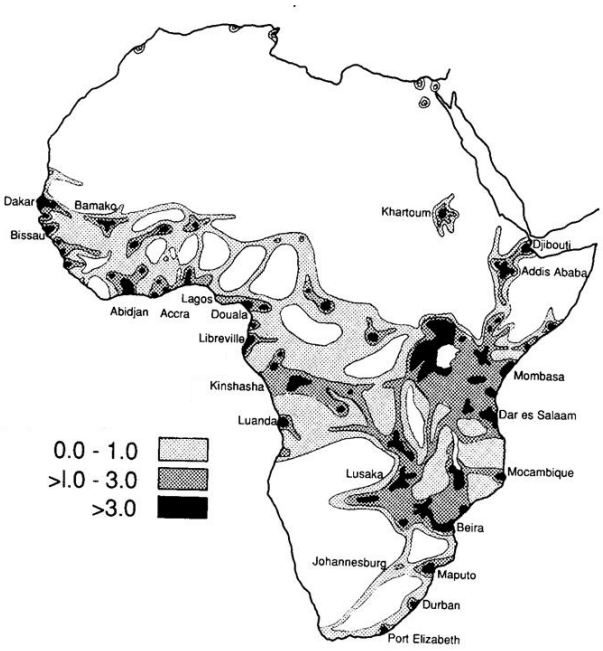

Figure 4: Minimal estimates of HIV infection rates in Africa in 1991. The higher incidence rate areas are correlated with traffic routes.

From Addis Ababa, truck drivers move south to Kenya and West to Somalia. In Kenya's capital, Nairobi, nearly 100% of prostitutes are infected with HIV [10]. Lake Victoria to the West of Nairobi exacerbates the diffusion as it is a commonly used mode of transportation to Uganda, Rwanda, and Tanzania. This war torn area is highlighted in Figure 4 with a large black spot near the middle of Africa. Civil unrest causes the movement of refugees and with them the HIV with which they are infected. Just as those that live near a major road are more likely to be infected

with HIV, those that live near the lake are more likely to be infected [38]. These details of spread lend themselves to identifying strategic places for epidemiological interventions where screening and treatment centers may be created.

1.2 Compartmental Models

In this section we present an overview of compartmental models for disease perpetuation and spread. For HIV we consider a standard SI epidemic process. We denote *S(t)* as the proportion of susceptible individuals at time *t*, and *I(t)* as the proportion of infected individuals at time *t* in a system of *N* individuals. The model holds that at every time step a fraction of individuals move from the susceptible *compartment* to the infected compartment. Additionally if we make the simplifying assumption that enter and exit rates are negligible (the number of births and deaths are equal), then we can use a constant replacement rate δ that models some fraction of infected individuals moving to the susceptible population. It is important to emphasize that this is aggregate behavior and so this is modeling the fact that some individuals within the infected group die and others are born into the susceptible population – not that some people are being cured of HIV. The infectivity of a disease with no interventions implemented is denoted λ_0. Also known as the *sufficient contact rate*, the infectivity is based on the connectivity of the population and transmission probability of the disease, and controls the number of new infections that occur within the system as seen visually in Figure 5 below.

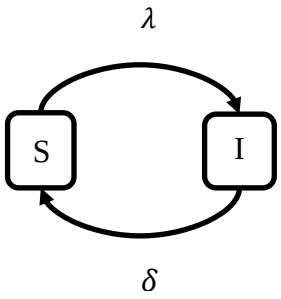

Figure 5: An SI model represents aggregate number of individuals in each of the two compartments: susceptible (S) and infected (I). Each time step some fraction of the susceptible population becomes infected relative to the infectivity coefficient $\boldsymbol{\lambda}$, and some fraction of infected become susceptible relative to the recovery rate (enter/exit rate) $\boldsymbol{\delta}$.

The number of new infections at time *t*, known as the *epidemic function*, is given by the formula

$$f(t, \lambda) = \lambda NI(t)S(t).$$

This comes from the fact that new infections occur based on the *contact rate* (the amount of mixing between the infected and susceptible individuals). Note that this formula assumes *random mixing* between compartments—every infected individual is equally likely to come in contact with a susceptible individual. Figure 6 shows epidemic curves for different values of λ with $\delta = 0$. As would be expected, even with a relatively low probability of transmission ($\lambda = 0.1$) the prevalence of the disease (the proportion of the population infected) is continually increasing. When the proportion of susceptible to infected individuals becomes low (as realized by the product $(t)S(t)$) the number of new cases in each time step declines.

Figure 7 shows how an epidemic for different values of δ with $\lambda = 0.5$. Now the system moves gradually to a steady state in which the number of new infected at each time step is equal to the number that enter/exit (recover). When very few individuals move from the susceptible to infected ($\delta = 0.1$), the steady state is high with 0.8 of the population being infected at any given time. For a higher enter/exit rate ($\delta = 0.4$) the steady state of the system is much lower.

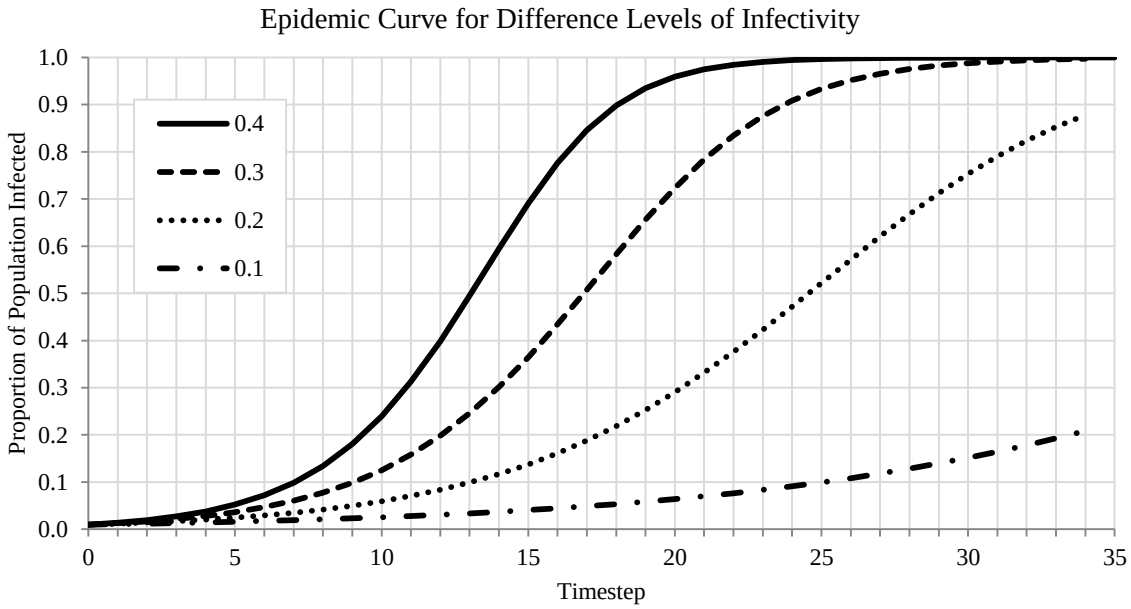

Figure 6: Epidemic growth over time for various values of infectivity. A highly infectious disease ($\lambda = 0.4$) infects nearly the entire population by time step 20. A less infectious disease ($\lambda = 0.1$) has only infected 0.2 of the population by timestep 35. Since the enter/exit rate is set to zero in this case, no infected individuals ever move back to the susceptible stage and the whole population gradually becomes infected no matter the value of λ.

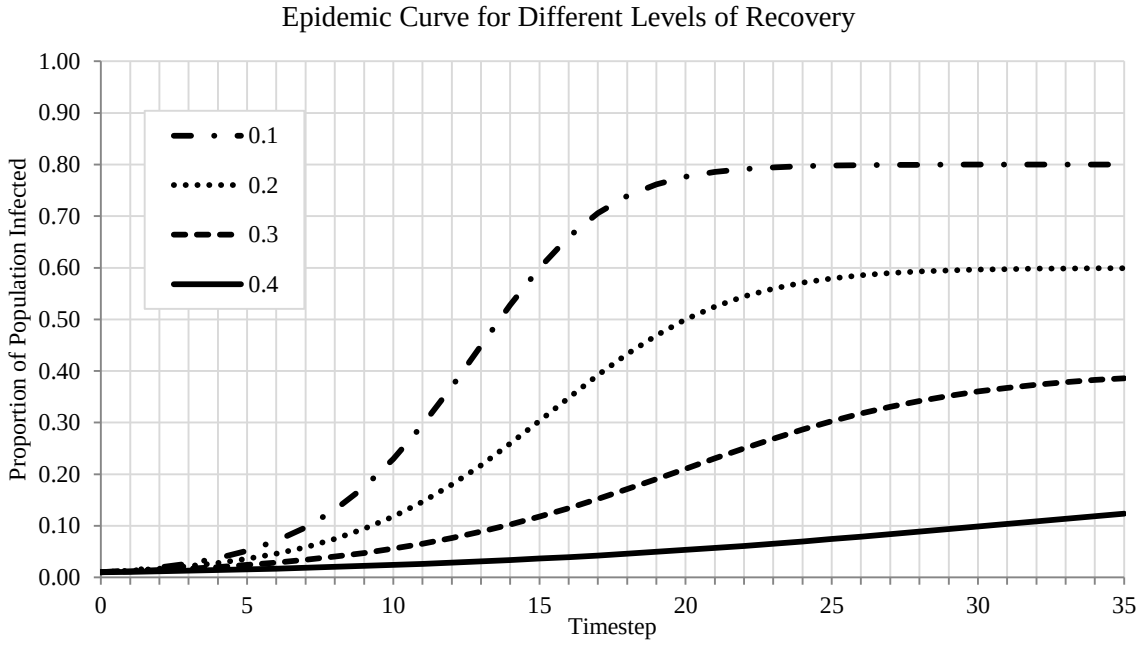

Figure 7: Epidemic growth over time for various values of enter/exit (recovery) rates. A high recovery rate implies that many people are moving from infected back to susceptible. Over time the system enters a steady state in which the number of new infected individuals is equal to the number of new susceptible.

SI models can be extended so as to allow for transition to other possible compartments. For example, an SIR model allows for individuals to move from the infected (I) compartment to the recovered (R) compartment. This may reflect immunity that is acquired after infection—an individual is no longer infected, but also not susceptible to reinfection, and so becomes recovered. This is typical for many rhinoviruses or seasonal influenza. In the case of seasonal influenza, individuals can also move directly from susceptible to recovered by means of vaccination.

The compartmental representation of an SIR model with vaccination is shown in Figure 8 and the epidemic curve is seen in Figure 9. Initially when there are many susceptible and no recovered, the number of infected are able to grow. As time proceeds many of the susceptible become vaccinated and are no longer able to become infected (by definition) and hence the number of infected in each time step begins to decline. Eventually all susceptible and infected move to the recovered state.

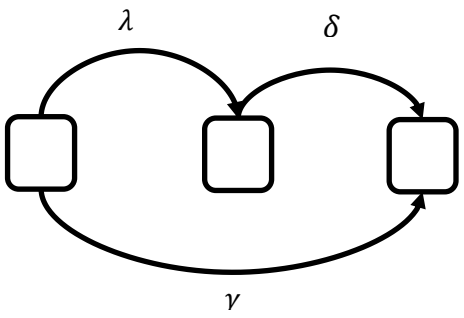

Figure 8: A graphical representation of an SIR model. This models individuals transitions from susceptible (S) to infected (I) to recovered (R). Additionally, individuals may move directly from susceptible to recovered via vaccination or natural immunity.

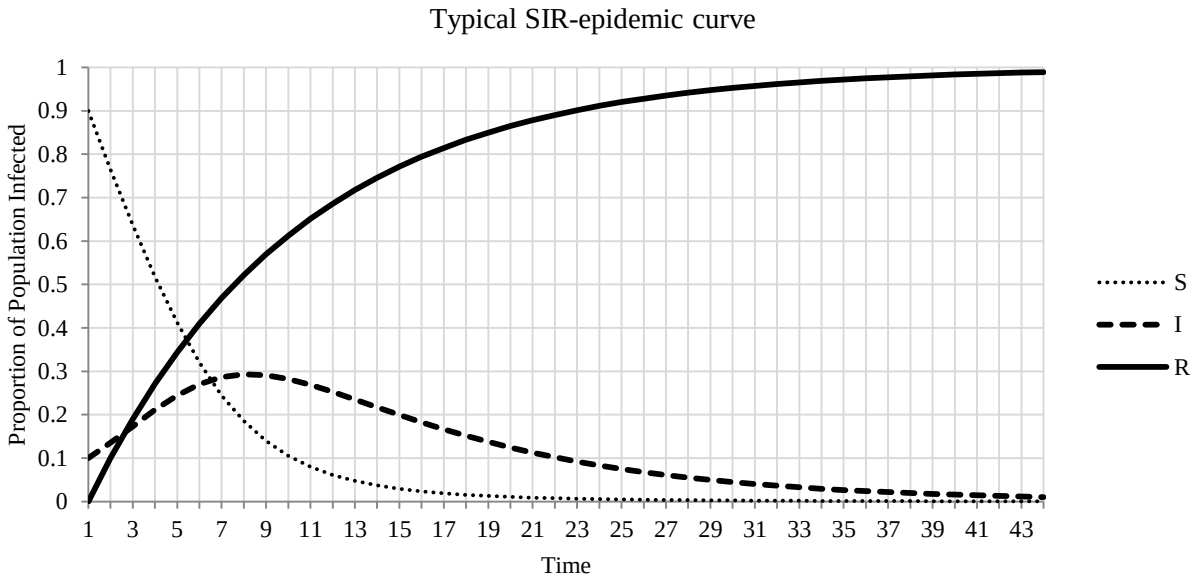

Figure 9: Specific SIR epidemic curve for values $\lambda = 0.5, \delta = 0.1, \gamma = 0.1$. Initially there are many susceptible, few infected, and no recovered individuals. The number of infected grows in the beginning as there are a large number of susceptible individuals. However, as time progresses and the number of susceptible decreases, either through infection or vaccination, less people become infected. Eventually the whole population is recovered and none are susceptible or infected.

Additional complexity can be modeled with additional compartments. There are SEIR models (the *E* stands for exposed) which models diseases in which individuals experience a latent stage of infection like some strains of influenza [39]. This means that they are infectious and able to spread the disease, but do not show symptoms. This latent stage thus makes it difficult to perform interventions like social distancing (isolating infected individuals) or vaccination programs (a vaccine is ineffective if an individual is already sick).

SIRD (the D stands for deceased) models are used for pandemic influenza that have relatively high mortality. The deceased compartment is similar to the recovered stated since individuals in these compartments are unable to cause further infections. However the additional compartment captures disease outcome of individuals in the population. The goal of these models

is to minimize the number of individuals that ultimately end up in the deceased compartment [40].

1.3 Intervening in Disease Diffusion

After a model has been selected and parameters properly set, epidemiologists and public health officials want to investigate strategies for intervening and stopping the spread of disease. In this section we present some compartmental models which investigate strategies for reducing disease burden.

1.3.1 Increasing Access to Anti-Retroviral Therapy

We can expand the simple SI model so that instead of a single infected stage there are several corresponding to varying levels of disease progression. Figure 10 shows a model that uses CD4 count (a proxy for the stage of HIV infection) as infected compartments. Since the lower CD4 levels represent individuals that are more infectious, it is cost effective to start anti-retroviral therapy (ART—a drug regime used to treat HIV) sooner since the cost incurred from treatment is outweighed by the cost of averted infections.

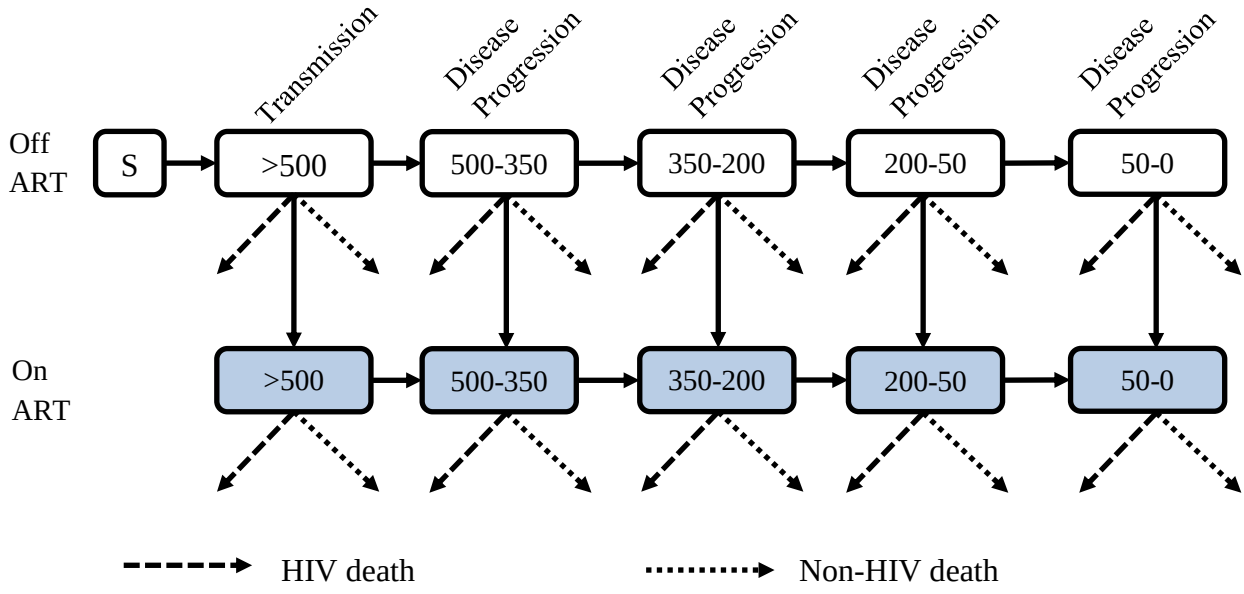

Figure 10: Another model that uses CD4 counts (a proxy for the stage of HIV infection) as infected compartments. Since the lower CD4 levels represent individuals that are more infectious, it is cost effective to start anti-retro viral treatment sooner since the cost incurred from treatment is outweighed by the cost of averted infections.

However, a government does not want to treat just a random subset of the HIV infected population, they typically want to treat the very sickest. Until recently the South African government had the threshold of treating individuals with CD4 count of less than 200 cells / mL [3]. This is the threshold for being diagnosed with AIDS. Using a compartmental model that take into account differences in infectivity due to treatment, epidemiologists at the South Africa Centre for Epidemiological Modeling and Analysis (SACEMA) have shown that increasing the threshold from 200 cells / mL to 350 or 500 cells / mL and specific age targeting would amplify decreasing incidence rates of the disease [41]. While the decrease in incidence seems trivial, SACEMA showed that the extra cost incurred by treating individuals with higher CD4 count levels (less sick) would be less in the long run due to avoided infections. Depending on the lifetime cost of treating HIV, about $12,000 with ART and $3,800 without ART, this could amount to as much as $2.4 million in net savings over the next 20 years.

1.3.2 A Mathematical Model for Optimal Resource Allocation of HIV

Another method for modeling public health decisions considers interventions that aim to interrupt some parameter of disease spread, connectivity or infectivity, at either an individual or population level. We examine the model in [42] to model optimal resource allocation. To begin, let each intervention method i be associated with a certain monetary cost c_i and some effect function $\lambda_i(c_i)$ on the disease incidence. The intervention works by either affecting another intervention or reducing overall infectivity of individuals or connectivity of a community.

With this framework in place we can calculate the number of infections averted over a time T by integrating over the difference of the number of infections that would have occurred without interventions $f(t, \lambda_0)$ and with interventions i, $f(t, \lambda_i)$. Additionally since the model time unit is 1 year, we need to take inflation of cost into account and so an annual discount rate r is used. Note that interventions may be getting cheaper as well and so r may be negative. This is generally taken to be 3% and does not have a large effect on the model.

The function for the number of infections averted then is

$$IA(c_i) = \int_0^T f(t, \lambda_0) e^{-rt}\, dt - \int_0^T f\big(t, \lambda_i(c_i)\big) e^{-rt}\, dt.$$

We can use this equation for the number of infections averted to define the optimization problem. Knowing the benefit from each averted infection W, and the number of infections averted $IA(c_i)$ from spending c_i the optimal resource allocation problem is

$$\text{Maximize } W \times IA(c_i) - c_i$$

$$\text{such that } c_i < B$$

where B is the budget for spending. The parameter W, also known as the *willingness to pay*, is a measure of how much a single averted infection is worth to the intervening party (most often the government). This metric can be captured in a number of ways: Quality of Life Years

(QALYs), a measure of economic output for a typical individual, or the ratio between Disability Adjusted Life Years (DALYs) and per Capita GDP.

Since there is only one variable c_i (how much to spend on intervention *i*), we can solve this problem analytically. For example, consider a population of 10,000 injection drug users (IDUs) in which the prevalence of HIV is 40% and sufficient contact rate $\lambda_0 = 0.0817$ [53]. Additionally consider a needle exchange program *ne* that changes the sufficient contact rate of HIV by a multiplicative factor relative to the amount spent:

$$\lambda_{ne}(c) = \lambda_0 \left[0.67 + 0.33e^{-0.0089\left(\frac{c}{N}\right)}\right].$$

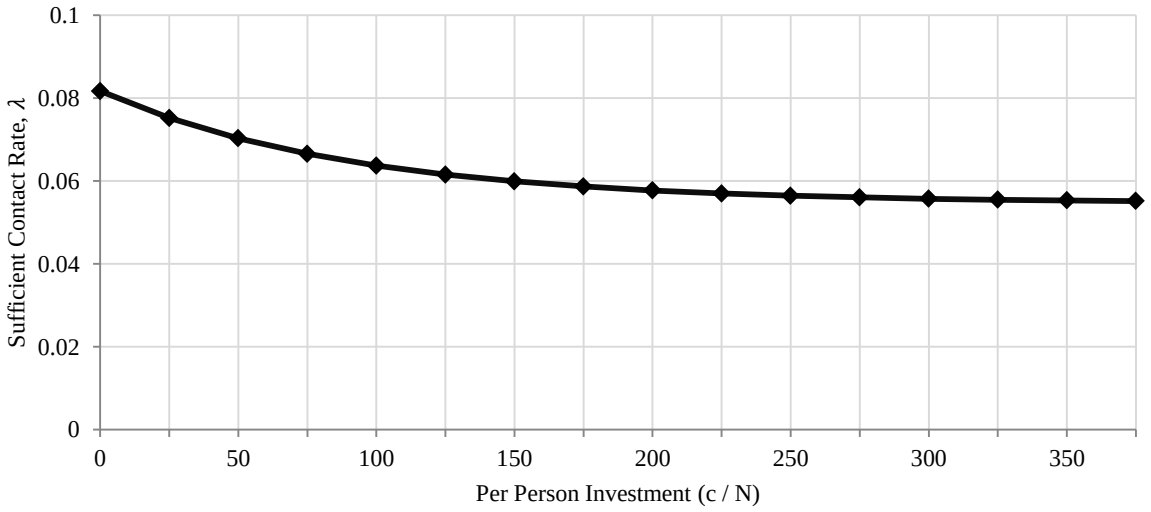

Figure 11: The production function for different levels of investment. The function exhibits decreasing returns to scale—each additional dollar spent provides less benefit then the previous. If no money is spent (*c = 0*), then the infectivity (sufficient contact rate) is 0.08. If $120 per person is spend, then the infectivity is approximately 0.06.

In Figure 11 it is easy to see that this particular intervention has a decreasing return to scales: each additional dollar yields less benefit. In the case of a needle exchange program, once an initial willing population has been located, additional willing participants may be difficult to find. With this production function we can use the epidemic function *f* (defined earlier) to generate the function *IA*.

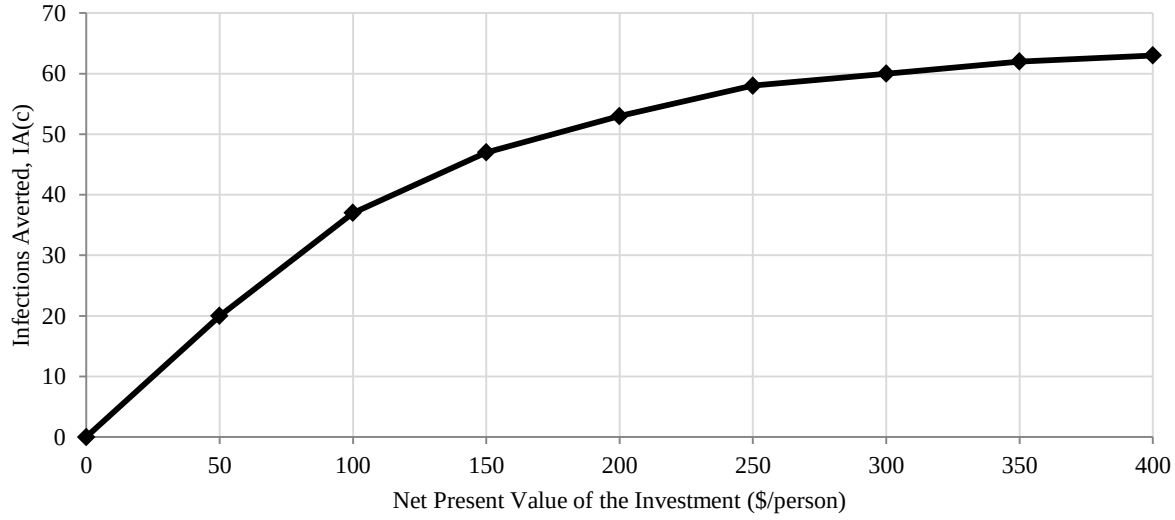

Figure 12: Infections averted for different values of investment. Increasing the investment per individual will increase the number of infections averted, but with decreasing return to scales. Spending at $120/person will avert approximately 40 infections.

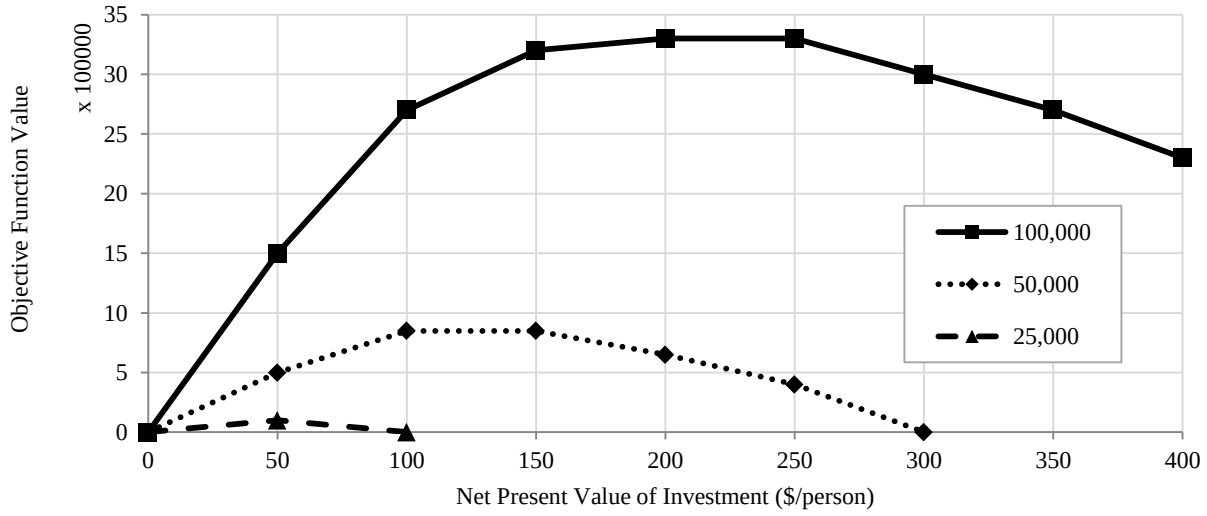

Figure 13: The objective function for different values of willingness to pay. The objective function has a greater optimal investment for greater values of willingness to pay *W*. For *W=$50,000* the optimal amount to spend is $120 per individual, which is $1.2 million in a population of 10,000 injective drug users.

To find the optimal amount to spend we use the above *IA(c)* function in the objective function as seen in Figure 13. The objective function when *W = 50,000* is maximized when the cost per person is $120. This is found exactly by taking the derivative of the objective function

with respect to cost and solving the equation set to zero. In words this means that if an infection costs $50,000 to the intervening agency (either through lost productivity, increased health care costs, etc.) then the optimal amount to spend is $1.2 million ($120 times the 10,000 person population). If less than this amount is available, all of the budget should go towards the needle exchange program. If more than this is available, the program should be allocated $1.2 million and the rest be re-appropriated to another intervention method.

Perhaps not surprisingly, when the benefit of an averted infection is more (W=$100,000), the optimal cost is higher. This makes sense as a greater benefit justifies the higher cost. Similarly, lower values for the sufficient contact rate yield lower optimal expenditures since the disease is less likely to be spreading. The same is true for lower prevalence: a population with a lower prevalence requires a lower optimal expenditure.

This mathematical model is limited in its scope however: it does not consider the benefit of many interventions implemented in combination. While it is possible to find a combination of several interventions through individualized analysis, the solution is not guaranteed to be optimal, nor is it likely to be. This is because intervention methods often interact with each other through mixing of target populations and referral to other interventions. For example, in the absence of all other interventions, HIV counseling and testing may convey little or no protective effects for uninfected individuals. When utilized alongside a national male circumcision program, however, counseling and testing may become a point of referral and a catalyst for the male circumcision program.

1.3.3 Optimal Resource Allocation for Multiple Intervention Methods

In the case of multiple interventions the simple SI compartmental model needs to be expanded so that target populations of interventions are each modeled. Additionally, the interaction between interventions is modeled via transition parameters between the different

populations. Zaric and Brandeau consider several interventions targeted at IDUs, IDUs on Methadone maintenance (a drug regime that relieves heroin withdrawal symptoms), and non-IDUs [43]. Note that Methadone maintenance has been shown to be very helpful in reducing heroin addiction (and consequently injection drug use) and so slots for the free drugs are most often full. Specifically they considered the effect of the following interventions:

1. Needle exchange for all IDUs;
2. Increasing the number of Methadone maintenance slots for all IDUs;
3. Increasing the number of Methadone maintenance slots for IDUs with HIV;
4. Increasing the number of Methadone maintenance slots for IDUs with AIDS;
5. Condom distribution to IDUs
6. Condom distribution to IDUs in Methadone maintenance
7. Condom distribution to the entire population

They model the three populations with a compartmental model where individuals progress through the disease stages non-infected, infected with HIV, infected with AIDs (Figure 14).

The transitions between compartments are based on the infectivity and size of the different compartments as in the previous model. Now however there are many infectivity constants for many different transitions and population interactions. Thus in addition to initial population size, the effects of each intervention need to be considered: whether through reducing infectivity (condom distribution, needle exchange), or reducing connectivity (methadone maintenance reducing IDU population). A detailed description of each intervention production function can be found in [43], but is omitted here for brevity.

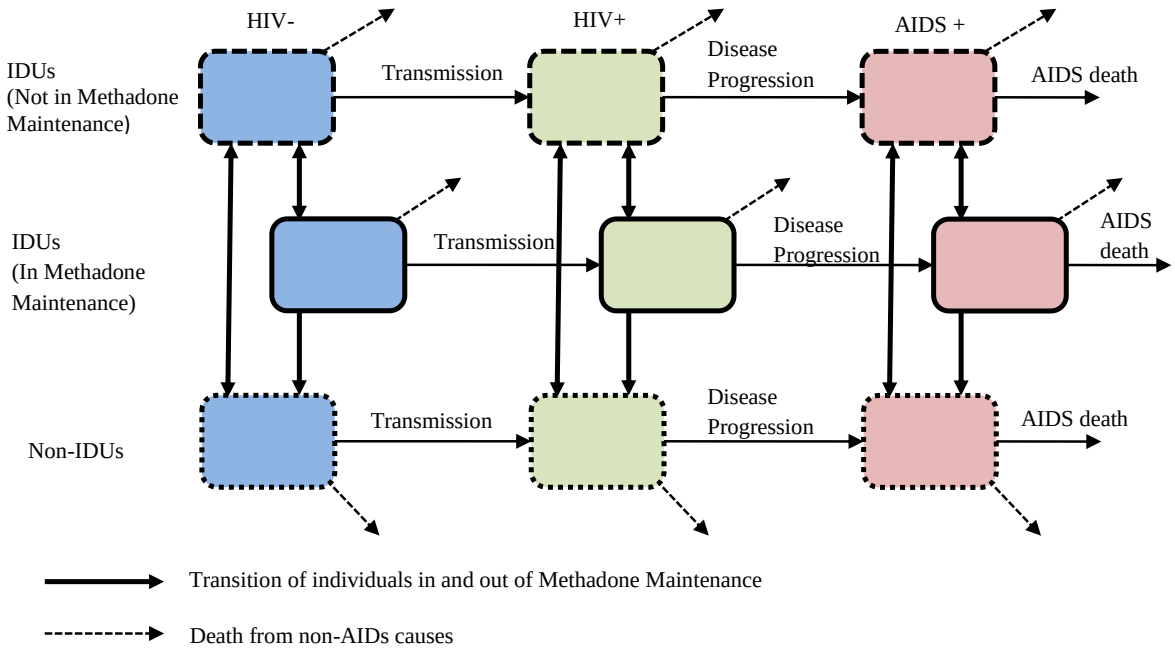

Figure 14: In order to find the optimal resource allocation of a portfolio of intervention methods, each of the target populations are modeled with an SI model. In this case the three populations are IDUs not in methadone maintenance, IDUs in methadone maintenance.

Once the compartmental model and the effect of different interventions on the model have been put into place we can generate the epidemic curve for specific allocation of resources (no spending resources being the base case). However, the nonlinear nature of the model makes a closed form solution unlikely, if not impossible. The resource allocation problem for multiple interventions then becomes a continuous knapsack problem which is known to be NP-hard [44]. Fortunately optimization theory and heuristic search allow us to find feasible, though not necessarily optimal, solutions to the problem.

1.3.4 Optimal Resource Allocation for Influenza Outbreaks

We can use influenza surveillance and intervention models as a catalyst for simulating and intervening in STD diffusion and perpetuation due to their similar nature: both are infectious diseases spread through contact (albeit a different mode of contact), and both are easily

preventable (influenza through vaccine and STDs through safe sex measures). In this way existing work may be applied to STD models of diffusion.

It should be noted that the differences between the two types of diseases are more than nominal: new influenza strains occur annually and so appropriate vaccines must be created. Transmission from an influenza infected individual to a susceptible individual can occur through casual contact (i.e., in a crowded market, or closed-system airplane). Most dissimilar is that influenza models tend to emphasize the diffusion of the disease—how influenza may spread across a country [45, 46]—whereas models of STDs are more concerned with the perpetuation of the disease. That is to say, influenza epidemiologists typically aim to isolate new strands so they do not infect a large percentage of the population. For STDs like syphilis and HIV many people are already infected and so models aim to reduce the *incidence rate*, the number of new cases in the population. However, we maintain that influenza models offer themselves as a proxy for infectious disease spread.

Ludkovski and Niemi considered the optimal resource allocation problem for disease interventions in a non-deterministic model [40]. Specifically, they consider the spread of flu within a boarding school of 763 students with two students initially infected. They simulate the epidemic with an SIR model using the Gillespie algorithm. This is a variation of the generic SIR model described earlier that uses continuous time steps instead of discrete, and non-deterministically simulates events. At every step, a value for τ (exponential distributed) is sampled and one of two "events" occurs – an individual moves from susceptible to infected, or an individual moves from infected to recovered. The propensity of each event is relative to the infectivity, λ_0, and recovery rate, δ.

For example, let $X(t) = \big(S(t), I(t), R(t)\big)$ be a triple that represents the number of susceptible, infected, and recovered in the system at time *t* respectively. The state of the epidemic is updated relative to

$$X(t+\tau) = X(t) + \begin{cases} (-1, 1, 0) \text{ with probability} \propto \dfrac{\lambda S(t)\, I(t)}{N} \\ (0, -1, 1) \text{ with probability} \propto \gamma I(t) \end{cases}$$

for a population of N and $\tau \sim Exp\left(\frac{\lambda S(t)\, I(t)}{N} + \gamma I(t)\right)$. They note that for large N (>1000) the model is essentially deterministic through the law of large numbers.

They assume that every day a decision can be made about what action to take: begin a vaccination campaign, isolate infected individuals (and incur some cost through lost productivity), or wait and see. The wait-and-see decision allows the policy maker to gather more information such as infectivity and recovery (and hence the basic reproductive number, R_0). The epidemic simulated many times for each of the interventions and a range of coefficient values to find a policy map (Figure 15) that minimizes the expected cost.

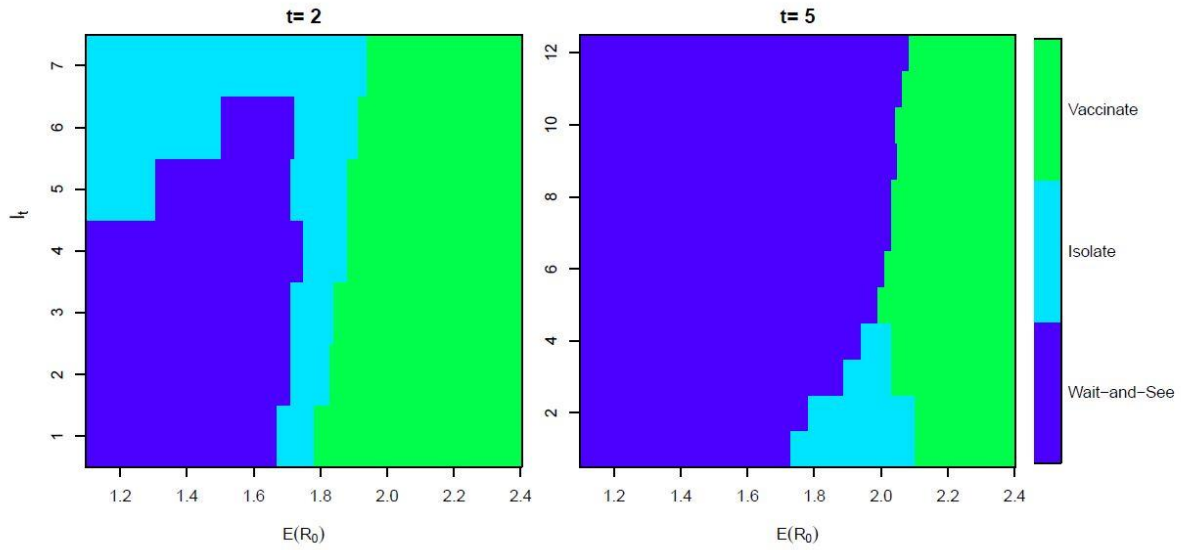

Figure 15: The optimal intervention is based on the expected basic reproductive number and the number of infected. If the basic reproductive number and the number of infected is small than the optimal strategy is to wait-and-see. If the basic reproductive number and the number of infected are high the optimal strategy is to vaccinate.

These policy maps show that the optimal resource allocation depends on the number of infected, the expected basic reproductive number, and the time of implementation. Ludkovski and Niemi take into account error in sampling methods that inform these values as well as perform sensitivity analysis. Their main contribution is this methods ability to evaluate and suggest an optimal allocation in real-time. This is necessary for real world influenza epidemics.

1.4 Agent-Based Models

Recently modeling efforts have shifted to agent-based simulations. These models simulate populations of individuals with agent-specific characteristics. The models allow agents to have interactions based on these characteristics and produce emergent behavior not typical captured by non-stochastic models. Agent-based models for sexually transmitted diseases simulate sexual relationships between agents, using the agent specific characteristics to create a dynamic sexual network. In this way, agent-based models are able to simulate how the disease diffuses through a network, and simulate possible actions to disrupt this process.

While it's clear that large scale network modeling is necessary to obtain robust results, it is not immediately obvious why previously developed agent-based models, e.g. of influenza, cannot simply be translated to apply to sexually transmitted diseases. The first reason for this is that many agent-based influenza models assume the contact network is known a priori [46, 47]. In agent-based STD models, contacts are frequently changing and hence these models must simulate the dynamic sexual network at the same time as disease diffusion. The second reason is that the possibility of infection is unique to each agent since the sexual partners of an agent are particular to that agent. This is not the case in agent-based models of influenza: the possibility of infection is specific to the location where an infected agent is found. All agents that are in the same location as an infected agent share the possibility of infection: in this way large-scale influenza models are able to aggregate infection events to specific locations [48, 49]. Because sexual encounters are not based on repeated random selection of prospective partners at a given time and location, large-scale models of influenza do not lend themselves to be used for large-scale models of sexually transmitted disease.

The most well-known agent-based simulation of HIV is STDSIM [50]. This particular model has been used to evaluate interventions for mass treatment of STDs [51], behavior change campaigns [52], condom distribution [53], and male circumcision[54] to name just a few. Auvert et al. used the agent-based model SimuAIDS to examine the relative importance of sexual behavior and biological factors on the spread of HIV [55]. Sloot et al. created the model Complex Agent Network (CAN) [56]. This model applies the research area of complex networks and applies it to agent-based simulation. In their discrete time step model, they impose a distribution of relationship durations and a power-law degree distribution for desired number of partners. They track incidence and prevalence over the order of several years and validate their model versus incidence of men-who-have-sex-with-men in Amsterdam.

CHAPTER II AGENT-BASED MODELING OF STDS

2.1 Introduction

Diffusion dynamics of sexually transmitted diseases are been influenced by sociological effects as discussed in the previous chapter. While compartmental models are very good at describing general epidemic trends, they can have difficulty modelling complex social phenomena and the interactions among them. For this reason, agent-based models are used to simulate individual-level behaviors and to gain insight as to how they may be interacting to contribute to disease dynamics.

In this chapter, we present a mathematical formulation for modeling HIV. Our goal is not to provide a fully validated model, but instead to show that this formulation can reasonably model many common disease-related sociological processes. We first we provide a non-exhaustive background summary of important sociological effects contributing to the epidemic. In the second section, we describe the mathematical formulation for modelling these effects. In the third section we show through simulation output that this framework can reasonably model many sociological processes including complex age-mixing patterns; a heterogeneous population of female sex workers, men-who-have-sex-with-men (MSM), and heterosexual agents; and society level changes in condom use behavior. We conclude in the final section with a discussion of the significance of the work and directions for future research.

Note that through-out the chapter we use the term “individual” and “agent” to distinguish between a real person in the world and a simulated person in our model respectively.

2.2 Background

One of the difficulties in HIV modeling is accounting for the multitude of behavioral changes at the societal level, and the myriad of changes to HIV response at the governmental

level. For example, evidence suggests that as knowledge about the existence of HIV proliferated through the country, individuals began to use condoms more frequently [9]. However, no there is little formal evidence and thus it's difficult to know the extent to which condom use affected the epidemic.

The high prevalence of age-disparate relationships among young women means that HIV is able to leap between generations with relative ease. Efforts have been made to discourage young women from forming high-risk relationships with older men, colloquially referred to as "sugar daddies", but gains have been minor due to the practice having relatively high societal acceptance [57–59].

The probability of transmitting HIV to a sexual partner changes over the course of infection, but is highest during the first three months. This fact has led to a debate in public health over the role which concurrency and partner turnover rates play in the epidemic [60, 61].

Poverty in general has socio-economic implications for HIV transmission. In addition to have decreased access to health care, poor individuals are more likely to have stressful jobs that are closely correlated with alcohol consumption and risky sexual behaviors. In some cases alcohol may even be used as currency for sex [24, 25]. Besides alcohol-for-sex and age-disparate relationships, women face a multitude of additional risks. Having less power in society, they are often unable to dictate the use of condoms in relationships, and are often the victims of rape [62]. Women typical are unable to end a relationship with an unfaithful partner, increasing prevalence of concurrent relationships in the sexual network and hence opportunities for HIV to spread.

Many of these processes have been modelled independently to understand their effect on the epidemic. However, it is becoming increasingly clear that the best responses to the epidemic will need account for these processes simultaneously.

2.3 The Mathematical Formulation

In this section we describe our mathematical formulation for a discrete-time, agent-based simulation. To explore the usefulness of this formulation, we implemented the model with the multi-agent simulation toolkit MASON[63]. We first describe the overall flow of the algorithm, and in subsequent sections describe how the model is flexible to additional levels of complexity to model complex sociological phenomena.

The time step of the simulation is one week. Each week the model progresses with three steps: (i) relationship formations and dissolutions, (ii) infections occur, and (iii) agents are removed and added. In short, agents form relationships based on individual characteristics such as gender, age, and desired number of partners. Transmission of HIV is controlled by the Infection Operator, and the progression of time is controlled by the Time Operator (individuals age are incremented, relationship durations are decremented).

The model initializes agents each with a sex, an age, and desired number of partners (DNP) according to a prior distribution. At each time step, we ask each agent two questions: would he or she like to form a relationship? If so, with whom? The heterogeneity of our model comes from the fact that each agent answers these question based on different criteria. For example, one agent may seek new relationships if their number of partners is less than their desired number of partners and he or she may want to form relationships with agents of the opposite sex (we will discuss other relationship forming rules in subsequent sections). The duration of a relationship is determined at formation as a random value taken from a prior distribution.

After we allow each agent to form a relationship, the Infection Operator performs initial infections, increments the number of weeks infected (for infected individuals only), and performs

infections between sero-discordant couples in the simulation. The Time Operator next increments the agents' ages, and decrements the duration of relationships in the network. If the duration of a relationship becomes negative (i.e., it has ended), the edge in the network is removed and the respective agent's number of partners are decremented. On the next step the agents are allowed to try to find a new partner. Each week, a fraction of individuals are removed and some new agents are born to replace them.

We repeat this process of first calling agents to form relations, second performing infections, and finally progressing time for the duration of the simulation. This process produces a dynamic sexual network and simulates the diffusion of HIV through a heterogeneous population. Figure 16 shows the pseudo-code for the algorithm.

While the model as described above is straightforward, we will show that this framework can be expanded to include more sophisticated processes for forming relationships, controlling demographic processes, and modeling disease characteristics.

```
Algorithm SimulateHIV

1:  initialize_population()
2:  repeat
3:      //agents form relations
4:      for agent from 1 to N do
5:          if agent.is_looking() then
6:              for other_agent from 1 to N do
7:                  if agent.is_looking_for( other_agent ) then
8:                      form_relationship( agent , other_agent )
9:              end other_agent for
10:     end agent for
11:
12:     //perform infections with operator
13:     infection_operator.perform_infections()
14:
15:     //progress time with operator
16:     time_operator.progress_time()
17: until time > endTime
```

Figure 16: Pseudo-code for the SimpactBlu algorithm. At each step, three things happen: (1) agents with less than the desired number of partners form new relationships; (2) Time progresses such that agent's ages are incremented and relationship durations are decremented by one week; (3) Infections occur in sero-discordant relationships.

2.3.1 Probability of Relationship Formation

We define a directional probability function P_{ij} as the probability that agent i forms a relationship with an agent j. Note that P_{ij} is not necessarily equal to P_{ji} since j may have different partner preferences (e.g. he or she may be interested in relationships with a narrower age gap). Additionally this probability is only relevant if both partners are interested in forming a new relationship (i.e., each had less than their DNP).

Consider a simple probability function applied to every agent which considers only the absolute age difference between two agents:

$$P_{ij} = e^{\alpha \cdot |\Delta \text{ age}|}$$

Where α is the probability multiplier. Even this simple probability function applied to every agent uniformly can yield the desirable result that the age difference in most relationships is relatively small. Figure 17 shows the probability of a relationship forming for different age differences and probability multipliers.

Figure 18 shows the age mixing scatter for a probability multiplier of -0.1. Each dot represents a potential relationship. The dot's x-value is the age of the male, and the y-value is the age of the female. The color of the dot is the probability the relationship forming with the above probability function and probability multiplier.

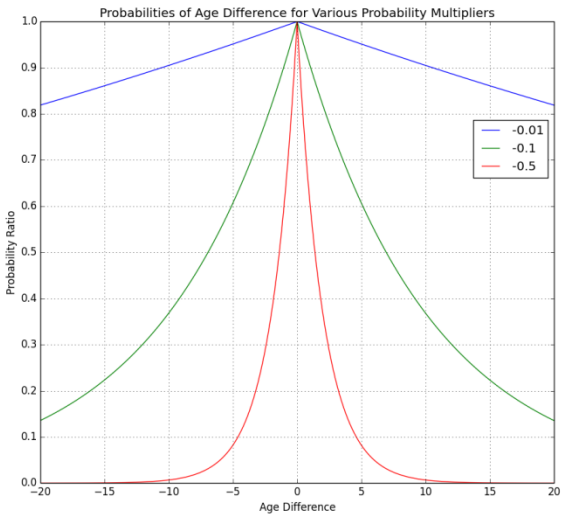

Figure 17: Probability of relationships formation for different probability multipliers. Age-disparate relationships can be made more or less likely this way.

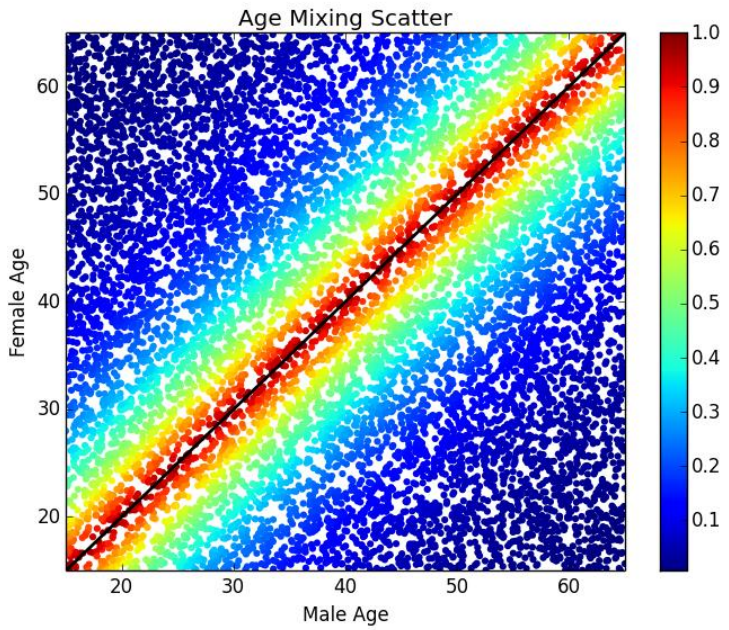

Figure 18: Age mixing scatter for a simple probability function and a probability multiplier of -0.1. Though simple, this probability function can produce age mixing patterns similar to those seen in the real world.

We can begin to add layers of complexity to the model by adding other factors into the probability function. For example, in addition to wanting to form relationships with agents of a similar age, agents are less likely to form relationships in general as they get older. To model this this we add an additional term that scales the probability of relationship formation based on the candidate couple's mean age. Hence, the probability function would be $P_{ij} = e^{\alpha_1 |\Delta \text{ age}| + \alpha_2 \cdot \text{mean\_age}}$ and the resulting age-mixing scatter would be Figure 19.

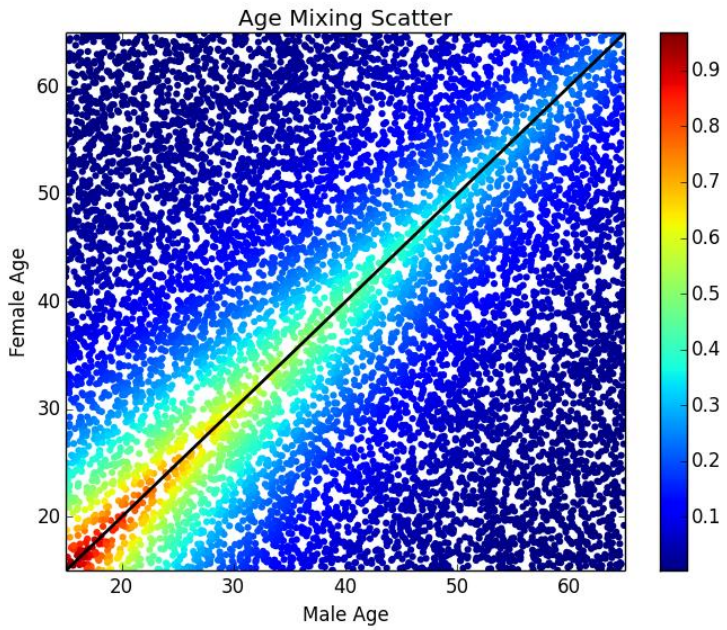

Figure 19: The age mixing scatter for a probability function that decreases with the mean age of the candidate couple. This reflects the real-life situation in which younger individuals form more relationships than their older counterparts.

These two simple examples have shown the usefulness of a generalized probability function: it offers flexibility as to which characteristics are significant in relationship formation, and by what amount. Figure 20 shows the age mixing graph for the probability

$$P_{ij} = e^{\alpha \cdot (|\Delta \text{ age} - \text{preferred\_age\_difference} \cdot \text{mean\_age}|)}$$

Where α is again the probability multiplier. Additionally the probability subtracts a preferred age difference from the actual age difference – this reflects that female agent may actually prefer an older male partner (perhaps for maturity or for economic security). The probability function multiplies the preferred age difference and the mean age of the couple to generate a larger preferred age difference for older couples. This reflects the fact that as men grow older, they increasingly prefer younger women.

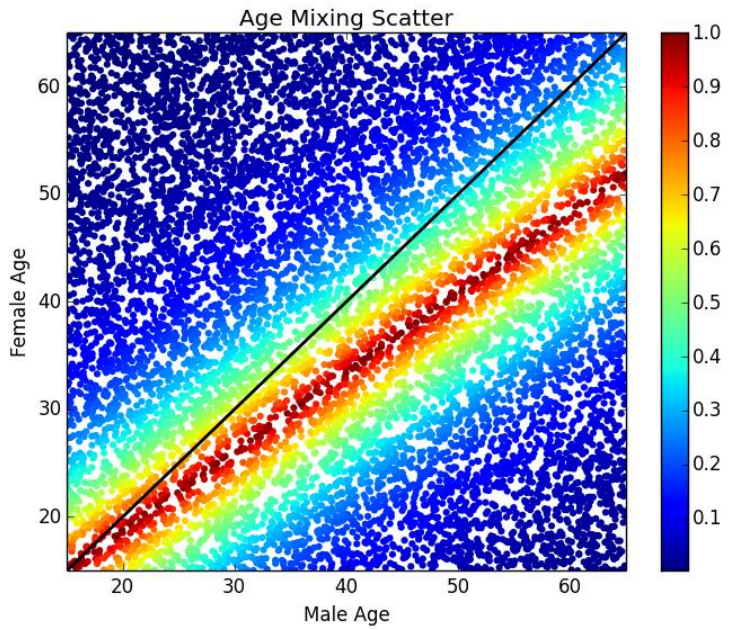

Figure 20: The age mixing scatter for a more complex probability function. This probability function additional considers that there is a preferred age difference which grows with mean age (*PM* = -0.1, preferred age difference = -0.2, preferred age difference growth = 1.5).

Let us finally consider the possibility that in addition to a preferred age difference that is larger for older couples, the preferred age difference becomes more dispersed for older couples. The following probability function models this idea:

$$P_{ij} = e^{\alpha \cdot \left(\left| \Delta \text{ age} - \frac{\text{preferred\_age\_difference} \cdot \text{mean\_age} \cdot \alpha_{\text{growth}}}{\text{preferred\_age\_difference} \cdot \text{mean\_age} \cdot \alpha_{\text{dispersion}}} \right| \right)}$$

Note that this equation is of the same form as the other, except the preferred age difference now grows and becomes more dispersed as the mean age of the couple grows. Figure 21 shows the resulting age mixing scatter plot.

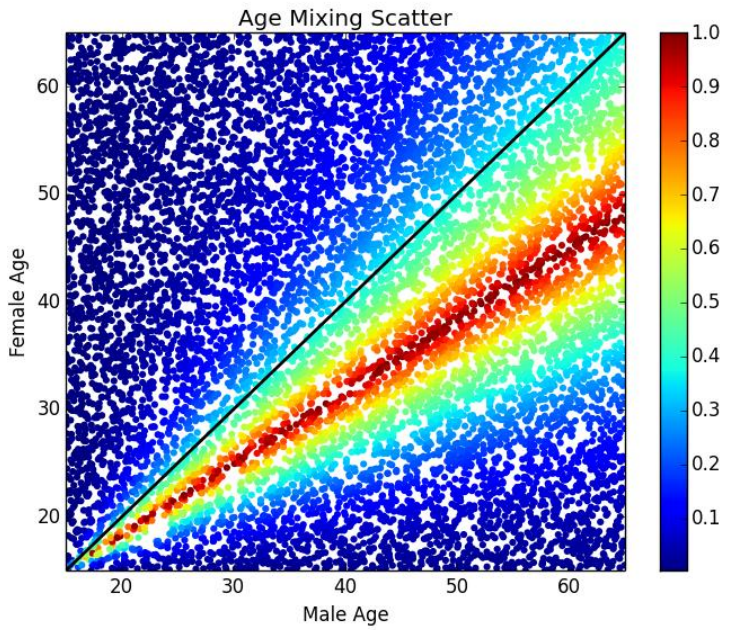

Figure 21: How preferred age difference can change with dispersion and growth. Here the baseline preferred age difference is -0.2, preferred age dispersion is -0.2, preferred age growth is 2.0, and the probability multiplier is -0.1.

The above figures showed the theoretical probabilities of relationship formations. Figure 22 is output from the model implementation and shows the flexibility of the model for simulating different age mixing patterns. We run each scenario for 1 year with a population of 1000 agents. For purposes of visualization duration of relationships was 10 weeks.

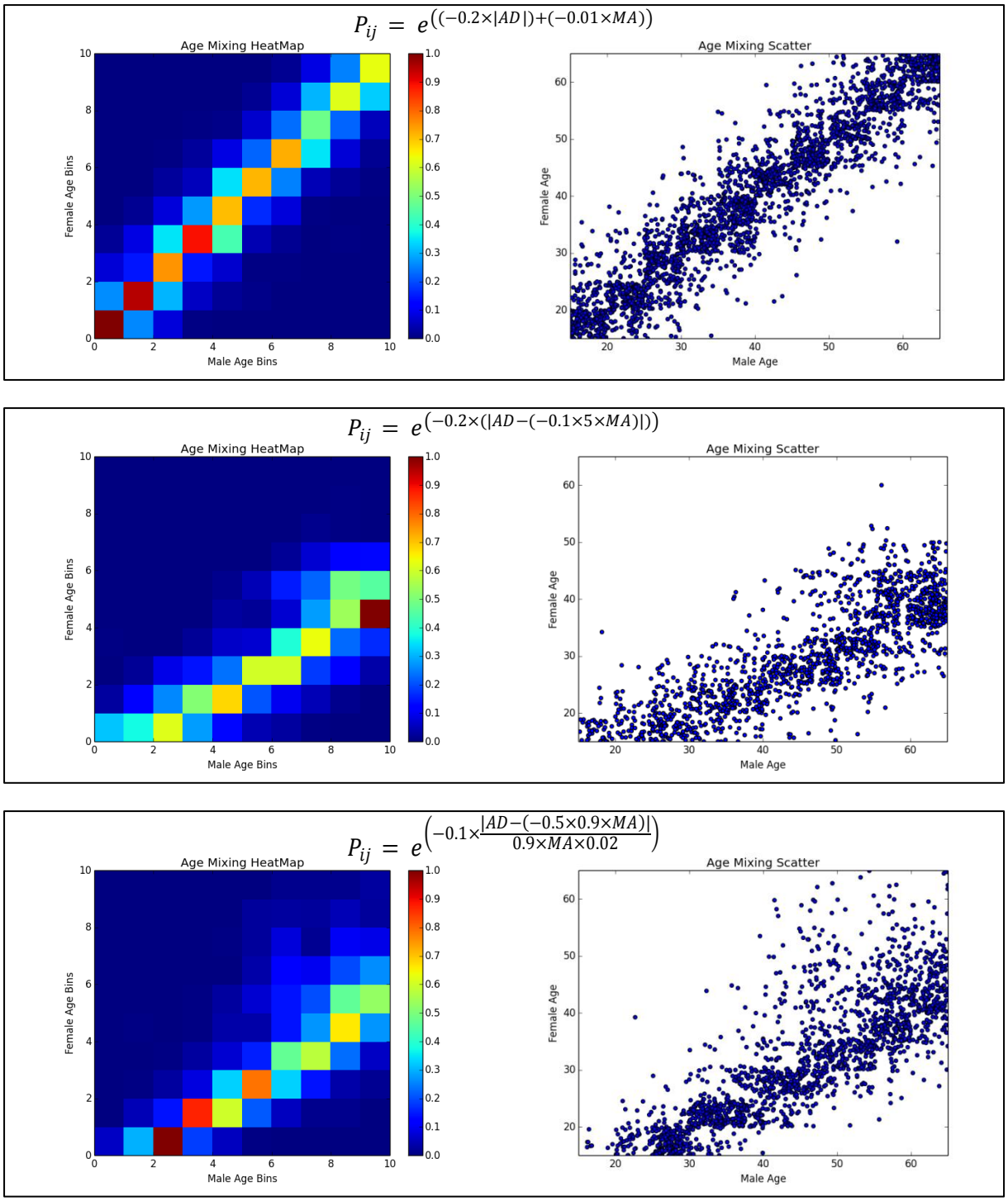

Figure 22: Age-mixing heat map and scatter for three different probability functions. Top: the simplest probability function that produces many relationships with agents of a similar age. Middle: a more complex probability function that produces relationships in which older men are paired with younger women. Bottom: the most complex probability function that produces relationships in which age matters less for older men.

We have shown a few different probability functions and the age-mixing patterns that these functions produce when applied to a whole population. In practice we create a heterogeneous population of agents, each with a probability function which governs the agent's personal behavior. In the model the population is defined by the proportion of different types of agents.

Note that some individuals form relationships independent of age. For example, men-who-have-sex-with-men (MSM) are less discerning of large differences in age in potential partners [56]. Female sex workers are likely to have sexual relationships with a wide range of ages – their discerning factor is the potential partner's ability to pay.

Table 2: The different types of agents and their associated probability function.

Agent Type	**Probability function**	**Notes**
Basic	$P_{ij} = 1$	This agent forms relationships independent of age – relies solely on his or her desired number of partners.
Cone	$P_{ij} = e^{\alpha \cdot \left(\left| \mathrm{AD} - \frac{\alpha_{PAD} \cdot \mathrm{MA} \cdot \alpha_{\mathrm{growth}}}{\alpha_{PAD} \cdot \mathrm{MA} \cdot \alpha_{\mathrm{dispersion}}} \right| \right)}$	α is the probability multiplier, α_{PAD} is the preferred age difference, α_{growth} and $\alpha_{dispersion}$ are preferred age difference growth and dispersion respectively. *AD* and *MA* are age difference and mean age.
Triangle	$P_{ij} = e^{\alpha \cdot \lvert AD - \alpha_{PAD} \rvert}$	Same as above
MSM	$P_{ij} = 1$	Always male and only forms relationships with other MSM agents
FSW	$P_{ij} = 1$	Has DNP of 16 and relationships only last 1 week

Table 2 provides an overview of all the agents used in our simulations along with the probability function they use to form relationships. All agents seek new relationships if their

number of partners is less than their DNP. All agents have a DNP draw from a power distribution except FSW agents who have a default of 16 (the average number of clients a typical FSW will have in a week)[64]. Note that implicit to all agent's probability functions is a variable indicating whether other agent is the correct sex for the agent's sexual orientation.

2.3.2 Operators

Though how agents form relationships is of obvious significance in disease diffusion, there are additional processes that influence the epidemic. For example, it is unlikely that an agent will form relationships based on the time since he or she became infected with HIV. However, since viral load peaks during the first few months of infection, recently infected individuals are more likely to transmit to their partners. Here we describe the simulation operators that control the various processes beyond relationship formation. As we did for the probability function of relationship formation, we first describe a simple implementation of the two operators used in our model, the Infection Operator and the Time Operator, and then show how they can be expanded to model processes that are more complex.

The main role of the Infection Operator is to propagate infection through network. At each time step, the Infection Operator iterates through the edges of the network and probabilistically transmits infection from HIV-positive agents to their HIV-negative partners. In the simple model, the probability of transmission is a constant value that does not change with time or individual. The Time Operator enforces the passage of time in the simulation by incrementing the age of all agents by one week. In order to maintain a constant size population the Time Operator removes agents from simulation when they are 65-years-old and adds a new 15-year-old agent to the simulation to replace them.

In the following sub-sections we discuss modification to the operators so that our model can more accurately simulate real behaviors.

As mentioned previously, the probability of HIV transmission varies with an individual's viral load. We modify the Infection Operator so that the infectiousness of an agent varies depending on his or her stage of infection. During the first 12 weeks, called the primary stage, an agent infects his or her partner with probability 0.032. After this the agent enters the latent phase for 384 weeks (approximately 8 years) and an agent infects his or her partners with the lower probability 0.0035. After this the agent enters final phase and infects his or partners with probability 0.0152 [16].

Additionally, the time until death for an individual infected with HIV, unless treated with ART, is about eight years, depending on the age of the individual. Therefore, a young agent infected in our model should not transmit to partners until she is removed at 65-years-old, but instead should be removed sooner. To model this, when an agent becomes infected, we assign a random number drawn from a Weibull distribution with scale 2.25 and a shape which is a function of age. This is consistent with data [65]. Figure 23 shows the distribution of time-until-death for agents of different ages.

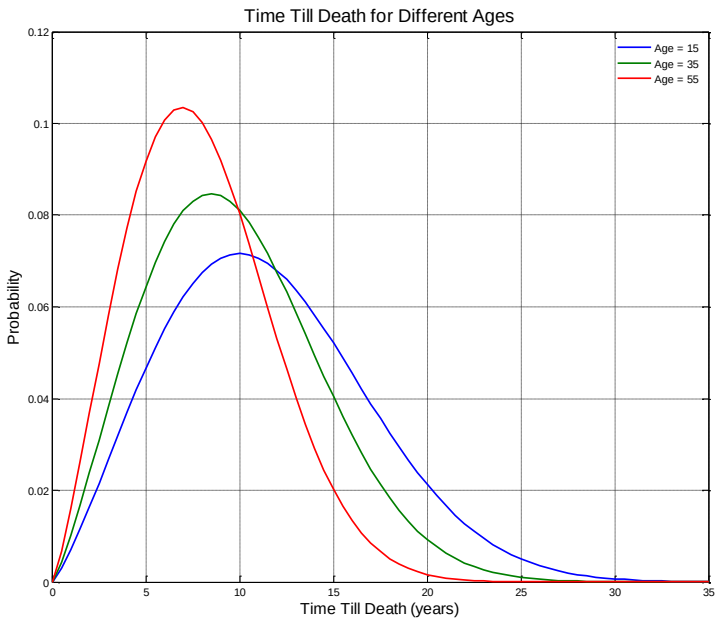

Figure 23: Time until death is drawn from a Weibull distribution with a scale of 2.25 and a shape that depends on age. Individuals that are younger at the time of infections are likely to live longer than their older counterparts are.

Similarly, non-AIDS-mortality is not a constant 65 years old. Moreover, the size of the population is not constant, but instead is constantly growing. We modify the basic Time Operator so that every year it removes a fraction of agents and adds a non-constant number of 0-year-old agents. ASSA2003, a demographic model produced by the Actuarial Society of South Africa, determines the fraction of agents removed based on age and sex mortality tables. ASSA2003 also determines the number of new agents based on the female age fertility tables from ASSA2003. The new agents enters the population at age 0, and are assigned an agent type (e.g. basic, cone, etc.) based on the population's type distribution (as discussed in the previous section).

2.3.3 Behavior Change

To account for the condom behavior change, our model includes gradual increasing condom use starting in the mid-1990s and peaking in the mid-2000s. Since exact values for the start date, end date, and maximum condom coverage are unknown, we use values that are

reasonable given the data [9]. Figure 24 shows our models assumption about the level of condom coverage: condom use begins at 0% in 1998, and reaches a peak of 15% in 2005.

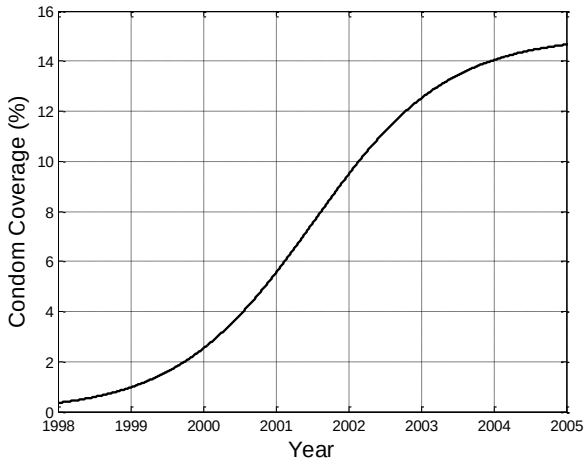

Figure 24: Individuals began using condoms as knowledge about HIV spread. Our simulation assumes a smooth increase in condom use from the mid-1990's to a peak around 15% in the mid-2000's.

Condom coverage of X% implies that X% of the population has their infectivity reduced (if they are HIV-positive) by 80%. While correct and consistent condom use may reduce infectivity by virtually 100%, this more modest value reflects incorrect or inconsistent use [66, 67].

In order to account for ART availability and the life-prolonging and infection reducing benefits, we modify our Infection Operator. When an agent becomes infected, in addition to being assigned a time of death, he or she is given a CD4 count at infection (Normal(1000, 250)), and a CD4 count at time of death (Normal(75, 25)) [3]. With these three pieces of information, we can interpolate an individual's CD4 count anytime between time of infection and time of death (assuming a linear decline in CD4 count).

We model the roll-out of ART with another operator. This operator proceeds in two steps every 4 weeks: (1) the operator tests a fraction of the population for HIV. If an agent is HIV-

positive and her CD4 count is below the threshold for treatment she is placed into the treatment queue. (2) If slots for treatment are available, the operator fills the slot with a patient waiting in the treatment queue. This is akin to these very sick individuals coming in to a clinic in order to receive treatment. In order to model the slow evolution of the availability of ART, the number of slots available increases gradually and smoothly starting at 10 slots in 2002 until 300 slots in 2013.

Table 3: Parameters used in the initial simulation model.

Parameter	Value	Unit	Notes / Justification
Simulation Constants			
Number of Years	30	Years	HIV was introduced to South Africa around 1985. We simulate until 2015.
Relationship durations	Power(52 , 4.2)	Distribution	
Desired number of partners	Power(8 , 10)	Distribution	
Sexual debut	15	Age	The age at which individuals first are able to form sexual relationships.
Population Constants			
Initial ages	Empirical	Distribution	
Initial population size	1000	Individuals	Largest population we can run in a reasonable amount of time
Proportion of MSM agents	0.04	Proportion	
Proportion of FSW agents	0.04	Proportion	
Proportion of agent type 1*	0.368	Proportion	
Sex	Male		
Type	Cone		Form relationship based on age difference.
*Preferred age difference**	0.9	Years	
*Probability multiplier**	-0.1		
*Preferred age difference growth**	0.05		
*Age difference dispersion**	0.004		
Proportion of agent type 2*	0.092	Proportion	
Sex	Male		
Type	Basic		Form relationships independent of age difference.
Proportion of agent type 3*	0.23	Proportion	
Sex	Female		

Table 3 continued.

Type	Triangle		Form relationship based on age difference.
*Preferred age difference**	2.0		
*Probability multiplier**	-0.9		
Proportion of agent type 4*	0.23	Proportion	
Sex	Female		
Type	Triangle		Form relationship based on age difference.
*Preferred age difference**	2.0		
*Probability multiplier**	-0.5		
Infection Operator			
Initial number of infected	2	Individuals	
Time of seeded population*	4	Year	The approximate year when HIV was introduced into South Africa.
Sex acts per week	2	Sex acts per week	
Length of phase 1	12	Weeks	
PTSA during phase 1	0.032	Probability	
Length of phase 2	384	Weeks	Approximately 8 years.
PTSA during phase 2	0.0035	Probability	
Length of phase 3	Infinity	Weeks	Agents remain in this phase until death.
PTSA during phase 3	0.0152	Probability	
CD4 count at infection	Norm(1000, 250)	Distribution	Normal distribution defined by a mean and standard deviation.
CD4 count at death	Norm(75,25)	Distribution	

Table 3 continued.

Condom Use			
Start year*	10	Year	
End year*	14	Year	
Maximum coverage*	35	Percentage	
Time Operator			
Fertility Rate	Empirical		Based on ASSA2008 model
Non-AIDS mortality	Empirical		Based on ASSA2008 model
AIDS-mortality	Weibull(1.2,scale)	Distribution	Scale is a function of age in years: 13+((15-infected.age)/10)
ART Treatment			
Start year	17	Year	Simulation time equivalent to 2002
End Year	25	Year	Time when the number of ART treatment slots stopped growing.
Maximum Coverage	50	Slots	One-third of HIV-positive individuals were on ART in 2012, this number reflects an approximation of that.

2.4 Simulation Output

Here we show output produced by the implementation of our model. Where possible, data informs parameter values. Where not possible, we choose parameter values from within a reasonable range or fitted based on a manual comparison between simulation output and actual sexual network in townships near and around Cape Town. We note that our goal is not to provide a fully validated model that can correctly predict future trends, but to merely show that output from the model spans a feasible range that includes observed outcomes.

Figure 25 and Figure 26 show a comparison between actual and simulated demographics and HIV prevalence respectively. The more complex Time Operator is able to produce demographic trends seen in real life. Similarly, the more complex Infection Operator is able to produce prevalence levels like those seen in South Africa.

2.4.1 Non-Trivial Age-Mixing

In addition to general epidemic trends the model is able to simulate complex age-mixing patterns seen in real-life. A sexual network survey of townships around Cape Town found that prevalence of age-disparate relationships was high for young women and continued to be high as they grew older. Young men on the other hand had more age-disparate relationships as they grew older. Figure 27 shows a comparison of age-mixing patterns between our simulated population and the actual population.

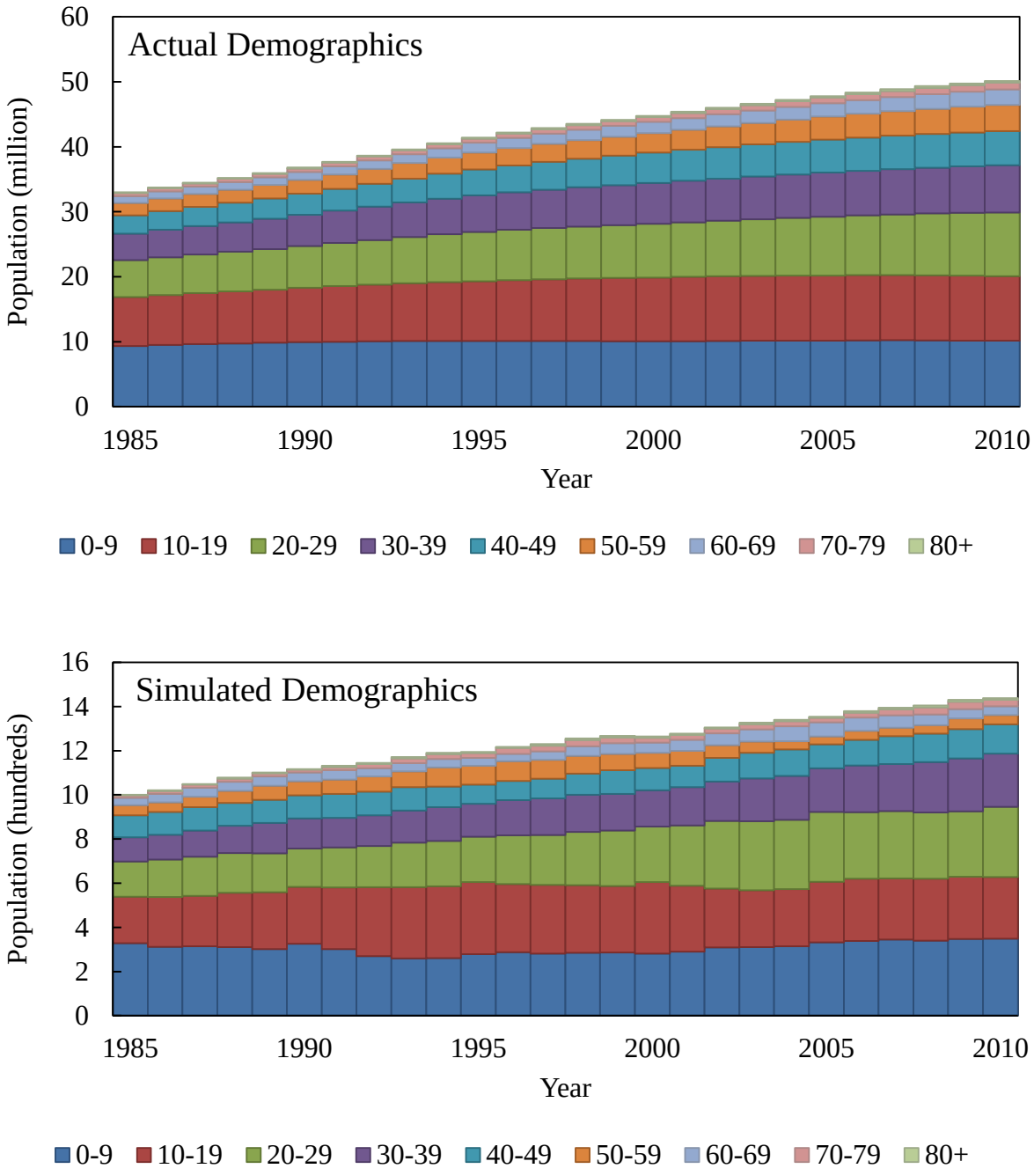

Figure 25: Demographic plots of the actual and simulated populations.

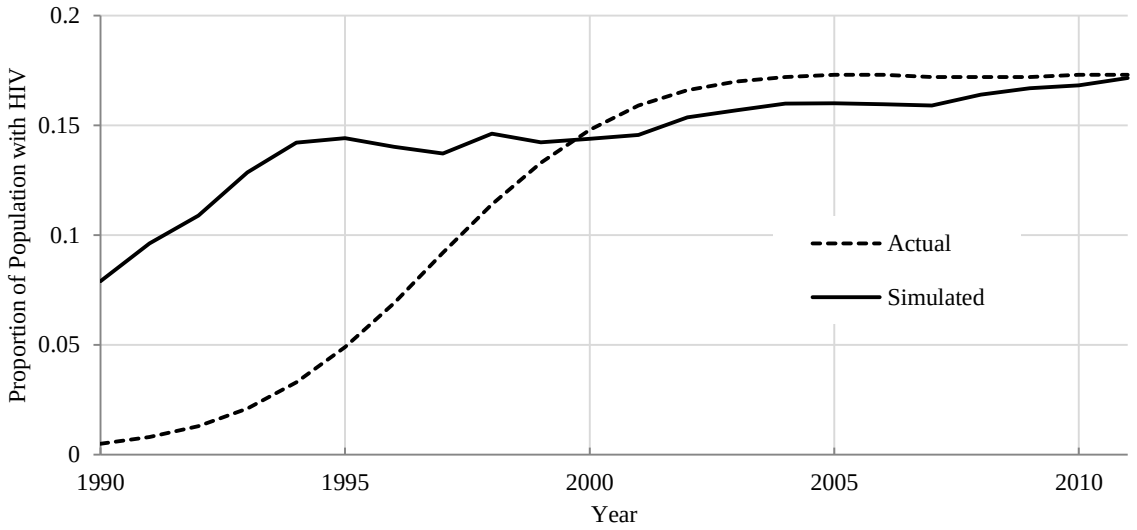

Figure 26: Comparison of simulated and actual HIV adult (15-49) prevalence in South Africa. The discrepancy implies that additional parameter inference is necessary.

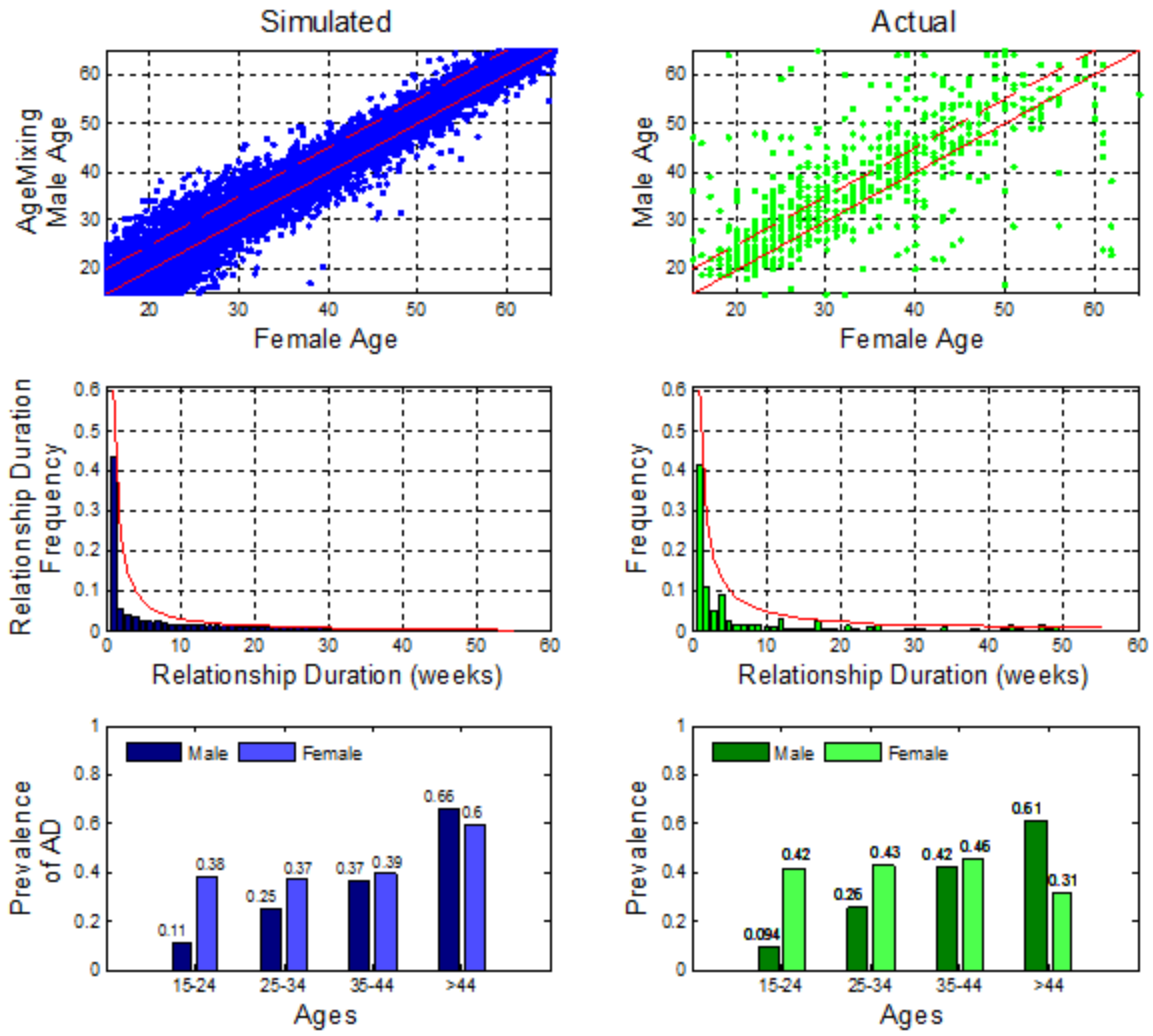

Figure 27: Comparison of the simulated sexual network and the actual sexual network seen from survey data collected in three disadvantaged communities near Cape Town. Our heterogeneous population allows us to simulate an age-mixing pattern in which proportion of age-disparate relationships is around 0.4 for women in all age categories, but increases gradually from 0.1 to 0.6 as men grow older. This is consistent with the sociological idea of "sugar daddies", in which older men provide economic support for younger women.

2.4.2 Relationship Durations

The structure of sexual networks is known to be significant factors in the epidemiology of STDs. Here we analyzed associations between standard indicators of sexual networks and the cumulative incidence of HIV. We explored the parameter space of *relationship formation* and

dissolution which generated a data set of sexual networks. We calculated concurrency, population partner turnover rate, median lifetime sexual partners, median age difference of relationships, and relationship duration from these networks and performed a regression analysis of these characteristics on the cumulative number of infections.

Our regression analysis suggests cumulative prevalence of concurrency, the median duration of relationships, and partner turnover rate are independent predictors of total number of infections, whereas median number of lifetime sex partners and median age difference of relationships are not. Additionally, the median duration of relationships seems to have a quadratic relationship with cumulative HIV incidence: if relationships in the system are short, HIV transmission is constrained by the limited number of sex acts; if relations are long, HIV is "trapped" in relationships.

This is an important distinction since the duration of a relationship (and hence the beginning of an individual's next relationship) will then determine the ability of the virus to diffuse through the network. The relationship between expected relationship duration in a simulation and the total number of infections is illustrated in Figure 28.

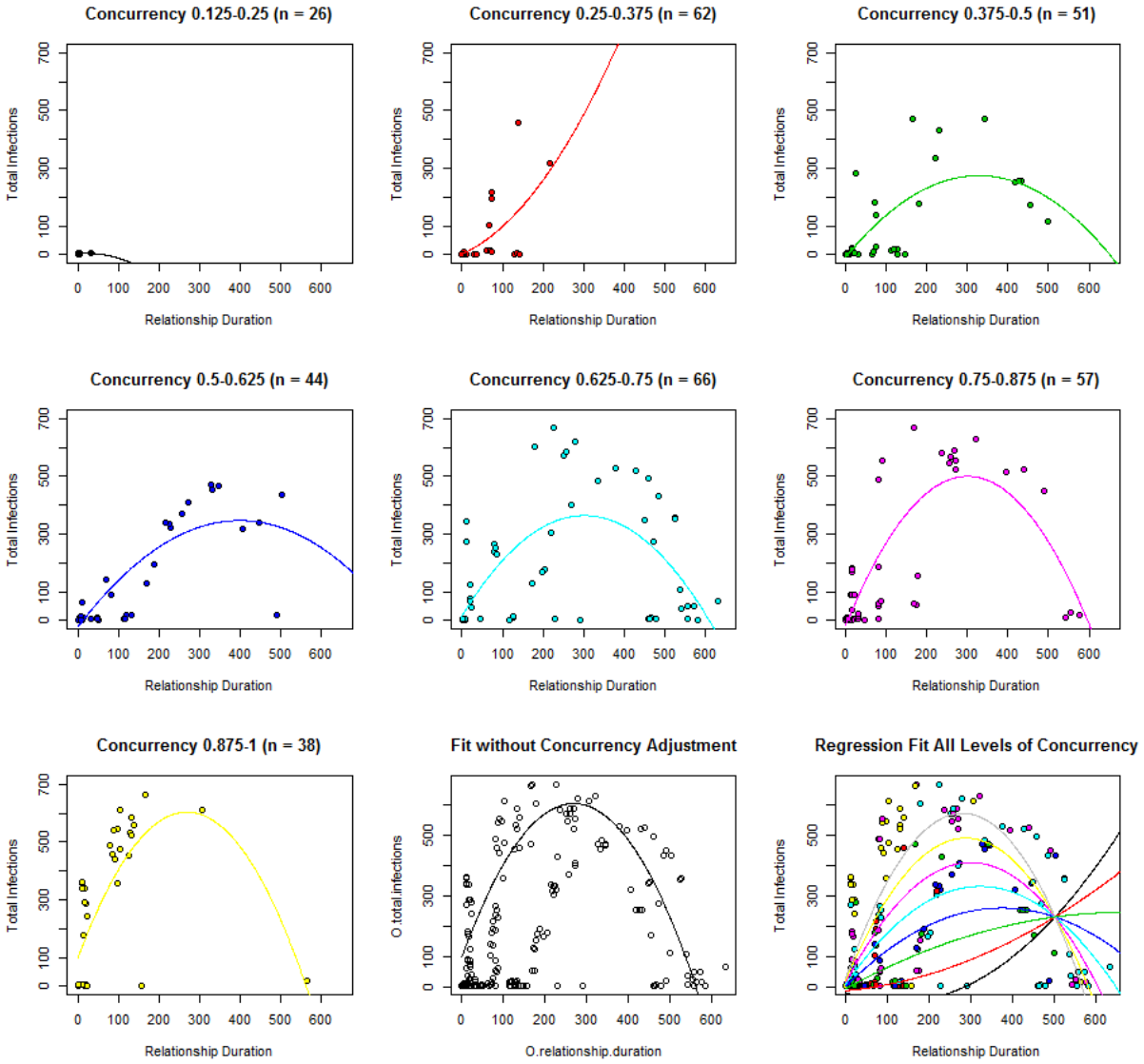

Figure 28: Simulation output showing the effect of relationship durations on total infections for different levels of network concurrency. Short relationships reduce the number of potential transmission events and thus reduce the total number of infections. Long relationships reduce the number of contacts an infected agent has and thus reduce the total number of infections as well. This parabolic relationship between mean relationship duration and mean total infections occurs independent of network concurrency (the proportion of agents with multiple partners).

In summary, we found associations between characteristics of the sexual network and total number of HIV infections. A behavioral change campaign then that effectively increases

average duration of relationships may see reductions or increases in HIV transmissions, and the relative impact of the intervention would depend on the network level of concurrency.

2.5 Discussion and Conclusion

In this chapter we've presented a mathematical formulation for simulating HIV. The formulation is flexible such that it is able to model sociological phenomena such as complex age-mixing patterns and behavioral changes in condom use. The model, parameterized correctly, can reproduce HIV prevalence trends and demographic shifts as seen in real world. While more work may be necessary to completely validate this particular implementation of the model, our goal was to instead show that the mathematical formulation was up to the challenge.

The naïve implementation with Java and MASON described here does not scale well to large populations. This makes the use of this implementation prohibitively expensive in terms of time for any meaningful modeling studies. Chapter 4 describes a shared-memory parallelization of the algorithm that greatly improves the performance of the model. Chapter 5 further describes a distributed-memory parallelization which further improves performance.

CHAPTER III A SIMULATION-BASED METHOD FOR EFFICIENT RESOURCE ALLOCATION OF COMBINATION HIV PREVENTION

3.1 Introduction

Over the past three decades there has been a wealth of operational research into effectively and efficiently combating human immunodeficiency virus (HIV). These interventions have had varying results. Condoms, for example, have been shown to decrease the probability of transmission per sexual act (PTSA) by 95%, but they tend to be used inconsistently. Male circumcision has been shown to reduce the PTSA by 50%, but provides consistent partial protection by design. Antiretroviral therapy (ART) is a medical treatment that slows the reproduction of HIV. ART has been associated with 96% reduction in PTSA, and has been shown to prolong the life of an infected individual. However, it is difficult to determine how to optimally distribute limited HIV prevention resources to prevention methods due to each method's different financial costs, levels of uptake and efficiency, and potential unintuitive interactions.

While the most intuitive solution is to spend at the point of maximal effect of each intervention, this is not possible in low-resource settings: in addition to the effectiveness of interventions, cost must be considered. In such settings the opportunity costs of allocating additional resources to one intervention over another might be great and so a greedy approach may not be appropriate. Differences in uptake, coverage, and consistency also support the notion that no single prevention method will be sufficient for disease eradication. Instead, a combination of interventions, known as combination prevention, is likely to be the most efficient use of public health funding [68].

Although combination prevention seems to be an obvious solution, the means by which we arrive at an optimal combination of preventions is not. High levels of complexity and heterogeneity in the process of HIV transmission (age-disparity within relationships, concurrent sexual relations, and infectivity of individuals based on stage of infection and treatment status) make traditional compartmental and differential equation (DE) models overly simplistic [69]. For this reason stochastic individual-based models that consider more explicitly the dynamic nature of a population's sexual network are better suited to the modeling of HIV combination prevention interventions. However, stochasticity such as non-deterministic transmission, formation, and dissolution events, make a closed-form solution to the problem of combination prevention difficult. Additionally, the problem of optimal resource allocation becomes intractable when considering diminishing returns of scale of spending, and subtle interactions between interventions.

In this chapter we present a method for finding a locally optimal combination of HIV prevention methods, and show that combination prevention performs better than any single intervention at reducing cumulative HIV incidence while working within a budget. Our research is novel in that we consider the objective of minimizing cumulative incidence in addition to respecting some given budget within an individual-based model. Our method uses artificial intelligence algorithms to find the best possible allocation of resources to prevention methods. Specifically we use simulated annealing, and a genetic crossover algorithm [44] to determine the best achievable intervention starting times and spending amount for condom distribution, male circumcision, and TasP programs.

In the next section, we discuss the agent-based model we used, a simplified version of the model presented in the previous chapter. We present the intervention methods and their

implementation, and the cost and effect of each within the model. In Section 3 we analyze the results of our optimization algorithms for combination prevention and in Section 4 we conclude with a discussion of the implications for policy and the areas of future work.

3.2 Methods

Our model is an event-driven, agent-based model that uses the modified next reaction method (mNRM) algorithm [70], a derivative of the Gillespie Stochastic Simulation algorithm [71]. The algorithm schedules events to occur relative to a unit-less hazard of each event. The time until an event is the time required for the cumulative hazard of the event to reach a random number between one and infinity. Thus, events with lower hazard are more likely to occur further in the future. We keep track of the time until every event, and perform each event in order.

The main purpose of the model is to simulate HIV transmission and the impact of HIV interventions. We conform to current recommendations for reporting of HIV modeling work [72], and follow the standard protocol suggested by Grimm et al. [73] to describe our model. This protocol, known as ODD (Overview, Design concepts, and Details), forms the structure of our methods description.

For purposes of reproducibility, we include a table of parameter, values, and justification in Table 4: Parameters used in the simulation. Parameter values are calibrated and validated in Section 3.2.8 Calibration and Validation. Values are loosely informed by behavioral and epidemiological surveillance from Cape Town, South Africa, but can be changed to explore other contexts.

Table 4: Parameters used in the simulation.

Parameter	Value	Justification
Population		
Population size	200 (100 male, 100 female)	This is the largest population we can run within a reasonable amount of time.
Initial infection	0.15	The approximate prevalence of HIV/AIDS in South Africa [3].
Age Distribution	70, 4	Scale, and shape parameters for Weibull [3].
Partnering Values	0.5, 0.5	α, β parameters for beta distribution. Set through experimental comparison to sexual behavior data [65].
Formation Event		
Baseline factor	2	See 3.2.8 Calibration and Validation
Current relations factor	0	
Mean age factor	-0.005	
Last change factor	0.014	
Age difference factor	0.1	
Mean age growth	0.4	
Mean age dispersion	0.154	
Preferred age difference factor	-0.18	
Dissolution Event		
Baseline factor	2.6	See 3.2.8 Calibration and Validation
Current relations factor	-0.23	
Mean age factor	-0.057	
Last change factor	-0.015	
Age difference factor	0.08	
Mean age growth	1.917	
Mean age dispersion	0.476	
Preferred age difference factor	-0.265	
HIV Transmission Event		
PTSA	0.032	[16]
Sex acts per week	2	[65]
Condom Distribution		
Risk reduction	0.8	This reduction incorporates inconsistent use [67]
Condom cost	$as = 2\exp\left(\frac{cd}{20} - 2\right)$	We experimented with different cost curves, but found little difference.

Table 4 continued.

Male Circumcision		
Risk reduction	0.5	Reduction for males only [74].
Circumcision cost	$cp = \frac{as}{50}$	[75]
Antiretroviral Therapy		
Risk reduction	0.96	[16]
ARV cost	$pa = \frac{as}{500}$	[76]

3.2.1 Purpose

The model was designed to explore the spread of HIV infections in complex and dynamic sexual networks. We built the model to address the question: which attributes contribute significantly to the diffusion of HIV, and what interventions are most effective in interrupting this diffusion?

3.2.2 Entities, State Variables, and Scales

The model considers two kinds of agents: males and females. Both kinds of agents have a notion of his or her:

1. Birth time (hence age)
2. Time since relationship change
3. Number of current relationships
4. Partnering value (described in 2.5 Initialization)
5. Time since infection
6. Exposure to a condom campaign
7. ART status (whether he or she has started taking ART)
8. Time of circumcision (males only)

3.2.3 Process Overview and Scheduling

Events occur one at a time according to the modified next reaction method. The events are:

1. Relation formation
2. Relationship dissolution
3. HIV transmission

For purposes of simplicity, mortality and replacement is not considered in this model.

As mentioned previously, events are scheduled to occur relative to the event specific hazard function (described in further detail in 3.2.6 Submodels). The order of events is significant since the firing of one event may enable or change another. The occurrence of some events affect the hazard of other events: the formation of a relationship between male *i* and female *j* may lower the hazard of formation of a relationship between male *i* and female *k* and thus the event will be scheduled to occur further into the future.

Additionally we have the notion of interventions which aim to interrupt disease spread by reducing the HIV transmission probability. Interventions (described in more details in 3.2.6 Submodels) are implemented at a specific starting time, and their coverage is relative to the amount of money spent.

3.2.4 Design Concepts

The model simulates the spread of HIV in complex sexual networks: events are specific to individuals (e.g. condom campaigns influence an individual's probability of HIV transmission, and relationships among individuals consider individual-level desirability of concurrency and

age-disparity), rather than to an aggregate sub-portion of the population. The individuality of events allows us to investigate the dynamics of an epidemic at a fine grain level.

3.2.5 Initialization

At initialization, 100 males and 100 females are introduced. The individuals are assigned ages from a Weibull distribution with scale 70 and shape 4 [77]. Each individual is assigned a random value from a beta distribution with $\alpha = 0.5, \beta = 0.5$. These values allow heterogeneity within our population so that some individuals with higher values are more likely to form relationships, and individuals with lower values are less likely to form relationships. Figure 29 shows the distribution of ages and partnering values at initialization.

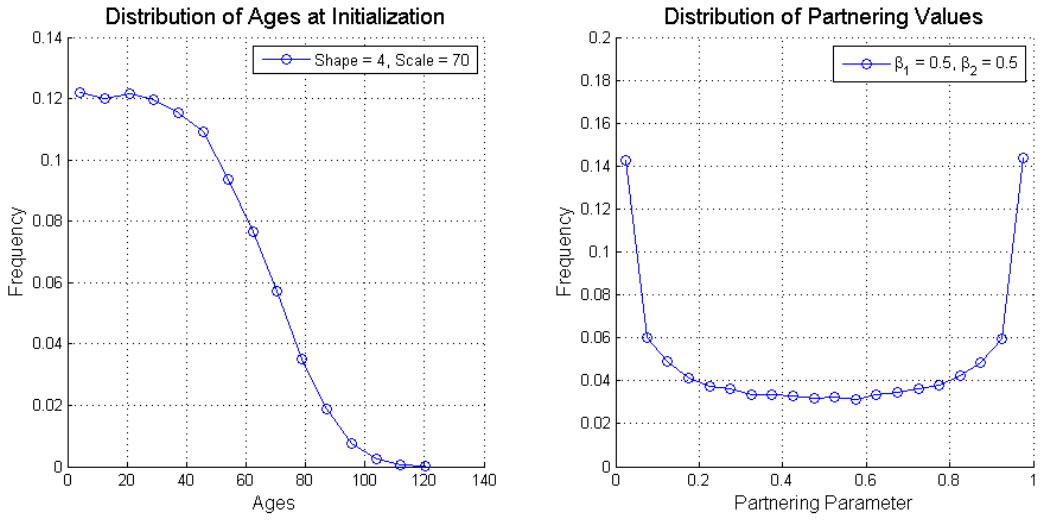

Figure 29: The distribution of ages (left) and partnering values (right) at initialization. Ages pulled from a Weibull distribution with scale 70, and shape 4, which is consistent with the age distribution of South Africa. Partnering values are pulled from a beta distribution with $\alpha = 0.5$ and $\beta = 0.5$, which produced a heterogeneous population similar to our observed sexual network (see Section 2.8 Calibration and Validation).

Relationships are allowed to form and dissolve until relationship dynamics are in a steady-state (two years). HIV is then introduced into the system through infecting 30 (15% of the population) randomly selected individuals [3].

3.2.6 Submodels

Each submodel represents one of the events or interventions that can occur. Each event has a specific hazard function that determines the time until it occurs.

3.2.6.1 Relationship formation

The event of relationship formation between male *i* and female *j* is based on the hazard function

$$h_{ij} = \exp(\alpha_1 u + \alpha_2 w + \alpha_3(x - 15) + \alpha_4 y + \frac{\alpha_5}{x \cdot \alpha_6 \cdot \alpha_6''} |m - f - x \cdot \alpha_6 \cdot \alpha_6'|).$$

Where *u* is the mean of the two individuals partnering values, *w* is the combined number of current relations, *x* is the mean age of the couple, *y* is the time since last change in relationship status (the last time either the male or female was an actor in a formation or dissolution event), *m* is male age and *f* is the female age. All others (i.e. all α_i) are constants with values set during calibration. For example, α_5 is the age difference factor, and α_6, α_6', α_6'' determine the preferred age difference. While HIV in men who have sex with men (MSM) is of concern, homosexual relationships are not considered in our model for simplicity. Relationships are only formed between individuals older than 15 years. Figure 30 shows a graphical representation of some elements of the hazard function.

This means that every relationship between every pair of individuals has a baseline of hazard of formation of $e^2 = 7.39$. This hazard is decreased multiplicatively based on the above

attributes. For example, consider a 22-year-old male (currently in one relationship, last ended a relationship 6 months [0.5 years] ago, and last started a relationship 1.2 months [0.1 years] ago) with a partnering value of 0.8, and a 19-year-old female (currently in no relationships, last ended a relationship 3 months [0.25 years] ago, and last started a relationship 2.4 months [0.2 years] ago) with a partnering value of 0.9. The hazard of a relationship forming is given by

$$\exp((2.0 \times 0.8 \times 0.9) + (0.1 \times 1) + \left(-0.004 \times (20.5 - 15)\right) + (0.01 \times 0.1) + \left(-0.1 \frac{|22 - 19 - (20.5 \times -0.181 \times 0.154)|}{20.5 \times -0.1812 \times 0.1544}\right) = 8.51$$

For random numbers 0.1, 1, 10, and 100 the time until relationship formation is 0.05, 0.43, 4.27, and 42.74 years respectively (random numbers are $(0,\infty)$ with expected value of 1). Note that even though the male is already in a relationship, there is a possibility of him forming another relationship with another female.

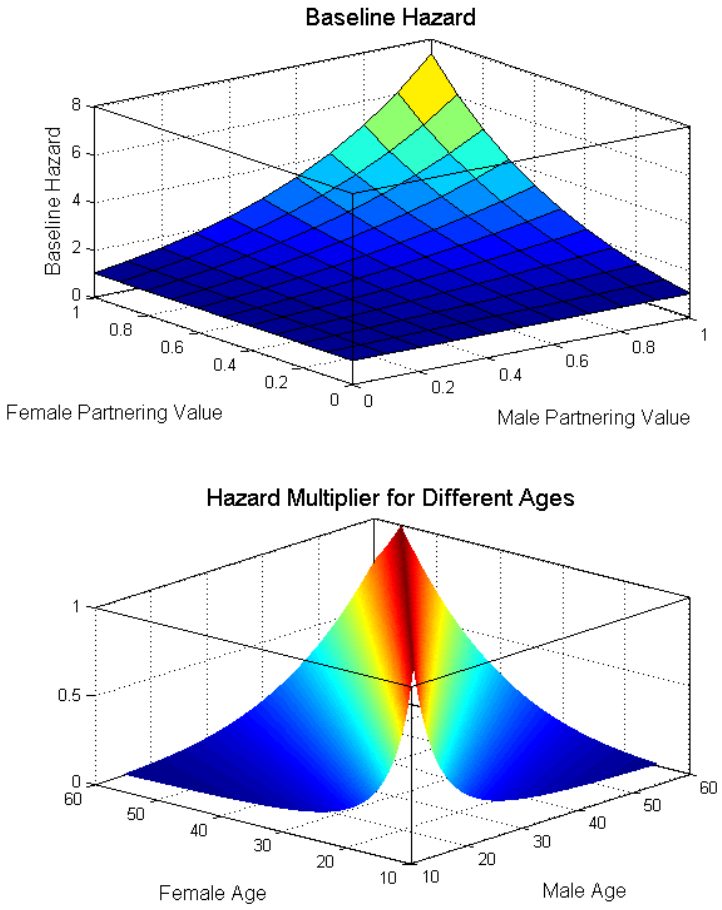

Figure 30: On the top, the baseline of a formation event is based on $\boldsymbol{\alpha_1}$ and the product of the two individuals partnering value. Individuals' with higher partnering values will have a higher baseline for forming a relationship. On the bottom, the hazard is decreased multiplicatively as two individuals' age difference moves further from the preferred age difference.

3.2.6.2 Relationship dissolution

Once a formation event occurs, the event of dissolving this relationship (breaking up) becomes possible. The hazard of a relationship between male *i* and female *j* dissolving is based on a hazard function of the same form as the formation hazard function, but with different constants (see Table 4). Our sexual network then emerges from a series of formation and dissolution events.

3.2.6.3 HIV transmission

Infection can occur in serodiscordant relations, i.e., relations in which one partner is infected and the other is not. The event is scheduled to occur relative to the hazard $-\log((1-PTSA)^S)$. Where S is the number of sexual acts per week, and $PTSA$ is the probability of transmission per sexual act.

3.2.6.4 Condom distribution

Unlike the random events, interventions are scheduled to occur at specific times (e.g. five years into the simulation) and is therefore independent of a hazard. We consider different targeting schemes for condom distribution which lead to different individuals possessing condoms. The intervention targeting strategies we considered were

1. Individuals currently in multiple concurrent relationships
2. HIV positive individuals
3. Younger individuals (males and females between 15 and 25)
4. Individuals who have a high perceived risk (their partners are in more than one sexual relationship)
5. Random individuals (no targeting).

At the start time of an intervention, we find targeted individuals and mark them as influenced by the condom distribution campaign. One influenced individual consumes one distributed condom. Note that a "distributed condom" does not equate to using a single condom in a single sex act, but is instead analogous to a single individual being supplied with many condoms for one year.

We make the assumption that we find targeted individuals with 0.8 probability (we account for the fact that finding specific individuals is difficult). Individuals influenced by a condom distribution campaign have their infectivity reduced by 80% [67]. While condoms are known to decrease infectivity by a significant amount [78], this lower number reflects the possible effects of inconsistent use.

We assumed a decreasing return to scale between individuals influenced and amount spent: $as = 2\exp\left(\frac{ii}{20} - 2\right)$ where *as* is the amount spent in thousands of USD and *ii* is the number of individual influenced by the campaign. This means that in order for a campaign to influence 60 individuals it would need to spend $42,000 per year.

3.2.6.5 Male circumcision

Male circumcision (MC) is similar to condom use in that it reduces the PTSA, but has the added advantage of being used consistently [76]. While condoms reduce PTSA by nearly 100%, male circumcision can reduce PTSA by about half as compared to without circumcision [74]. We implemented a single MC campaign which does not target any group; at the start time of the intervention random males were chosen to be circumcised. PTSA *to* males influenced by the MC campaign is reduced by 50%. Unfortunately, circumcision does not seem to hold any benefit to females other than that their partners are less likely to become infected [79].

We assumed a linear relationship between circumcisions performed and amount spent: $cp = \frac{as}{50}$, where *as* is the amount spent in thousands of USD and *cp* is the circumcisions performed. This comes from the fact that a single circumcision costs about $50 to perform [75]. This means that in order for a campaign to reach 60 males it would need spend $3,000.

3.2.6.6 Antiretroviral treatment

TasP as an intervention method not only reduces HIV related deaths, but also has the ability to reduce the infectivity of an individual by means of decreasing his or her viral load [41]. Therefore, treating a significant portion of the population with ARV can decrease HIV incidence. Our implementation of TasP finds HIV infected individuals with probability 0.8 and reduces their infectivity by 96%.

We assumed a linear relationship between patients on ARV and amount spend: $pa = \frac{as}{500}$, where *as* is the amount spent in thousands of dollars and *pa* is the number of person years of ARV supplied. This comes from the fact that ARVs cost about $500 per person per year [76].

3.2.7 Search Heuristics

The optimization problem we aimed to solve had an objective of minimizing cumulative incidence with the constraint that the amount spent could not exceed the prescribed budget of $1,000,000 / year (about $150 per person per year). Therefore a solution is a set of starting times and amount of money to spend on each intervention. The quality of a solution depends on cumulative incidence averaged over 10 runs. The cost depends on the two parameters “starting time” and “spending amount”. The cost of a solution is determined by the number of years each campaign is implemented (calculated as the number of years between the start of the campaign and the end of the simulation) multiplied by the number of condoms distributed, or individuals on ARVs. Males circumcision does not incur a yearly cost – cost is calculated just once. A feasible solution spends less than the budget. The optimal solution has the minimal cumulative incidence possible.

The simulated annealing algorithm is a walk through the parameter space. Our implementation always accepts improving moves, and accepts non-improving moves with probability $\frac{\exp(e-e_{new})}{T}$, where e is the quality of the current solution, e_{new} is the quality of the new solution, and T is the temperature of the system. Temperature decreased relative to the current time step k at a rate of $T(k) = 0.96^k$. Maximum number of steps was 100.

The genetic algorithm produces 10 random solutions, assesses their quality, then produces a new set of 10 solutions by performing a crossover of the best 5 solutions. This procedure is repeated for 20 generations. These values were chosen through experimentation to minimize run time and maximize quality. Crossing over two solutions means taking the first *p* values of the first solution, and the last *n-p* values of the second solution, where *n* is the total number of start times and spending amounts, and *p* is a uniform random [0, n].

We first applied the search heuristics to find the best combination of condom distributions, and then applied them to find the best combination of random condom distribution, male circumcision campaign, and a roll out of TasP.

3.2.8 Calibration and Validation

Inference of appropriate parameter values, or calibration, was done in three steps: (1) the simulation was run for a specific set of formation and dissolution parameters for 50 years (to ensure relationship equilibrium and to have a large number of individuals who became sexually mature within the simulation). (2) From the resulting sexual network we calculated the distribution of partner ages, age differences within relationships, total number of lifetime sexual partners, level of concurrency in the sexual network, and the duration of relationships of *males* in the simulations. (3) We then compared these summary statistics to the responses from *males* that

took part in the Cape Town Sexual Network survey [76]. We compare to only male data because of possible gender-related sampling bias. The study took place from July 2011 to February 2012 and was located in three disadvantaged communities near Cape Town, South Africa. Table 5 contains the actual values from the survey compared to simulated values from our model.

Table 5: A comparison of summary statistics of data and a simulated network.

Statistic	Actual Data	Simulated Data
Age of partner median (IQR)		
Median	26	27.8
lower quartile	21	21.3
upper quartile	39	34.9
Age of partner breakdown (%)		
<=24 years old	35.6	34.7
25-34 years old	24.9	38.7
35-44 years old	14.3	12.7
45+ years old	25.2	12.1
Age difference median (IQR)		
Median	3	4.2
lower quartile	0.5	1.7
upper quartile	6	8.6
Age-disparate (%)		
non age-disparate	65	49.5
age-disparate: 5-9 years	17.8	28.4
age-disparate: 10+ years	17.2	22.1
Total lifetime sex partners (%)		
1	8.7	14.5
2-5	42	62.4
6-14	22.1	22.4
15+	27.2	0.7
Concurrent relationship in past year (%)		
Yes	41.5	12.4
No	58.5	87.6

Table 5 continued.

Duration of relationships		
Median	17	27.1
lower quartile	1	8.6
upper quartile	43	95.7
Duration of relationships breakdown		
1 week	26.6	8.0
2-39 weeks	48.5	52.1
40+ weeks	25.1	39.9

3.3 Results and Discussion

3.3.1 Condom Distributions

Independent runs of the condom distribution strategies (Figure 31) show that all strategies have an effect on reducing cumulative incidence. The most effective strategy seems to be targeting HIV-positive individuals and individuals in concurrent relationships (high risk). Targeting the younger population seems to have less effect, likely because the number of targeted individuals is low. This results in unused condoms and higher cumulative incidence. We hypothesized that Specific age group targeting would have an effect through protecting a large cohort and averting infections to the younger population (<15 years) reaching relationship formation age. This did not seem to play out in the simulations however.

The fact that some condoms go unused implies that a better scheme would be a combination of condom targeting strategies in which each intervention spends at their maximal level of effectiveness and allocates the saved funds to other strategies. That is to say that it may be worthwhile to delay the start time of a certain intervention (and consequentially save some of the budget) since these individuals may not be infected for many years into the future. For

example, it may be practical to delay the start of an intervention that targets individuals with a high perceived risk because they are unable to become infected until their risky partner becomes infected. This in turn reduces cost and allows more of the budget to be allocated to another condom distribution such as one that targets HIV positive individuals.

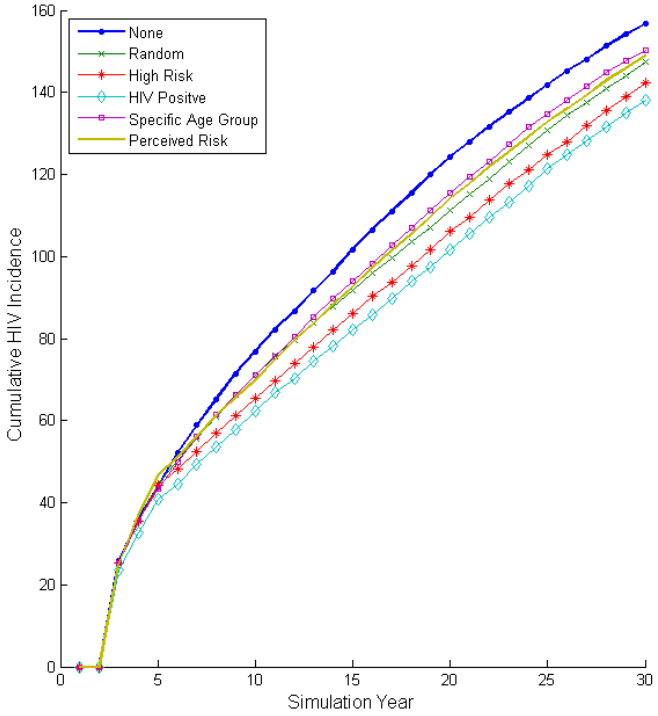

Figure 31: The cumulative incidence for the five described targeting strategies for condom distribution and the "no interventions" strategy averaged over 50 runs. Thirty individuals were infected with HIV from simulation year 2.1 to 2.9. Interventions were set to begin at year five, and attempted to distributed 54 condoms. All interventions reduce the cumulative incidence relative to the "no interventions" scenario, although targeting HIV-positives and those with high risk seem to be the most effective. The other interventions reduce cumulative incidence from doing nothing, but not much difference can be seen between random, high perceived risk, or age-specific targeting. However, with the exception of random targeting, all of the interventions are wasteful as none use all the allocated condoms. The cost was the same for all interventions at $996,000 which is within our $1,000,000 budget.

The optimization algorithms found a solution to the combination condom prevention problem for different prevention start times and amount spent as seen in Table 6. The total cost is $987,385 (about the same as the independent runs of condom interventions), but the cumulative incidence of combination prevention (Figure 32) is lower than targeting high risk individuals and much lower than no interventions.

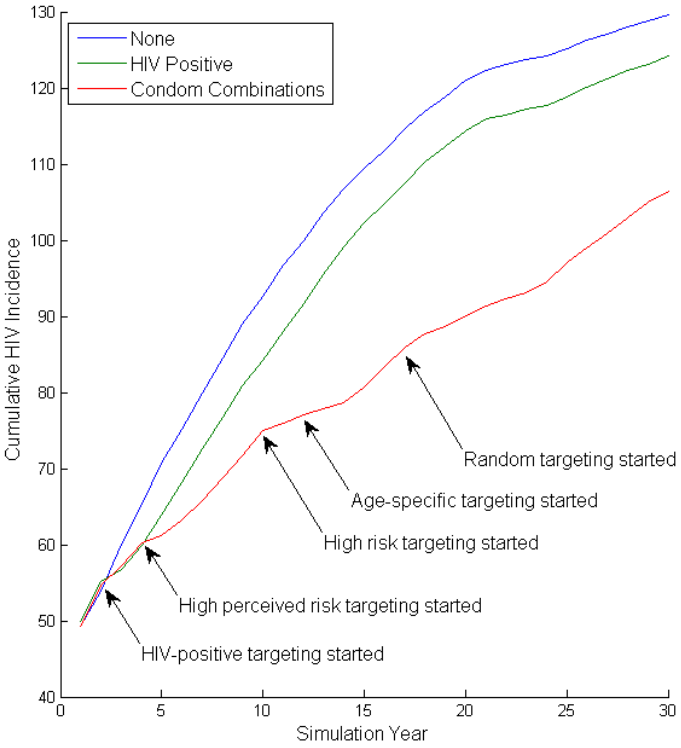

Figure 32: The cumulative incidence for no interventions, for targeting HIV-positive individuals, and for a combination of condom targeting strategies averaged over 50 runs. Forty individuals were infected with HIV from simulation year 0.3 to 1. Interventions were allowed to start at time 2. The figure shows the overall trend that condom combination prevention has a lower cumulative incidence than high risk targeting, which has a lower cumulative incidence than no intervention at all. The reason for this is that the condom combination prevention accounts for diminishing return and allows each intervention to be funded at the best level and is able to redirect unused resources to other interventions.

Table 6: The starting time and number of condoms to distribute for each intervention for our combination condom prevention strategy. The cost for this combination of condom distributions interventions is $987,385.

Intervention	Start Time	Condoms
Random	17	42
High Risk	10	10
HIV Positive	2	40
Age Specific	12	1
High Perceived Risk	4	42

The combination prevention has a lower cumulative incidence because it is able to fund each intervention at its locally optimal cost-effect point and therefore distribute more condoms to more people with less waste. Additionally, the susceptible or infected population is not a single group but a combination of groups. Therefore the intervention that targets many different groups in combination is likely to be the most effective. This is perhaps why the random strategy performed well in independent runs: it was able to reach many different groups. However this intervention is not implemented until later in the combination prevention solution. This is likely because combination prevention allows us to target these groups specifically through the other interventions, diminishing the necessity of a random distribution campaign.

3.3.2 Combination Prevention

Figure 33 shows the cumulative incidence under scenarios for male circumcision, TasP, the random targeting condom distributions, and the combination of prevention strategies. Table 7 shows the values of this combination prevention solution.

The solution to combination prevention performs many male circumcisions, likely because each is relatively cheap. It also spends heavily on TasP, which is comparatively expensive, but also has the most dramatic effect on HIV cumulative incidence within our model. However, combination prevention achieves the best reduction in cumulative HIV incidence.

Table 7: The starting time and spend variable (condoms distributed, circumcisions performed, or patients on ARV respectively) on each intervention for our combination prevention strategy. All preventions start early, but have different levels of implementations as indicated by the spend variable.

Intervention	Start Time	Spend Variable
Condom Distribution	5	28
Male Circumcision	5	100
TasP	5	64

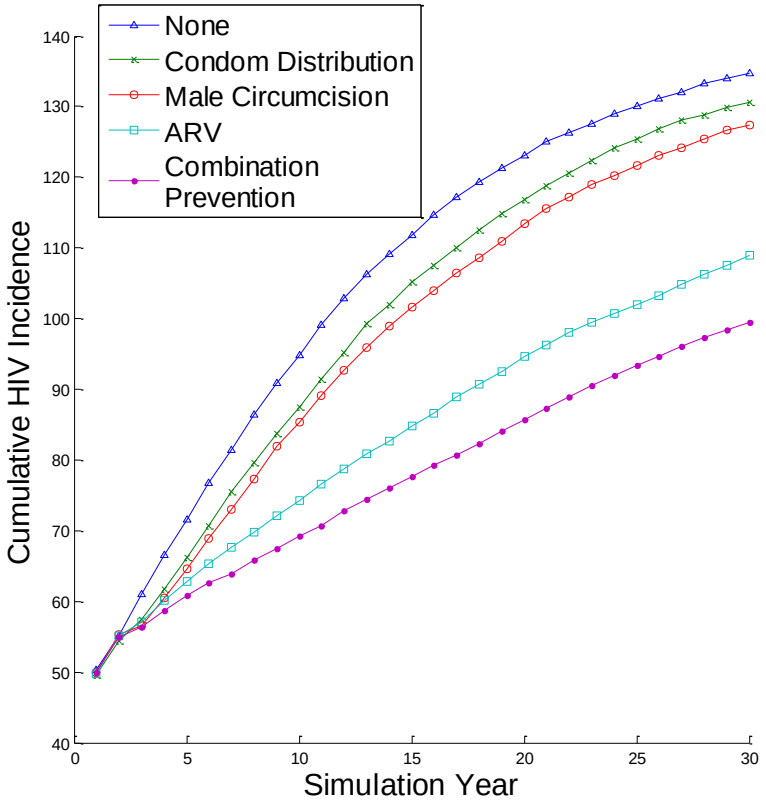

Figure 33: The cumulative incidence for no interventions, random targeting condom distribution intervention, male circumcision, TasP, and combination prevention. Our combination spends heavily on TasP, but also relies on condom distributions and male circumcision to achieve an even lower cumulative incidence. This shows that funds may be better allocated to a combination of prevention methods instead of any single interventions. The total cost was $995,870 for the combination prevention scheme.

3.4 Conclusions and future work

No current intervention is likely to be a silver bullet to the HIV epidemic, and none is likely to be found. Therefore a combination of prevention methods is likely the most effective solution. While the most intuitive strategy is to spend maximally on each intervention, this is not always possible due to limited resources. In this chapter, we have shown that combination of prevention can be more effective at minimizing cumulative HIV incidence than any single strategy, and described a method for finding the best possible combination prevention.

Other metrics of the quality of interventions should also be considered: cumulative incidence only tells one story. Additional consideration should be given to more time sensitive outcomes like the number of AIDS orphans averted or number of orphan years averted. These other metrics may provide greater support for the life-prolonging ART treatment intervention and yield a different combination of prevention interventions. Future work will consider multi-component objectives.

Due to the precise nature of the algorithm large populations can take a significant amount of time, even when run on a cluster. In the next chapter we present our algorithm for parallelizing the mathematical formulation which allows the modeling of much larger populations.

CHAPTER IV A PARALELLIZED ALGORITHM FOR SIMULATING DYNAMIC SEXUAL NETWORKS

4.1 Introduction

Epidemiologists are increasingly using agent-based models to simulate complex and heterogeneous human behavior and its effect on the diffusion of sexually transmitted diseases [80]. This is due in part to the fact that agent-based models of a sexually transmitted disease (STD) epidemic can capture more fine-grained complexities that might otherwise be understated in statistical or compartmental models [81, 82]. One of the challenges however is the large computational cost: obtaining a distribution of model outcomes requires many simulation runs, and obtaining robust results that are free of small population effects requires that each run uses a sufficiently large population.

In this chapter we present a parallel algorithm and implementation for simulating large-scale dynamic sexual networks and STD transmission through them. First, we describe the algorithm and how it works. Second, we present the algorithm's implementation in Python, and describe our method for calibrating parameter values. Next, we present a parameter exploration and show empirically that a model with higher levels of population heterogeneity requires a larger number of agents to obtain robust results. We conclude with a performance analysis to show that our model indeed scales well to large population sizes, enabling it to model highly heterogeneous populations.

We use disease parameters informed by the HIV epidemic in Southern Africa, though our goal is not to create a fully validated model of HIV transmission to be used for predicting future epidemic trends. Rather our algorithm is meant to simulate a generic STD in an agent-based

environment that is flexible to a variety of epidemic scenarios and scales well to large population sizes. The model was developed according to principles of good epidemiology modeling [72]. The implementation is open-source and available through our GitHub repository (github.com/seanluciotolentino/SimpactPurple).

4.2 Simulating Sexual Networks

4.2.1 Process Overview

The central components of an STD simulation are (i) relationship formation and dissolution, (ii) infection propagation, and (iii) demographic continuity (i.e., birth and death). In our model, each of these components is performed by an operator: the *relationship operator*, through a process described below, forms and dissolves relationships between agents; the *infection operator* considers all infected agents and probabilistically infects their partners in the network; and the *time operator* ensures demographic continuity by removing older agents from the system and inserting younger agents as needed. Each operator is applied in turn once per time step for as long as the simulation is allowed to proceed.

The simulation is initialized by specifying a population size (the initial number of agents that are created). When an agent is created, it is assigned a sex (Bernoulli (0.5) – the approximate sex ratio in South Africa [2]), an age (Uniform (15, 65) – used in CAN model [56]), a *desired number of partners* (DNP, power distribution – used in CAN model [56] with parameter values set through calibration; see Implementation and Calibration), and a sexual behavior index (Uniform(1,5) – chosen for simplicity). The sexual behavior index models an agent's preference for partners of a similar type. We note that due to a paucity of data, the sexual behavior index is used to create additional heterogeneity in a few circumstances and is only included in models where indicated – in all other models, agents form relationships solely on other characteristics. All agents in the base model are heterosexual. An agent is

considered to be looking for a partner if his or her number of partners (initially zero) is less than his or her DNP. In this way, the DNP can be thought of as the agent's target degree in the sexual network at any given point in time.

4.2.2 Probability of a Relationship

The probability of two agents i and j forming a relationship is based on the function $P_{ij} = \exp\left(M_A \times \left(\left|AD - \left(PAD \times PAD_{growth} \times MA\right)\right|\right)\right) \exp(M_{SB} \times SBD)$ where $M_A < 0$ is a probability scaling factor for the significance of age on relationship formation. AD is the age difference of the couple from the male perspective (a value of -5 means that the female is five years older than the male), PAD is the preferred age difference, PAD_{growth} is the preferred age difference growth, and MA is the mean age. $M_{SB} < 0$ is a probability scaling factor for the significant of sexual behavior, and SBD is the difference in sexual behavior indices of the couple. Note that *AD, MA, SBD* are calculated based on the candidate couple, while M_A, *PAD*, PAD_{growth}, M_{SB} and are parameters of the simulation.

A probability function of this form means that two agents with an age difference near the preferred age difference, and similar sexual behavior indices are more likely to form a relationship. The values for M_A and M_{SB} scale the probability of a relationship forming such that age difference or sexual behavior index is more or less significant. The PAD_{growth} parameter allows the preferred age difference to increase as men grow older. A visualization of the probability function is provided in Figure 34.

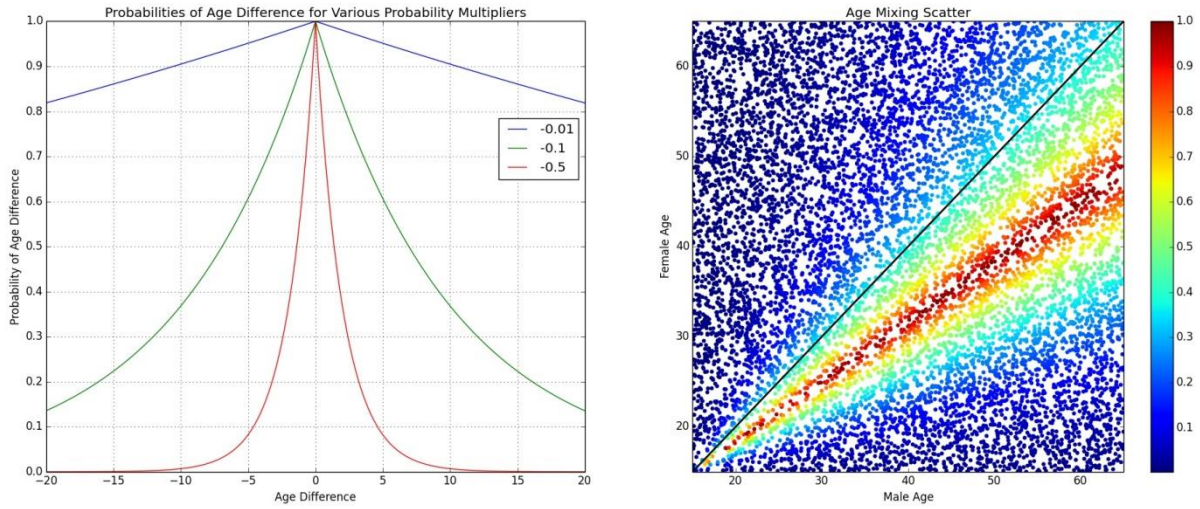

Figure 34: Left: the relative probability of relationships formation for different PM values and a preferred age difference of 0. Right: the relative probability of relationship formation for different combinations of male and female ages. Here M_A is -0.1, M_{SB} is 0, PAD is -0.2, and PAD_{growth} is 1.5.

While this is a simple probability function that accounts for differences in age and the mean age of the couple, the probability function can easily be altered to incorporate many more characteristics including sexual orientation, race, socio-economic status, and geographic location.

4.2.3 Relationship Operator

The naïve method for forming relationships is to consider each agent with fewer partners than his or her DNP, and then iterate through all other agents to find a suitable partner. This solution has the advantage of simplicity, but does not scale well due to its intrinsically quadratic run time. To create a more scalable solution, we limit the number of potential partners considered for each agent at any given time step, while allowing them to form relationships with agents across the age spectrum.

The model keeps track of agents looking for a relationship with *queues*, a list of objects that is ordered by some criteria. We create a *grid of queues* where the dimensions of the grid reflect the attributes that we want to use to inform relationship formation. The base model creates a 2x10 grid of queues (Figure 35) based on 2 sexes and 10 age categories (ages 15 to 65, grouped by 5). At initialization, agents are created based on age and sex distributions. Their age and sex in turn determine their respective queue, which represents their birth and sex cohort. The agents are placed in their queues and ordered based on the time since they were last allowed to form a relationship (which is initially the same for every agent).

The relationship formation procedure then takes place in two phases. In the first phase, a limited number of agents seeking new relationships are recruited from their queues (with agents who have been waiting the longest being recruited first) and used to populate another queue, called the *main queue*. The agents placed into the main queue are ordered based on their respective age and gender cohort. In the second phase, the relationship operator considers each agent in the main queue and attempts to match him or her to agents that are still in the queues.

The matching of an agent in the main queue occurs in three steps. First, the relationship operator takes the top agent from the main queue, referred to as the *suitor*, and sends a message to each queue that the suitor is looking for a match (Figure 36). Second, each queue orders their agents (in parallel) based on each agent's affinity to the suitor (Figure 37), a binary outcome of a random draw from that agent's probability function relative to the suitor.

In the third step, each queue responds with a match; the first agent in their ordering (Figure 38) or none, if no agent in the queue was willing to form a relationship with the suitor. From the returned possible matches, the suitor chooses a new partner based on his or her

probability function value towards each of them. The duration of the relationship is given by $d_{ij} = MA \times D_{scale} \cdot EXP(D_{shape})$ where MA is again the mean age of the couple, D_{scale} is a constant scaling factor, and $EXP(D_{shape})$ is a random value from an exponential distribution [57]. Each agent that is forming the relationship is removed from his or her respective queue if he or she is no longer looking for partners (i.e. if their number of partners is equal to their desired number of partners). This way they won't be recruited to be suitors and won't be returned as matches in future time steps.

The matching procedure for the relationship operator then proceeds by iterating through the main queue and making relationships for each suitor. Since suitors are ordered in the main queue based on the queue from which they came, the next suitor is often similar, in terms of age and sex, to the previous suitor. Consequently, queues do not need to reorder their agents since the probability relative to the agent of this particular age and sex has already been calculated. Queues can then return matches in constant time. The fact that previous probability calculations are recycled enables significant speed up as shown in Figure 41.

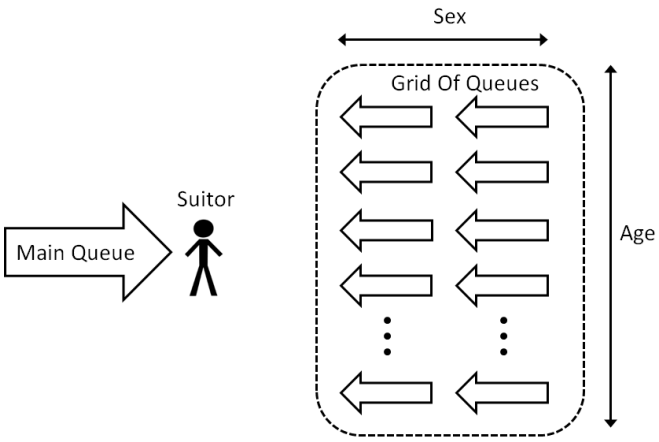

Figure 35: The simulation is made up of a grid of queues, which holds all the agents, and a main queue that holds agents waiting to be matched. We refer to the agent at the head of the main queue as the suitor.

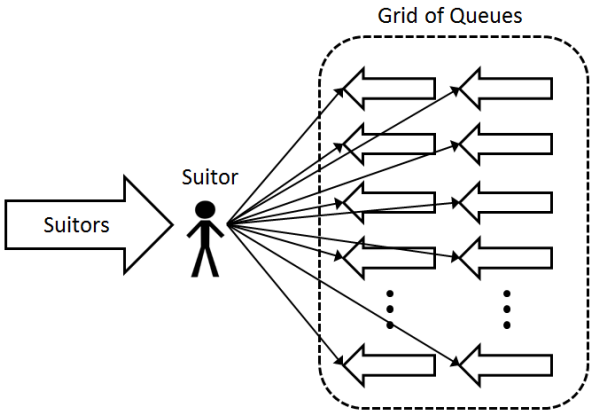

Figure 36: A message is sent to each queue, asking for a match for a particular suitor. Note that while the agents in our base model implementation are strictly heterosexual, the model supports homosexual matching.

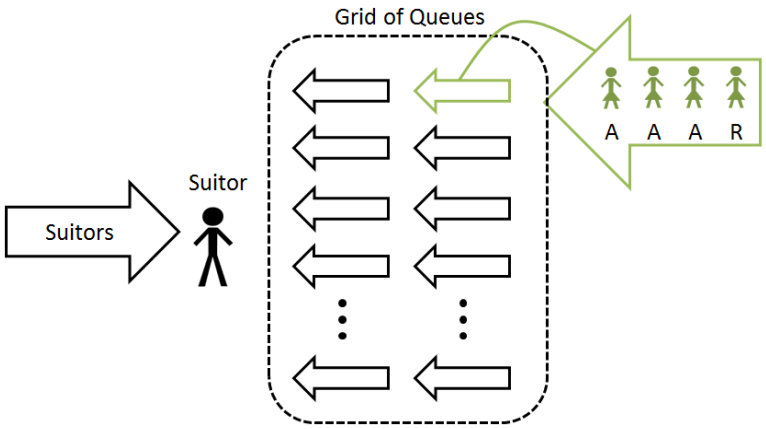

Figure 37: Each queue considers a suitor in parallel by ordering their agents relative to each agent's acceptance (A) or rejection (R) of a relationship with the suitor. The acceptance is randomly determined relative to an agent's probability function.

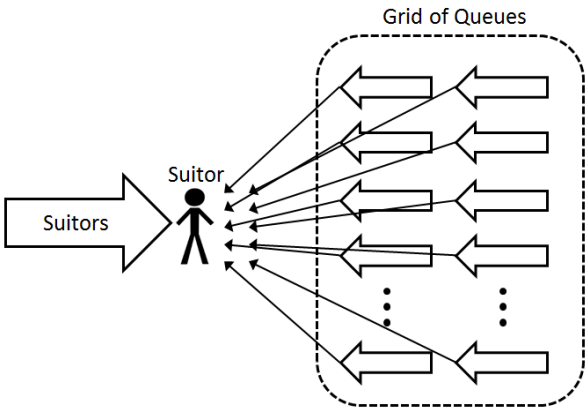

Figure 38: Queues return a possible match for the suitor. The suitor chooses a new partner from these matches randomly based on the probability function.

To summarize, the relationship operator first recruits new suitors into the main queue and second matches suitors to candidates returned from the queues. This recruit-match strategy leads to significant speed improvements by reordering queues in parallel, and mandating that reordering only occurs when the next suitor in the main queue is from a different age-gender cohort than the previous suitor.

4.2.4 Infection and Time Operator

After the initial population has been created, the infection operator seeds the STD by infecting a few agents chosen at random. At each time step the infection operator iterates through the list of infected agents and propagates infection to uninfected partners (Bernoulli(0.01) [83]).

The time operator ensures the passage of time by ending relationships and removing and replacing agents that have reached the maximum age (65 years old in our model). When a relationship ends, the two agents return to their respective queues, which were assigned based on their age and sex at initialization. The oldest queue is queried for agents who have exceeded the maximum age: these are agents are removed and then replaced by a 15-year-old agent of random sex and random DNP. Thus the population size remains constant over time, and the demographics remain approximately similar throughout the simulation. Relationships for agents being removed are ended, and the surviving partners are allowed to form new relationships. As simulation time progresses, the relevant time window also changes and the time operator creates new queues as needed. (A simulation with 5 year age bins will need to create a new queue every 5 simulation years to hold the new 15-year-old agents).

4.3 Implementation and Calibration

The model described here was implemented in Python (version 2.7.5) using multiprocessing, numpy [84], networkx [85], and matplotlib [86] modules.

Our goal is to simulate a large-scale network that approximates the behavior of a real-world dynamic sexual network. Here, we attempt to infer reasonable parameters values (enumerated in Table 8) so that the output of our simulator, once all parameters are established, corresponds to values found in existing sexual behavior surveys. Unfortunately, collecting comprehensive and reliable data about these networks is both difficult and costly, and thus

generally data are quite sparse. So parameter values are informed by literature where possible. Where it is not possible, appropriate values are inferred empirically using approximate Bayesian computation (ABC) [87, 88]. The simulation parameters for the number of agents recruited into the main queue at each time step were set manually to achieve relationship equilibrium quickly.

Here we briefly describe the ABC method used to infer parameter values. Given a set of parameters to be inferred $\theta = \{\theta_i \mid i = 1,2, \dots, n\}$, prior distributions for those parameters $\pi = \{\pi_i \mid \pi_i \text{ is the distribution of } \theta_i\}$, and a vector of existing data (summary statistics) to compare against, the ABC algorithm works by repeatedly performing 3 steps:

1. Create parameter set θ^* by sampling from each parameter's respective prior distribution.
2. Run the simulation with the sampled parameters.
3. Compare summary statistics from the simulation to those derived from existing data. If simulation output is within a pre-specified distance bound, accept the parameter set, otherwise reject it.

After many samples we have a large set of accepted parameter sets to construct the parameters' posterior distributions and their resulting sexual networks.

Table 8: The parameter values used in the simulation. Parameters inferred using the ABC method are represented by θ_i. All other parameters are taken from literature.

Parameter	Value	Description	Justification
Probability scaling factor for age difference (M_A)	θ_1	Coefficient in the probability function that determines the baseline probability of a relationship forming for deviation away from the preferred age difference.	
Preferred age difference	θ_2	Coefficient in the probability function that determines the age difference for which the baseline probability of a relationship forming is highest.	
Preferred age difference growth	θ_3	Coefficient in the probability function that determines the amount that preferred age difference grows with mean age.	
DNP Distribution	$\theta_4 \cdot \text{Power}(\theta_5)$	The distribution of desired number partners; also known as the degree distribution.	Distribution used in the CAN model [56].
Duration Distribution	$\theta_6 \cdot \text{Exp}(\theta_7)$	When a relationship is formed, the duration of the relationship is pulled from this distribution.	Duration of relationships are approximately exponential [57].
Probability scaling factor for sexual behavior difference (M_{SB})	0.0	Coefficient in the probability function that determines the relative significant of the sexual behavior indices of agents.	Not used for calibration because not enough data is available (used for simulating increased heterogeneity in later sections).
Age Distribution	Uniform(15,65)	The distribution of ages when agents are initially created.	Arbitrary; A uniform distribution was chosen for simplicity.
Sex Distribution	Bernoulli(0.5)	The distribution of sex when agents are initially created.	The approximate sex ratio in South Africa [2].
Initial recruitment rate	0.02	The initial proportion of agents recruited from queues to populate the main queue.	Set experimentally to allow the simulation to quickly reach equilibrium.
Warm-up period	20	The number of weeks that the simulation uses the value of *initial recruitment rate*.	Set experimentally to allow enough time for the simulation to reach equilibrium.
Recruitment rate	0.005	Proportion of population to be recruited for the main queue every week.	Set experimentally so that the number of new relationships formed is similar to the number of relationships dissolved.

Table 8 continued.

Probability of infection	0.01	The probability that an infected agent will infect their partner in a given week.	A reasonable value within the range of reported values [20].
Initial infected	0.01	The initial proportion of the population that is infected with the STD.	Arbitrary; a small value was chosen to investigate diffusion through the network.
Seed time	20	The time at which initially infected agents begin to transmit to their partners.	Chosen through experimentation – this value represents the amount of time for relationship formation to reach equilibrium.
Age of Removal	65	The age at which agents are removed from the simulation.	Value used in the CAN model [56].
Age of Introduction	15	The age of the agent being introduced into the simulation when replacing an outgoing agent.	The approximate age of sexual debut [2].

The data used for comparison come from national population-based household surveys conducted in 2002, 2005, and 2008 [2]. The purpose of these surveys was to monitor sexual behavior in South Africa. Demographic, social, and behavioral information was obtained from 23,369 individuals through personal interviews. We compare summary statistics of the data to simulation output: the prevalence of multiple sexual partners (defined by whether individuals have had multiple sexual partners in the past year) in each sex, and the prevalence of age-disparate relationships among young individuals (defined by whether individuals less than 20-years-old had a partner that was five or more years older) in each sex. These summary statistics are proportions and were chosen because they were determined in the report to be significant factors contributing to the HIV epidemic. While ideally we would compare the distributions (e.g., the distribution of age differences in relationships), the only data available are summary statistics (proportion and 95% confidence interval) about the population as a whole.

Distance is calculated as the sum of the absolute value of the difference between simulation summary statistics and survey summary statistics (note that these are proportions and hence normalized). There are a total of 26 summary statistics: 18 for multiple partners (3 age groups x 3 time points x 2 sexes), 4 for generational relationships (1 age group x 2 sexes x 2 time points), and 4 for age-disparate relationships (1 age group x 2 sexes x 2 time points). A total of 10,000 30-year simulations were run with populations of 10,000 individuals. We used the arbitrary distance threshold of 250 resulting in 1561 accepted simulations.

Figure 39 shows the simulation output compared to survey summary statistics for young individuals having an age-disparate relationship in the past year (additional graphs comparing simulation output for multiple partnerships are in the Appendix: APPENDIX A. FULL ABC CALIBRATION OUTPUT). The graph implies that our model is able to reproduce the age-disparate relationship trends seen in the survey data. In particular, young women have more age-disparate relationships than their male counterparts.

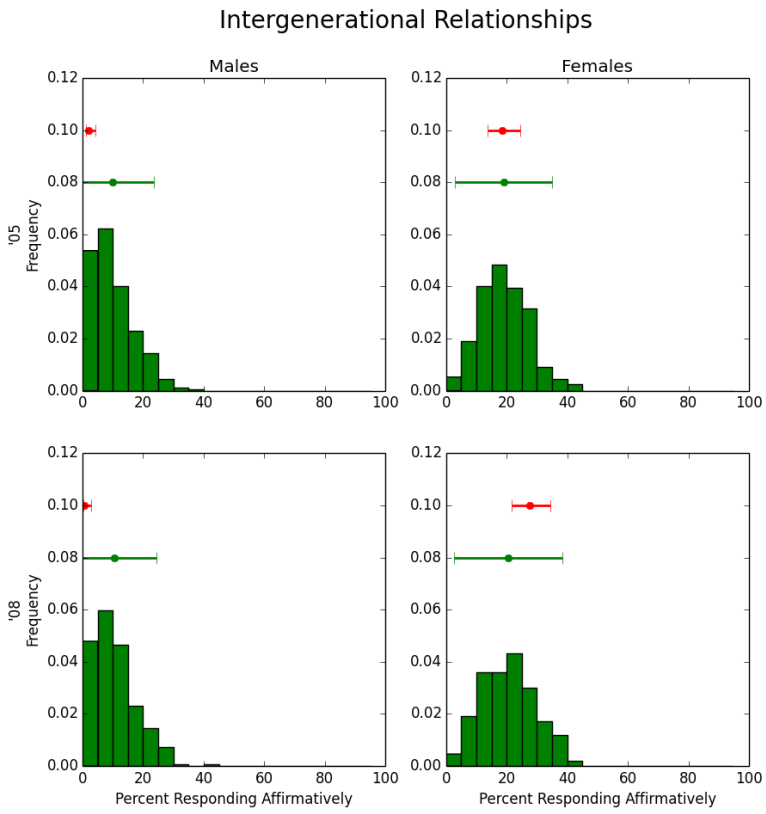

Figure 39: Age-disparate relationships in the past year among individuals 15-24 years old. Top graphs show data from 2005, and bottom graphs show data from 2008. Red dot and error bars show mean and standard deviations obtained from survey data, green dot and bars show the corresponding values from the 207 accepted simulations. Note that the confidence placement of the confidence intervals along the y-axis is arbitrary. The bar graph shows the distribution of output from accepted simulations. The figure shows that the simulation is able to produce trends like those seen in the real world.

4.4 Reducing Variation in Model Output

The calibration of the previous section shows that our model can reasonably reproduce a real-world sexual network with respect to summary statistics of the age-mixing pattern and degree distribution. In this section we investigate the effect that population size has on model output. To do this we model three scenarios with varying levels of heterogeneity: (1) agents form relationships based only on their sex – relations are independent of the agent's age and sexual behavior index; (2) Agents form relationships based on age and sex, but not their sexual behavior

index; (3) Agents form relationships based on their age, sex, and sexual behavior index. For each scenario we use three population sizes: 10^2, 10^3, and 10^4. For each scenario and population size we run the simulation 10 times to produce a distribution of model output.

Figure 40 shows disease prevalence over time for each of the nine models (three heterogeneity scenarios with three population sizes each). For the simplest scenario in which agents only form relationships based on sex, 1000 agents may suffice to accurately describe epidemic trends. However, the two more complex scenarios, which include age and sexual behavior indices in the probability function, require as many as 10,000 agents to reduce variability substantially. Additionally, the figure suggests that using a smaller population and averaging over many simulation runs is not a satisfactory solution: robust results are obtained through large population sizes. The true sexual network seems to be global [89] and have a high degree of heterogeneity [90]. In order to get the same level of invariability in the model as what we believe is true in the real world simulations need to use large populations of agents.

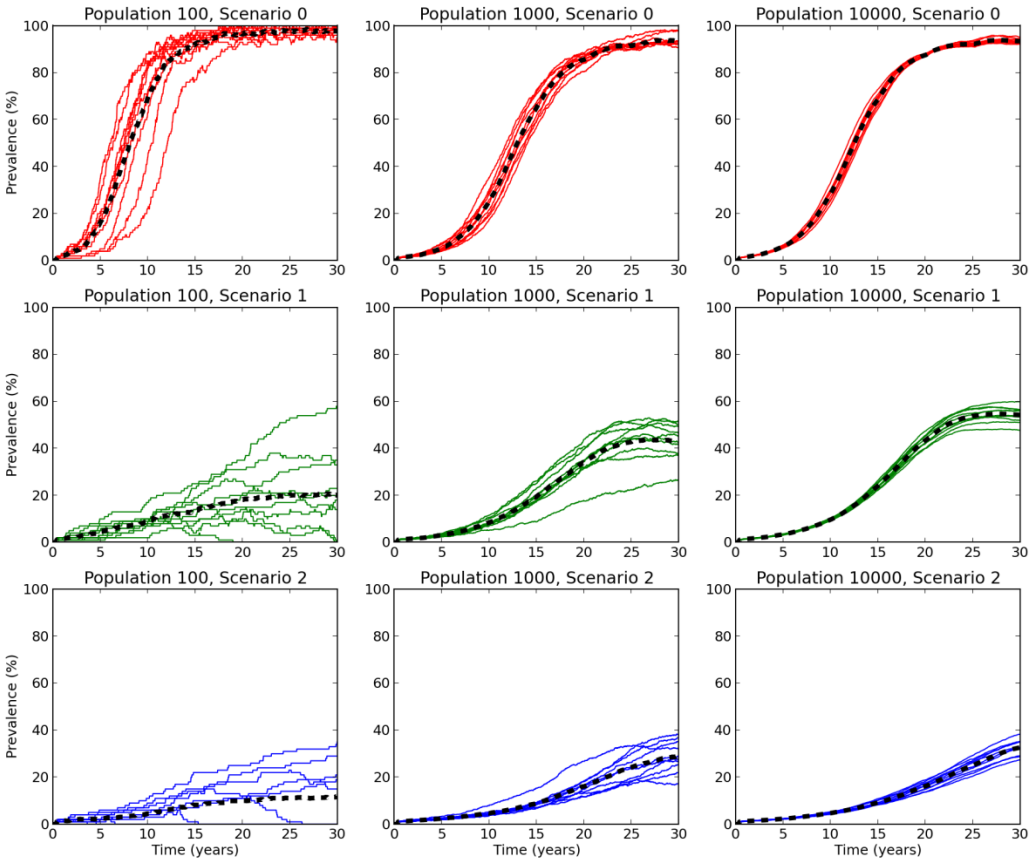

Figure 40: Ten prevalence curves for each of three scenarios with three different population sizes. Average of the 10 runs is shown with black dotted line. Too few agents increases variation in model output and produces unmeaningful results.

4.5 Performance Analysis

The previous section showed that as heterogeneity in behavior agents increases, the number of agents in the simulation must also increase. Here we analyze the performance of our algorithm and show that it scales well to large population sizes. The model implementation was run for different population sizes over 30 years. Parameter values were determined by the ABC method and are shown in Table 8.

Figure 41 shows the amount of time required to run each population size. Simulations were run on a 12-core, 3.2 GHz computer with 16 GB of memory (the computer was oversubscribed with processes). Even though it exhibits quadratic runtime, the quadratic coefficient is sufficiently small that larger population sizes can be run in reasonable time. For example, a 30 year simulation with 150,000 agents can be run in about six hours.

Increasing the number of queues (by decreasing the size of age-cohorts) increases the age-mixing precision, but at the cost of increased run time. For example, on the same computer a 30-year simulation with 1,000 agents and the default 20 queues takes approximately 9 seconds to run. Using 50 queues causes the simulation to take 14 seconds, and using 100 queues takes 17 seconds.

In simulation runs where grid queues are forced to resort for every suitor (as opposed to saving accept/reject decisions for the next suitor) runtime increased substantially: 10,000 agents required 70 minutes. With resorting the same population size required only 4 minutes.

Since the number of relationships grows quadratically with the number of agents in the simulation, memory consumption also exhibits quadratic behavior. Figure 42 shows the quadratic relationship between memory consumption and population size, and suggests that the size of the population is limited by computing capacity, not memory constraints.

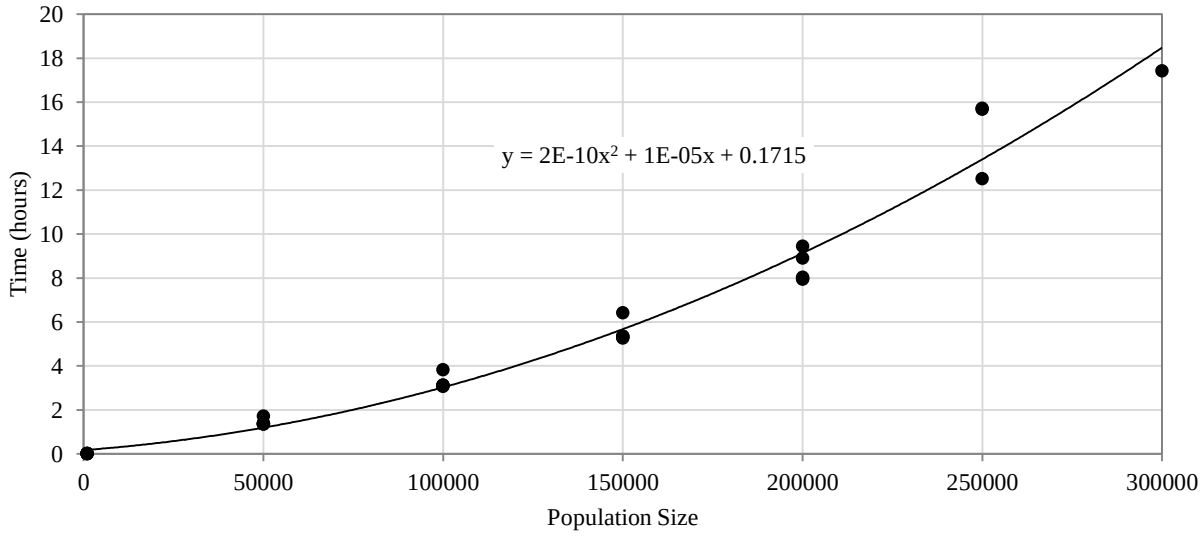

Figure 41: Run times for simulation runs with varying population size. Simulations were run over 30 years on a 16 core 3.2 MHz computer. The elapsed time grows quadratically, but the quadratic coefficient is sufficiently small that larger populations are capable of being simulated.

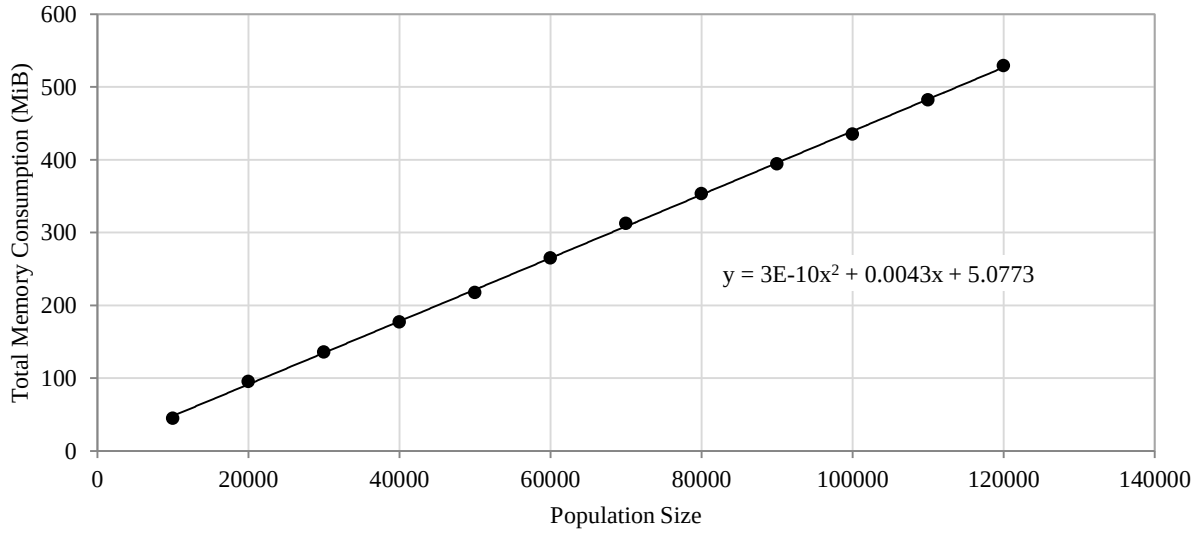

Figure 42: Memory consumption with varying population size. Since the number of relationships grows quadratically with the number of agents, so does the amount of memory consumed.

4.6 Discussion

While agent-based models can generate more complex and detailed projections than deterministic models, the stochastic nature of the simulations can make small population sizes produce biased, unstable dynamics. Simulating larger populations reduces model variability, but can take a prohibitive amount of time to run. Here we've presented a parallel algorithm and implementation that can run multi-year simulations with large populations in a reasonable amount of time on commodity hardware. Other agent-based models of STDs are not capable of simulating more than approximately 10,000 agents [91]. Note that direct comparison to other agent-based sexual network simulators is difficult since many, such as CAN [56], STDSIM [50], and EMOD [92], do not report the computational aspects of their implementations. McCormick et al. do report comparable runtimes of their model in supplementary material, but discussion of hardware and exact parameters is absent [93].

In our implementation speed up is obtained in two ways. First, the implementation minimizes the time spent simulating unlikely events (such as very age-disparate relationship) by partitioning agents based on their sex and age. This allows us to efficiently find matches for suitors in parallel. Second, the simulation avoids redundant calculations by caching and exploiting the accept/reject decision from the previous suitor.

The model is capable of producing a broad range of networks with demonstrated similarity to those observed in the real world. We acknowledge though that several simplifying and limiting assumptions were made that preclude the model from making real world predictions in its current form: that the population size remains constant over time and maintains a uniform age distribution is inconsistent with demographic data for South Africa [54]; seeding initially infected individuals at random is inconsistent with high risk transmission clusters [3]; and a constant probability of transmission is inconsistent with strong evidence that HIV infection probability varies according to viral load and other factors [83]. These assumptions were made for simplicity, but do not detract from our goal of efficiently simulating dynamic sexual networks. We plan to address these limiting assumptions in future work.

In the next chapter we use geographic partitioning to divide the country into simulations of smaller communities and distribute them across multiple machines. In this way we scale the model implementation further and simulate migration between communities.

4.7 Conclusions

The model and implementation is a novel simulation algorithm for large-scale agent-based modeling of sexually transmitted diseases. The model is flexible to many epidemic scenarios and able to simulate many complex social phenomena observed in real-world sexual networks. The implementation takes advantage of multiple processors and scales well to larger population sizes. Unlike ordinary differential equation models, the model can produce fine-grain cross-sectional distributions of the population (such as the percentage of agents that had more than two partners in the last year). And unlike standard agent-based modeling approaches, we do not simulate all agents as unique individuals. Through the use of queues we keep the individual-level characteristics necessary for simulating fine-grain processes while also eliminating some of the computational overhead intrinsic to agent-based modeling.

CHAPTER V SIMULATING MIGRATION AND SEXUAL NETWORKS IN A DISTRIBUTED ENVIRONMENT

5.1 Introduction

We have shown that agent-based models of HIV transmission are well-suited to simulating individual level processes, like complex age-mixing patterns or heterogeneity of sexual behavior. Similarly important are the geographic location and migration patterns of individuals because they can determine the spatial distribution of a sexually transmitted disease [94]. How and where individuals migrate affects sexual network connectivity, bridging geographically disparate network components. The mobility of a population can indirectly determine epidemic persistence through seeding and reseeding infected communities and can undermine localized intervention attempts [94, 95]. Mobility and sexual risk also seem to be related: whether because of loneliness or less family contact, mobile individuals have an increased number of sexual partners and engage in more sexually risky behavior [96, 97]. The interaction of population mobility and increased sexual risk has had a large impact on the initial HIV epidemic in South Africa [22, 98–100]. Any attempt then to eradicate HIV must consider the impact that mobility and migration have in the perpetuation of the disease [23].

Agent-based models are well suited to simulating a mobile population: Wood et. al developed an ABM for simulating migration in Burkina Faso which used the Theory of Planned Behavior as a basis [101]. Silveira et. al created an ABM which describes the economics of rural-urban migration in an Ising-like model [102]. It is necessary to use a large population in their implementation though to avoid “small world” phenomena that can emerge purely from having too few agents in the model. However, increasing the number of agents in a model also increases

the amount of time required to run the model. In the previous chapter we presented a parallelized model and implementation that takes advantage of multiple processors on a single computer and significantly reduces the amount of time required for larger populations. This too has limits though which suggests that obtaining further speed improvements will necessitate distributing model computations onto a cluster of computers.

In this chapter we present our multi-scale model of HIV transmission in a large dynamic sexual network. Our algorithm geographically partitions a model world so that dynamic sexual networks for different regions can be simulated in parallel on separate nodes of a cluster. The advantage of this approach is two-fold: large populations mean that additional heterogeneity can be modeled with less chance of introducing small-world effects; and geographic components of HIV transmission such migration and mobility can be modeled. The novelty of the model comes from the use of geographic partitioning which allows us to distribute the simulation on a cluster of computers and to simulate migration processes. In the next section we discuss our model, describing (1) the simulation of a small community as a single network, (2) larger communities as multiple small communities, and (3) a country as multiple large communities. In Section 3 we present a performance analysis of the model implementation, and perform an exploration of HIV prevalence and persistence in a range of migration scenarios.

5.2 Methods

The model described in this chapter is an extension of the model described in the previous chapter that simulates a single closed community on a single compute node with multiple processing elements (cores). We extend the original model by first simulating large communities (i.e. >500,000 agents – too large for a single compute node) as multiple interconnected smaller communities, each on a separate compute node of the cluster; and second

simulating an entire country as a network of large communities connected via cyclically migrating agents. For completeness we briefly describe our previous model for simulating a dynamic sexual network on a single node, and then describe our methods for extending the model to multiple nodes and connecting them via migration.

5.2.1 Small communities as single networks

Our model uses the function $P_{ij} = \exp\left(PM \times \left(\left|AD - \left(PAD \times PAD_{growth} \times MA\right)\right|\right)\right)$ to calculate the probability of a relationship forming between two agents i and j, where PM is a probability multiplier, AD is the age difference of the couple from the male perspective, PAD is the preferred age difference defined from the male perspective, PAD_{growth} is the preferred age difference growth, and MA is the mean age. Note that *AD* and *MA* are calculated based on the candidate couple, while *PM*, *PAD*, and PAD_{growth} are parameters of the simulation. This probability function allows relationship formation to be informed by the preferences of both agents, and by fine-grain details about the agents, like their age.

The model keeps track of agents looking for a relationship with *queues*, a list of objects that is ordered by some criteria. The model creates a *grid of queues* where the dimensions of the grid reflect the attributes that we want to use to inform relationship formation. At initialization, agents are created based on age and sex distributions. Their age and sex in turn determine their respective queue, which represents their birth and sex cohort. The agents are placed in their queues and ordered based on the time since they were last allowed to form a relationship (which is initially the same for every agent).

The relationship formation procedure then takes place in two phases. In the first phase, a limited number of agents seeking new relationships are recruited from their queues (with agents

who have been waiting the longest being recruited first) and used to populate another queue, called the *main queue*. The agents placed into the main queue are ordered based on their respective age and gender cohort. In the second phase, the relationship operator considers each agent in the main queue and attempts to match him or her to agents that are still in the queues. For each potential match, a random number is drawn from a uniform distribution and compared to the probability function described above. When two agents form a relationship a random value from an exponential distribution is drawn to determine the duration of the relationship. The matching procedure for the relationship operator then proceeds by iterating through the main queue and making relationships for each suitor. Figure 35 is a visual representation of the model.

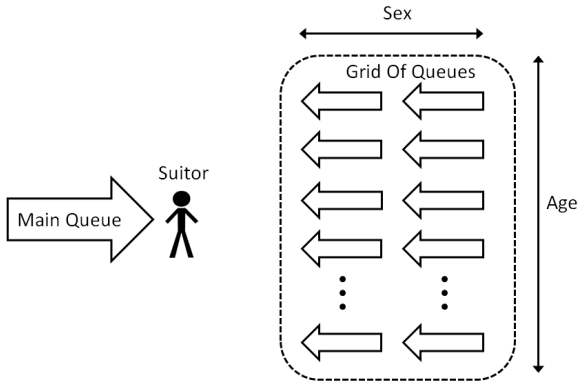

Figure 43: The simulation is made up of a grid of queues, which holds all the agents, and a main queue that holds agents waiting to be matched. We refer to the agent at the head of the main queue as the suitor.

After the initial population has been created the infection operator seeds HIV by infecting a few agents chosen at random. At each time step the infection operator iterates through the list of infected agents and propagates infection to uninfected partners. While there is substantial evidence that the probability of HIV infection changes with the viral load and CD4 count of an HIV-positive individual, our model assumes a constant probability of infection for simplicity.

The time operator ensures the passage of time by ending relationships and removing and replacing agents that have reached the maximum age (65-years-old in our model). When a relationship ends the two agents return to their respective queues, which were assigned based on their age and sex at initialization. The oldest queue is queried for agents who have exceeded the maximum age and then replaced by a 15-year-old agent of random sex and random DNP. Thus the population size remains constant over time, and the demographics remain approximately similar throughout the simulation. Relationships for agents being removed are ended, and the surviving partners are allowed to form new relationships. As simulation time progresses, the relevant time window also changes and the time operator creates new queues as needed.

The progression of these three operators simulates a dynamic sexual network with infection propagation: the relationship operator forms relationships based on a probability function; the infection operator propagates infection through the sexual network; the time operator dissolves relationships and removes and replaces older agents. The model's implementation places each queue on a separate processor core parallelizing it to a single compute node. This enables us to simulate populations about to 700,000, at which point the amount of time required (approximately 20 hours on a 16-core 2.6GHz computer) is prohibitive.

5.2.2 Large communities as multiple small communities

In order to simulate larger communities (>500,000) we distribute the computation across multiple nodes of a cluster. We extend the original model by simulating a large community as a *group* of small communities. The group is composed of a single *primary* community and multiple *auxiliary* communities, each referred to as *sub-communities*. Each sub-community is placed on a separate node of the cluster. The primary node maintains the data structure for the

sexual network, and each sub-community works to build the network. This is to avoid time-consuming update messages about the state of a distributed network.

Each sub-community follows nearly the same process of relationship formation as before with two exceptions: (1) during the recruitment phase, instead of a recruited agent being placed into the sub-community's main queue by default, the recruited agent is sent to the main queue of a sub-community randomly chosen from the group. Note that the recruited agent may still be placed in the main queue of the sub-community from which it originated. (2) After the recruiting phase, sub-communities similarly iterate through the main queue matching suitors and agents in the queues. However, auxiliary communities send relationship matches to the primary community to be added to the sexual network. The primary community adds the relationships to the network after checking that neither agent formed another relationship this round. This check is done to ensure that agents do not have more relationships that their respective DNP.

After relationships are formed the primary community, the only sub-community in the group with knowledge of the sexual network, performs infection propagation and removes relationships that have ended. Each sub-community, in parallel, removes and replaces agents that are beyond the replacement age. The distributed version for simulating larger communities is represented visually in Figure 44.

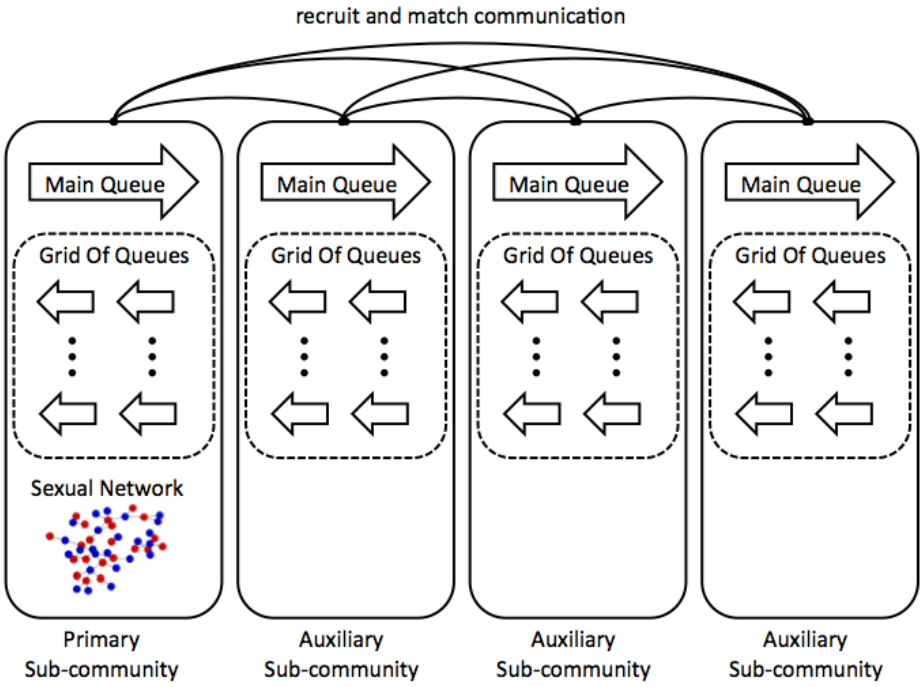

Figure 44: A large community is simulated as a group of sub-communities. Each sub-community recruits agents from their grid of queues to populate one of the main queues in the group. Relationship matches made by auxiliary sub-communities are sent to the primary sub-community to be added to the sexual network. The primary sub-community performs the infection propagation and expired relationship removal steps. Each sub-community removes old agents from their respective queues in parallel.

5.2.3 Multiple communities as multiple large communities

To simulate HIV propagation at a national level we consider different provinces as separate, but interconnected large communities. The communities are connected via cyclically migrating agents that travel between their home and work communities. A community determines which and to where agents migrate based on South Africa's 2011 census [77]. The data indicates the number of individuals in each province that resided in another South African province during the previous census in 2001. We use this number as a proxy for the relative pull, or *gravity*, between the provinces. The gravity is normalized to determine the probability that an agent initialized in community i migrates to community j. The migration network is represented visually in Figure 45.

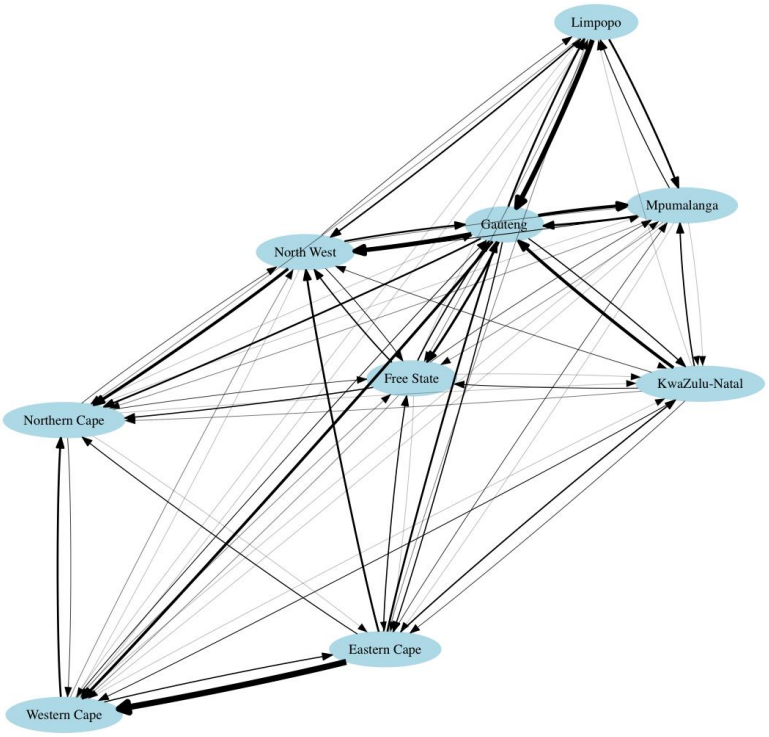

Figure 45: A visual representation of the migration network between provinces. Each province is connected to every other province through migration. Darker arrows represent more migration, while lighter arrows represent less migration. For readability self-looping arrows have been omitted.

5.2.4 Calibration

Our goal in this work is to develop a model that is capable of simulating sexual networks informed by age mixing and migration patterns that scales well to larger populations. Where possible, literature informed parameters values. Where no literature is available we used the approximate Bayesian computation (ABC) method[87, 88] to infer reasonable values that produced a sexual network that is approximately similar to real life. The parameter values are given in Table 8. Comparison to the real-world network can be found in APPENDIX A. FULL ABC CALIBRATION OUTPUT.

Table 9: The parameter values used in the simulation. Parameters are taken from literature or inferred using ABC.

Parameter	**Value**	**Description**	**Justification**
Probability multiplier	-0.2	Coefficient in the probability function that determines the baseline probability of a relationship forming for deviation away from the preferred age difference.	ABC
Preferred age difference	-0.1	Coefficient in the probability function that determines the age difference for which the baseline probability of a relationship forming is highest.	ABC
Preferred age difference growth	0.1	Coefficient in the probability function that determines the amount that preferred age difference grows with mean age.	ABC
DNP Distribution	$1.2 \times$ Power(0.1)	The distribution of desired number partners; also known as the degree distribution.	ABC; distribution used in the CAN model [56].
Duration Distribution	$30 \times$ Exp(1)	When a relationship is formed, the duration of the relationship is pulled from this distribution.	ABC; duration of relationships are approximately exponential [57].
Age Distribution	Uniform(15,65)	The distribution of ages when agents are initially created.	Arbitrary; A uniform distribution was chosen for simplicity.
Sex Distribution	Bernoulli(0.5)	The distribution of sex when agents are initially created.	The approximate sex ratio in South Africa [2].
Initial recruitment rate	0.02	The initial proportion of agents recruited from queues to populate the main queue.	Set experimentally to allow the simulation to quickly reach equilibrium.
Warm-up period	20	The number of weeks that the simulation uses the value of *initial recruitment rate*.	Set experimentally to allow enough time for the simulation to reach equilibrium.
Recruitment rate	0.005	Proportion of population to be recruited for the main queue every week.	Set experimentally so that the number of new relationships formed is similar to the number of relationships dissolved.
Probability of infection	0.01	The probability that an HIV-positive person will infect their partner in a given week.	A reasonable value within the range of reported values [20].
Initial infected	0.01	The proportion of the initial population that is infected with HIV.	Arbitrary; a small value was chosen to investigate diffusion through the network.

Table 9 continued.

Seed time	20	The time at which initially infected agents begin to transmit to their partners.	Chosen through experimentation – this value represents the amount of time for relationship formation to reach equilibrium.
Age of removal	40	The age at which agents are removed from the simulation.	Largest value possible with 5 years age bins, 2 sexes, and 16 cores per nodes.
Age of introduction	15	The age of the agent being introduced into the simulation when replacing an outgoing agent.	The approximate age of sexual debut [2].
Number of years	30	The number of years simulated.	The approximate time between South Africa's first cases of HIV and the present.
Time home	3	The amount of time that a migrant agent will spend at home community.	Reasonable value based on previous models [98].
Time away	15	The amount of time that a migrant agent will spend at their away community.	Reasonable value based on previous models [98].
Migration scale	1.0	The relative "pull" or "gravity" between communities.	Values from the 2011 South Africa census [77].

5.3 Performance Analysis

Large community simulations were run with different population sizes and an increasing number of nodes. Each compute node has 64 GB of memory and 16 2.6 GHz cores. Each additional compute node, up to five total nodes, reduces the amount of time required to run a simulation as seen in Figure 46. Runtimes cease to improve after five nodes however, and each additional compute node exhibits a diminishing return on speed up.

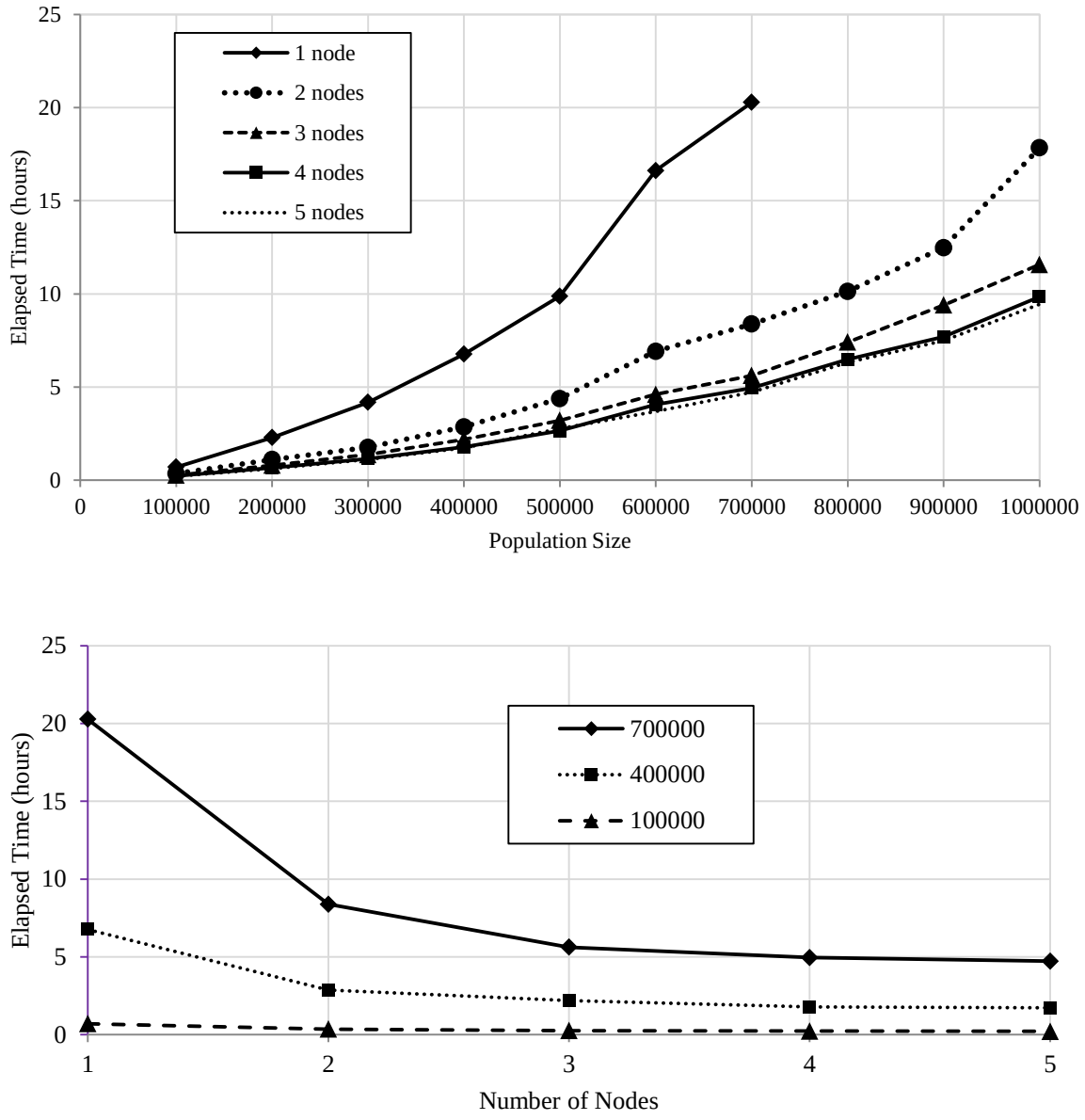

Figure 46: Top: the amount of time required to run different population sizes with varying number of compute nodes in a cluster. Bottom: up to four additional compute nodes can reduce runtime, at which point additional parallelism does not seem to be beneficial.

To assess the computational overhead of migration we ran two migration scenarios with increasing population sizes. Both scenarios simulated three inter-migrating communities. In the first scenario each community is on a single node (using a total of three nodes for the simulation), and in the second scenario each community across is distributed across two nodes

(using a total of six nodes for the simulation). Figure 47 shows the amount of time required to run each scenario for different population sizes. Note that the population size is the total number of agents in the simulation with agents evenly distributed among communities (e.g. for a simulation with 30,000 agents each community has 10,000 agents).

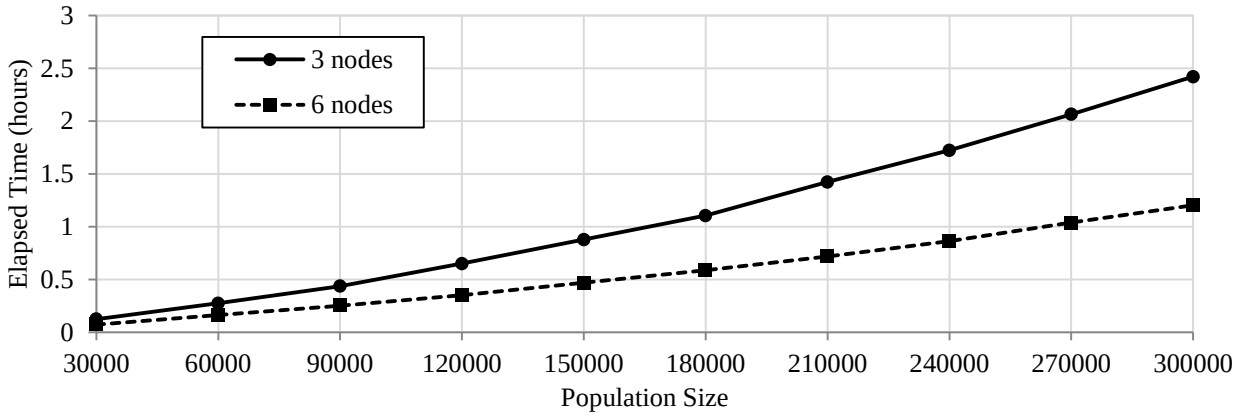

Figure 47: Runtimes for a simulation with three inter-migrating communities. The first scenario uses three nodes, and the second uses six nodes. The runtimes for the two scenarios suggest that the computational overhead of migration is not very large.

5.4 Parameter Exploration

We explore the parameter space of the model by simulating various migration scenarios. In particular, we vary the relative probability of migration between provinces, the lengths of time that agents stay at their home and work location, and the spatial distribution of initial infections in a model with 3 large communities connected by migration. To explore the effect of each of these three parameters, we randomly select a value from a discrete range and fix all other parameters with default values (enumerated in Table 10). Each simulation then runs for 30 years with 124,000 agents – approximately $1/100^{th}$ of the actual population. We run 100 such simulations, and investigate their effect on HIV prevalence in the different communities.

The relative probability of migration between communities is varied by scaling the matrix obtained from the South African 2011 census [77]. The value for migration indicates the power that the migration matrix is taken to: the value of 1.0 uses the matrix as it is obtained from the

census, while 0.1 uses the matrix raised to 0.1. Values less than 1.0 produce more migration than the census, while values greater than 1.0 produce less migration.

Table 10: Ranges of the values used in parameter exploration.

Parameter	Values	Default	Description
Migration scale	[0.1, 0.5, 1.0, 2.0]	1.0	The power that the migration matrix is raised to in the simulation.
Time scale	[1, 3, 5, 7]	3.0	The amount of time that migrating agents spent in their home community. The amount of time that a migrating agent spends in their away community is 5 times the amount of time spent in the home community.
Spatial distribution of initial infections	[isolated, geographically dispersed, population dispersed]	Geographically dispersed	The geographic dispersion of the initially infected agents: isolated indicates that all initially infected are in a single province, dispersed indicates that initially infected are selected from all provinces regardless of population density, population dispersed indicates initially infected are selected from all provinces relative to population density.

The amount of time spent home and away appears not to have a large effect on disease prevalence (Figure 49), while the amount of migration does (Figure 48). More migration produces lower prevalence values in the seed community (Community 0 – left) because infection events are occurring in other communities instead of within. For the non-seed communities (Community 1, middle, and Community 2, right) the relationship between the amount of migration and 30 year prevalence is non-linear: too much migration (values of 0.1) results in the infection less diffusion, and too little migration (values of 2.0) results in the community not being seeded with infection at all.

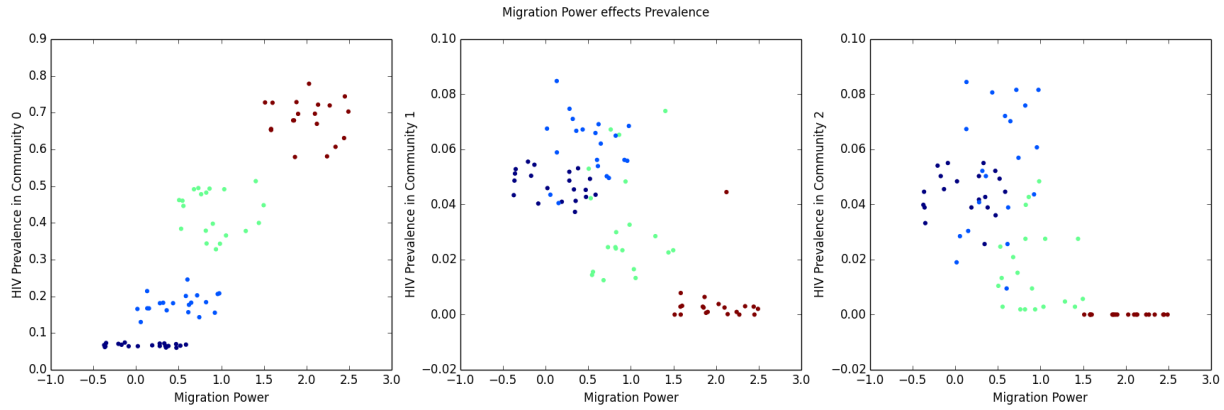

Figure 48: The effect of migration on 30-year prevalence for a 3 community simulation.

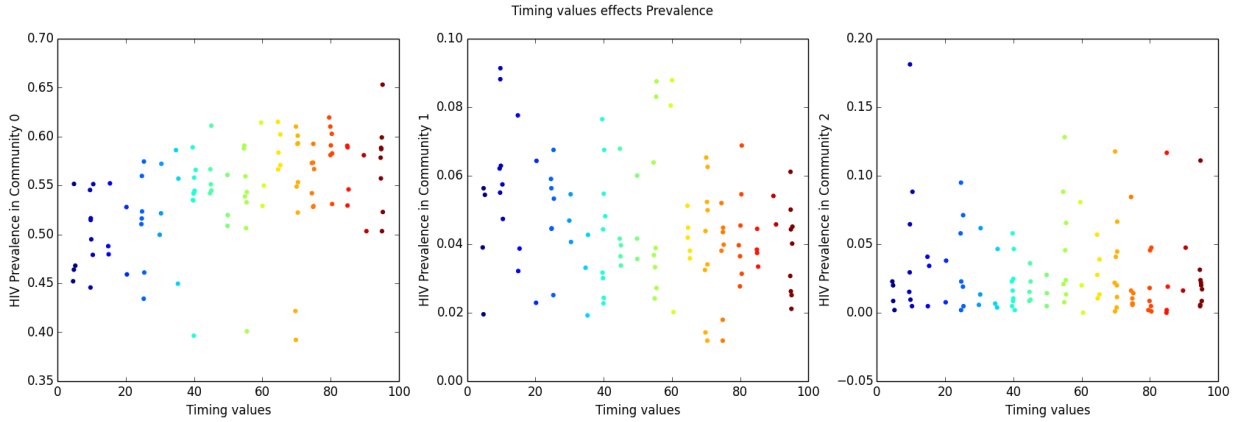

Figure 49: The effect of time spent home and away on 30-year prevalence for a 3 community simulation. The different values don't seem to have a large impact on disease prevalence.

We expanded the exploration of the migration parameter by using all 9 provinces in the simulation and fixing all other parameters. The number of agents used in the simulation was 472,000. We again ran simulation 100 times in order to obtain a distribution of disease prevalence after 30 years. Figure 50 shows the distribution for each of the 9 provinces in 5 different migration scenarios. Infection was initially seeded in the Gauteng province arbitrarily.

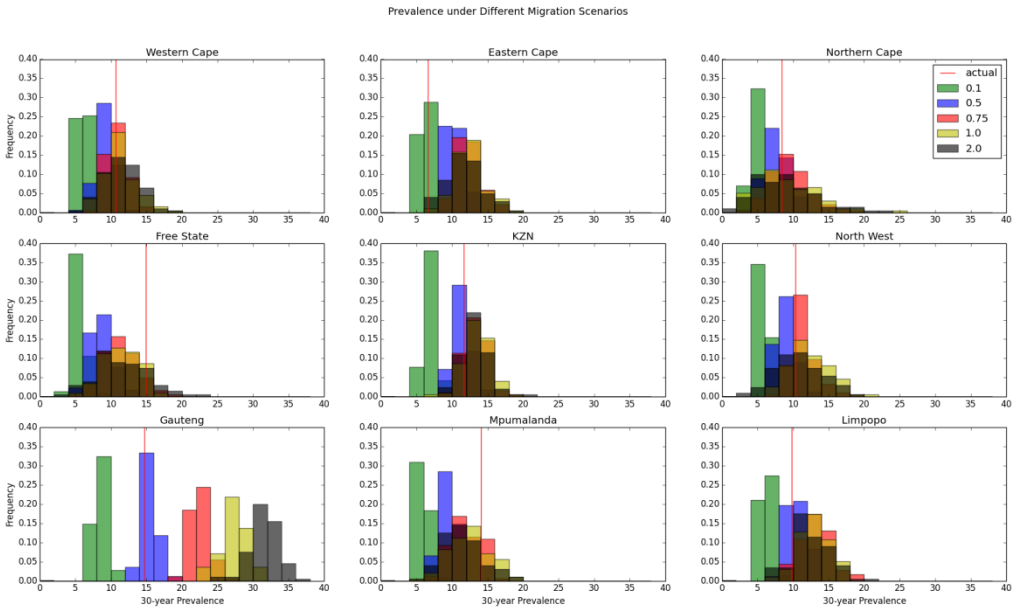

Figure 50: The distribution of disease prevalence after simulating for 30 years under 5 different migration scenarios for the 9 provinces of South Africa.

The role that migration plays in Gauteng, the seeded community, is readily apparent: more migration means that the infection is spread to other provinces and hence is not able to spread as extensively within as with less migration. The influence of migration on disease prevalence in the other provinces is less apparent, but it seems that the values for 0.75 and 1.0 (which produces migration patterns most similar to real life) produce distributions that are higher on average. This implies that the real life patterns perhaps contributed greatly to the diseases diffusion and deviation in either direction (more migration or less) would have dampened the epidemic outcome.

5.5 Discussion

Agent-based models of sexually transmitted diseases are able to simulate fine-grain processes such as complex age mixing behavior and geographically specific migration that are known to contribute significantly to disease persistence. However, the large amount of

heterogeneity in agent behavior requires that agent-based models use large population sizes in order to avoid small-world effects. Our model uses multiple cores on a single node, and multiple nodes on a cluster to distribute the work of building a complex dynamic sexual network. When simulating a large community as multiple small communities, the model scales well with each additional compute node used. When additionally simulating migration between large communities, the model continues to scale well with larger population sizes.

Our goal in this work was to create a model that was able to simulate complex age-mixing patterns and geographically specific migration patterns. However, we admit that the model has not been vigorously validated and hence is not suitable for forecasting epidemic trends. Future work will focus on using more realistic parameters such as a non-uniform age distribution and a probability of infection that changes with time since infection.

5.6 Conclusions

In this paper we presented a parallel and distributed algorithm for simulating dynamic sexual networks. While agent-based models of migration and agent-based models of sexually transmitted diseases have been developed previously, to our knowledge this is the first agent-based model that simulates disease propagation in a migrating sexual network. Additionally, because the simulation is distributed across several nodes of a cluster the model is able to scale well to larger population sizes and thus avoid small-world phenomenon.

CHAPTER VI CONCLUSIONS

In this thesis we have shown how agent-based models can be used effectively and efficiently to simulate the diffusion of sexually transmitted diseases. The algorithms presented here are effective at simulating fine-grain processes difficult to capture in compartmental models, and they are efficient through the use of parallelism and distributed computing. In chapter 2 we presented the mathematical formulation for simulating dynamic sexual networks. We showed that the model and implementation were able to simulate important sociological processes such as age-mixing and concurrency. In chapter 3 we used a simplified version of the mathematical formulation and machine learning algorithms to find good combinations of HIV prevention strategies. In chapter 4 we presented a parallelized algorithm for the model and showed that the implementation scales well to larger population sizes. In chapter 5 we geographically partitioned the sexual network and simulated them in parallel on separate nodes of a cluster. We took advantage of the geographic partitioning to additionally simulate migration and movement of individuals. We conclude with a discussion of agent-based modelling, its uses for finding good combinations of prevention methods, and how we can scale it to large population sizes.

6.1 Agent-Based Modelling

Generally speaking, models try to explain and give understanding to processes or phenomena seen in the world. Agent-based models attempt to understand these processes by simulating individuals and the individuals' behaviours from which the process emerges. This is in comparison with compartmental models that aggregate individuals into groups (or compartments) and use more coarse-grain view of a system to describe a process.

For example, an agent-based model might simulate the behaviour of 100 individual wolves and 10,000 individual sheep, each with unique location in the simulated world, to explain how a predator-prey system works. A compartmental model on the other hand might aggregate the wolves and sheep into two compartments and use the total number of animals in each to explain the same system. Choosing a model type then depends on the level of detail desired: if the starting location of the animals is thought to be important (e.g., if animals are so far apart that wolves have difficulty finding sheep) then an agent-based model is a good option. However, if location is not thought to be important (e.g., all animals are randomly intermixing) then the extra granularity gained by simulating the actions of individuals is likely unnecessary and a compartmental model might be a better choice.

In this work, we chose to use agent-based models to simulate HIV transmission because we are interested in modelling fine-grain processes that may otherwise be lost in a compartmental model. For example, we are interested in simulating HIV transmission in a highly heterogeneous population – i.e., a population where all the individuals have characteristics and behaviours that are unique. To do this we simulate individuals in a given population with individual agents, and assign them characteristics, like gender, age and an intrinsic sexual activity drive. These agents move around in a simulated world and form and dissolve sexual relationships with other agents based on the assigned characteristics. In this way the agents produce a dynamic (i.e. changing over time) sexual network through which HIV is able to spread. In this way agent-based models are intuitively similar to how the real world operates: HIV diffusing through a population is the result of discrete events (like forming a relationship or becoming infected with HIV) happening to distinct individuals. These discrete events contain randomness, but are informed by individual characteristics – their individual sexual drive or a

preference for older partners. This means that, like in real life, different agents experience different events at different times.

6.2 Combination HIV Prevention

Each prevention method has a different financial cost of implementation, as well as varied community acceptance. The important question for a government on a fixed budget is which programs will be effective in a community, and in what combination and in what order should they be implemented?

Our work on simulating combination HIV prevention investigated not only the overall effect on important variables, but also potential interactions among interventions. For example, in the absence of all other interventions, HIV counseling and testing conveys little or no protective effects for uninfected individuals. When utilized alongside a national male circumcision program, however, counseling and testing becomes a point of referral and a catalyst for the male circumcision program.

The implication is that a better allocation of scarce public resources is possible through modeling and simulation. For each of the possible prevention methods there exists a point of diminishing return at which more money invested provides little pay off and is better allocated to other programs. For example, distributing 2 million condoms may reduce the total number of new infections significantly, but doubling the number of condoms distributed will not halve the number of new infections. When each prevention method is used optimally there are no lost opportunity costs for spending more on one method of prevention versus another.

6.3 Simulating large populations

While an agent-based model is somewhat intuitive, a modeller faces many questions while developing the model. For example, how many agents are needed to adequately simulate the underlying processes of HIV propagation? A tempting solution is to simply use the largest population size possible. However, as the number of agents in the simulation increases so does the amount of time required to run the model – and a model that takes months or years to run is not very useful. Large population sizes are necessary though to avoid small population phenomena: processes that emerge purely from having unrealistically few agents being modelled. For example, consider a purely heterosexual agent-based model of HIV transmission. If we use a population with 4 agents whose sex is randomly assigned then our model will fail to see any transmission in approximately 12.5% of simulations. This is because in approximately an eighth of those simulations all the agents will be the same sex. It's for this reason that larger population sizes are necessary to create robust and reliable results from simulations.

In an attempt to simulate very large populations (millions of agents), we've developed parallel algorithms that distribute the model's workload among multiple processors on a single computer and among multiple computers on a cluster of machines. Running the agent-based model in a high performance setting enables us to significantly speed up the simulation of large population sizes. With these new algorithms, simulations with large population sizes that used to take months now only take hours.

All of these challenges are computational in nature. We can develop more efficient algorithms for simulating larger and more dynamic population. We can build more sophisticated models that more closely match sexual network and demographic data. However, these point to a larger challenge: how do we simulate a process that is governed by highly volatile rules that are

constantly changing? We can collect more data and build more models, but the reality is that effectively simulating sexual networks means effectively simulating human behaviour – and effectively simulating human behaviour is a hard problem. This does not mean that modelling should not be done – modelling efforts have already saved lives. It means that all assumptions made when developing a model should be carefully documented, and the implications of these should be thoroughly investigated.

If we employ useful tools like sensitivity analysis and approximate Bayesian inference to explore the range of answers that models produce, given the data and additional assumptions; if we explicitly acknowledge the gaps in our knowledge and our suspicions of biased data; if we clearly state the intentions and limitations of our models; then the use of models will no longer be a straw man treasure hunt for the fountain of truth or unscientific attempt at predicting the future. Models can be what they are: a systematic exploration of plausible trends and phenomena in a stylized model world; a representation of a system that helps us to understand the findings of previous empirical studies; an aid in narrowing our focus for follow-up empirical experiments.

APPENDIX A. FULL ABC CALIBRATION OUTPUT

In this section we provide the entire output from the approximate Bayesian computation (ABC) in CHAPTER IV A PARALELLIZED ALGORITHM FOR SIMULATING DYNAMIC SEXUAL NETWORKS, Section 4.3 Implementation and Calibration. The algorithm calibrates the model by finding sets of parameter values that produce the most desirable output. The algorithm repeatedly chooses values for parameters based on prior distributions and then runs the simulation for model output. After many iterations the parameter sets that produced the best model output defines the posterior distribution for parameters.

The graphs below are the full output from the ABC method. We show the distribution of model outputs, posterior distributions for parameter values, and comparison of model output to data for each summary statistic. The method was run with a population of 10,000 agents, and 10,000 parameter sets were run. We used an arbitrary acceptance quality threshold of 250, resulting in 1,561 accepted simulations (16% simulation runs).

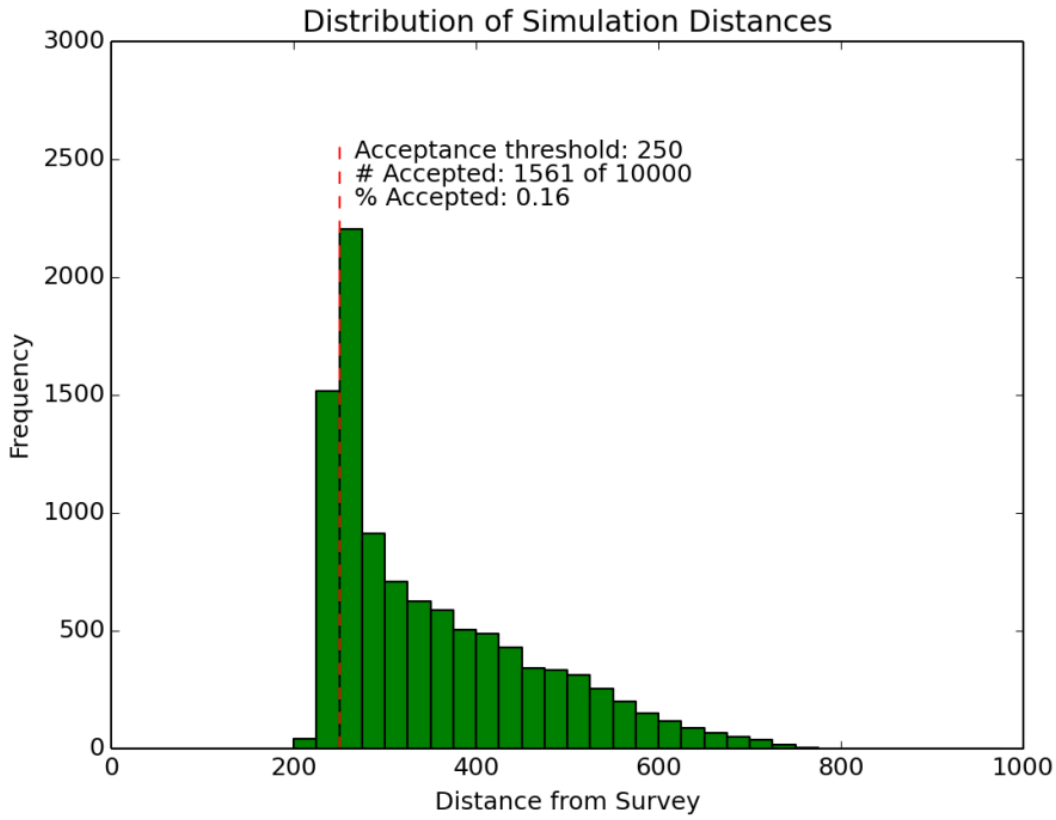

Figure A1: Distribution of distances values for the 10,000 simulation runs. Accepted simulations were those with distance less than 250, resulting in 1561, or 16% of all, simulations.

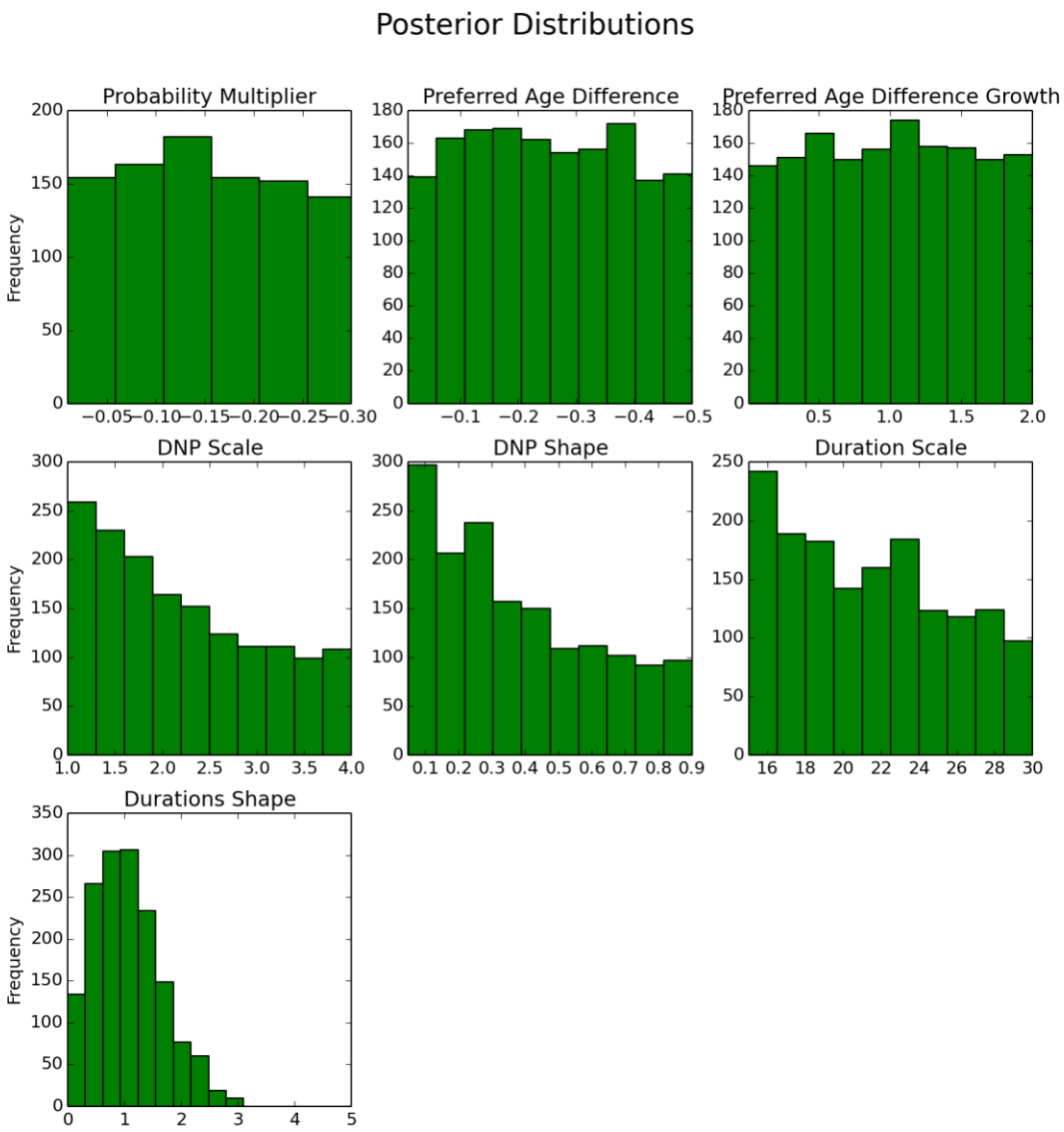

Figure A2: The posterior distributions for each of the inferred parameters.

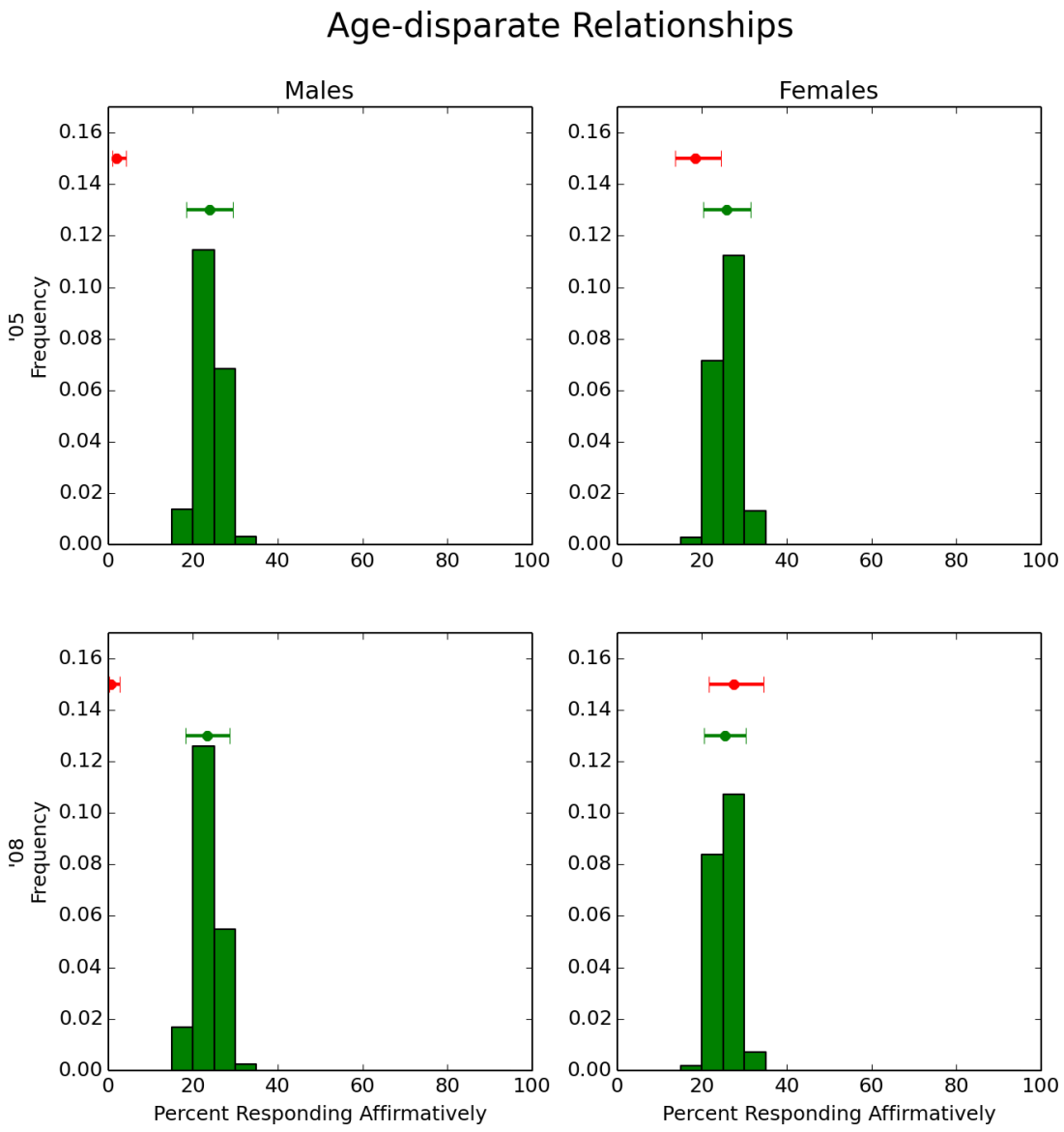

Figure A3: Age-disparate relationships in the past year among individuals 15-24 years old. Top graphs show data from 2005, and bottom graphs show data from 2008. Red dot and error bars show mean and standard deviations obtained from survey data, green dot and bars show the corresponding values from the 207 accepted simulations. Note that the confidence placement of the confidence intervals along the y-axis is arbitrary. The bar graph shows the distribution of output from accepted simulations. The figure shows that the simulation is able to produce trends like those seen in the real world.

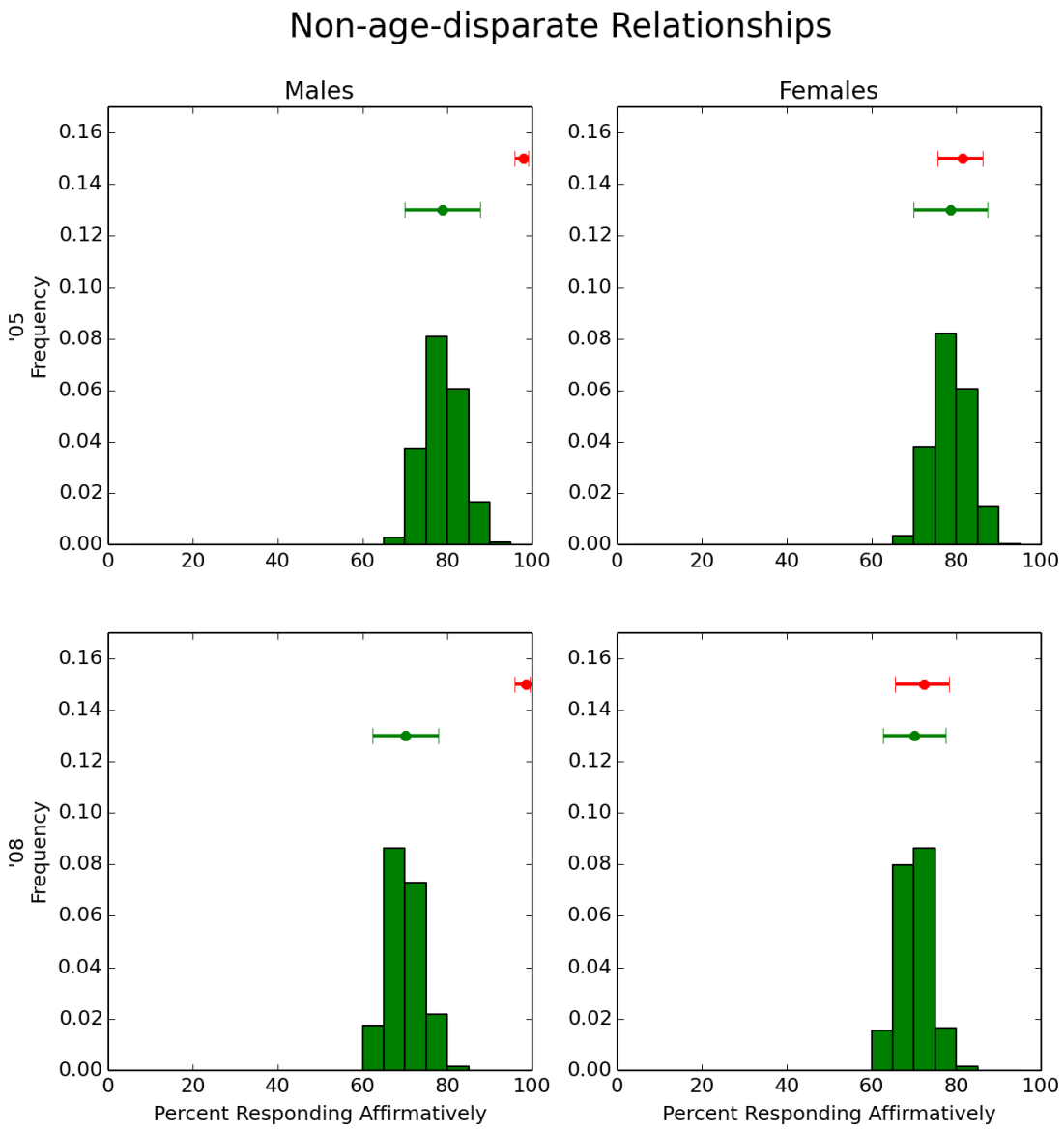

Figure A4: The distribution of the values for non-age-disparate relationships in the accepted simulations (green bars) for different sexes and survey years. The green dot-and-bar chart represents the average and one standard deviation of the distribution, while the red dot-and-bar char represents the average and two standard deviations for the actual survey data. The values that the simulation produces are similar to those seen in the survey data.

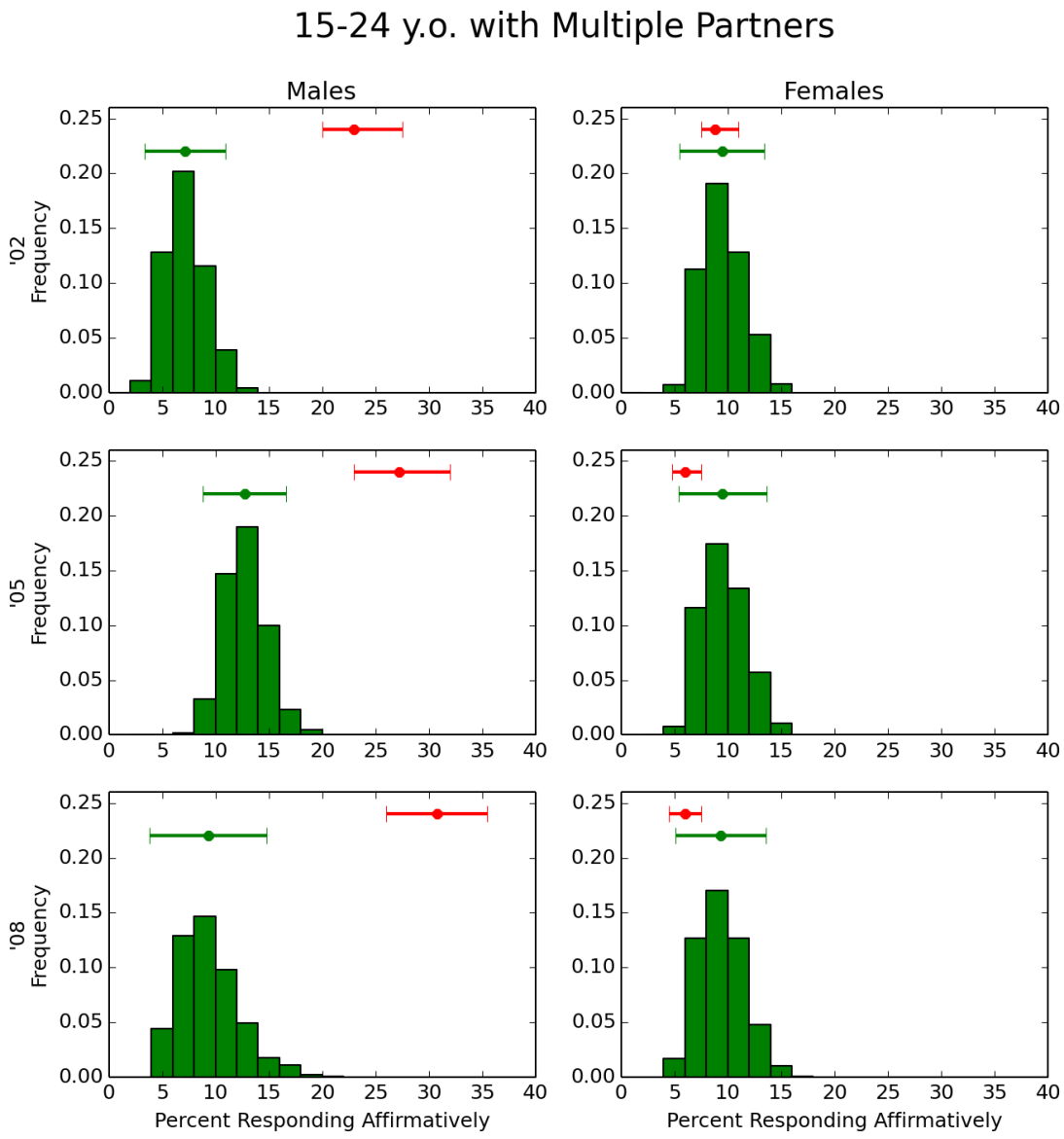

Figure A5: The distribution of 15-24 year old agents that had multiple partners in the past year (green bars) for different sexes and survey years. While the simulation values for males do not seem to align with survey values, this is likely due to bias in the data – i.e. young male agents tend to overestimate the number of sexual partners that they have had.

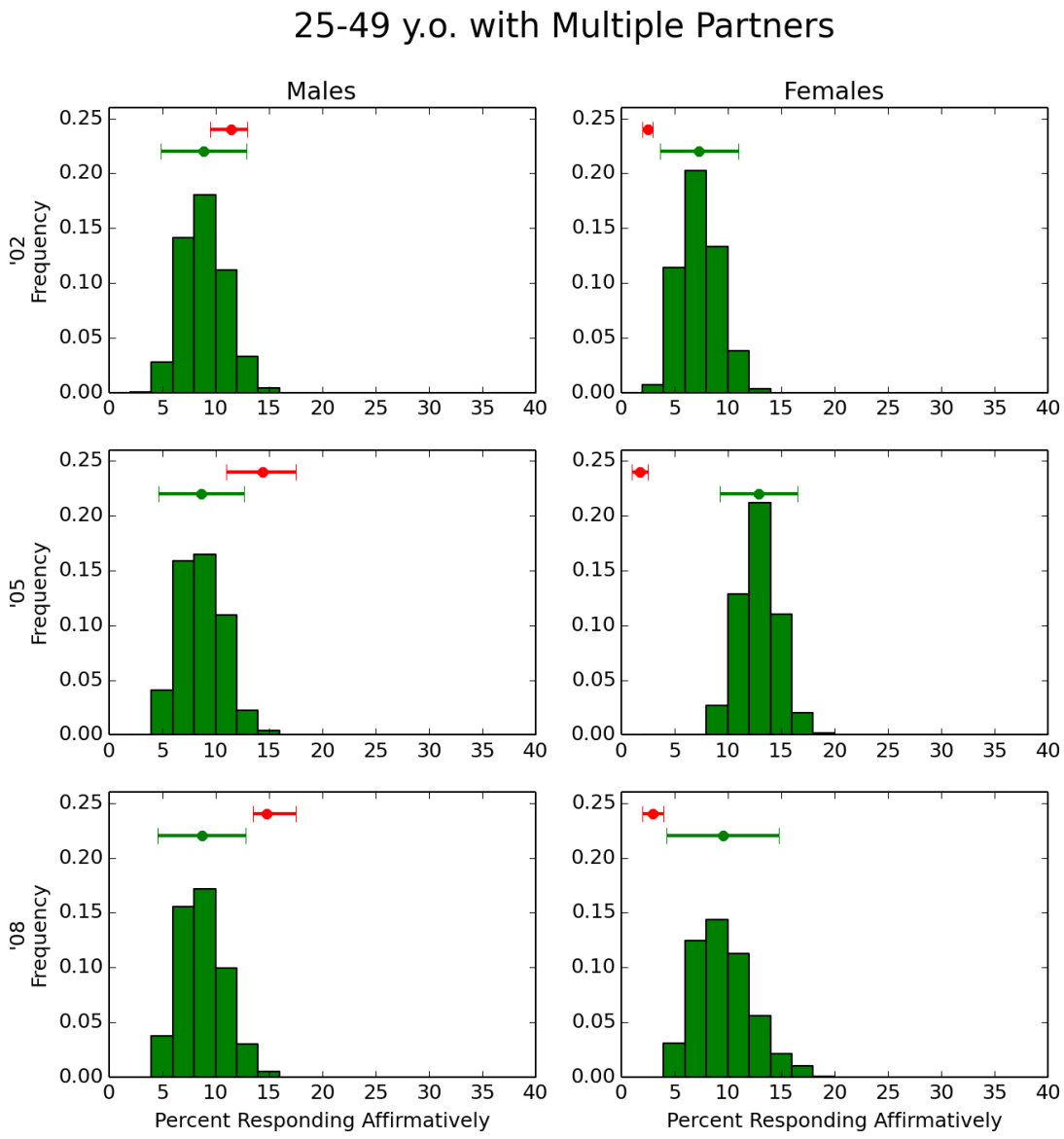

Figure A6: The distribution of 25-49 year old agents that had multiple partners in the past year (green bars) for different sexes and survey years.

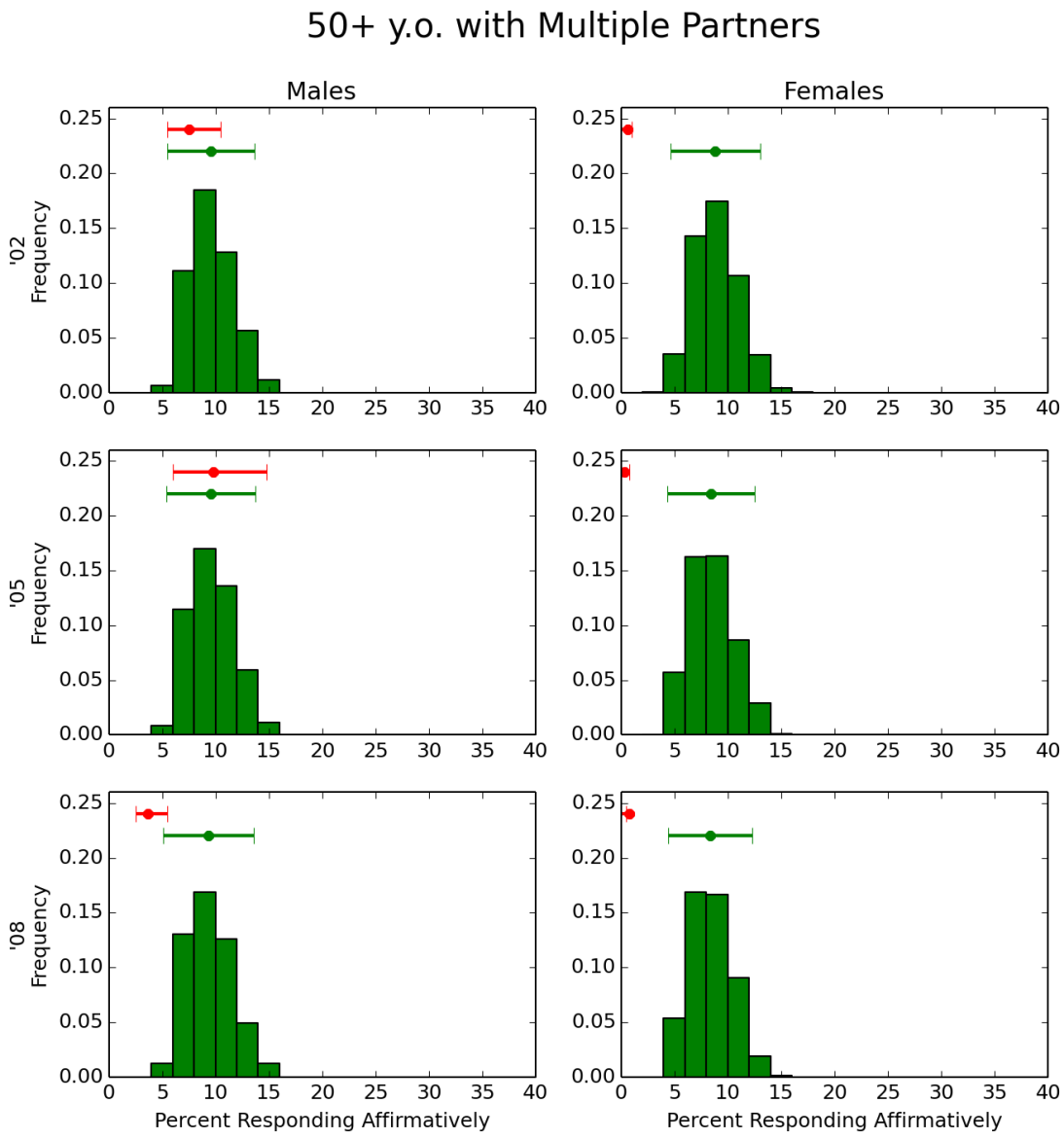

Figure A7: The distribution of 50+ year old agents that had multiple partners in the past year (green bars) for different sexes and survey years.

In order to assess the usefulness of the distance metric we reran the analysis using a random subset of simulation runs (as opposed to selecting high quality simulations runs). The figures below (blue bar charts) indicate that using the distance function is useful in determining the posterior distribution of parameter values.

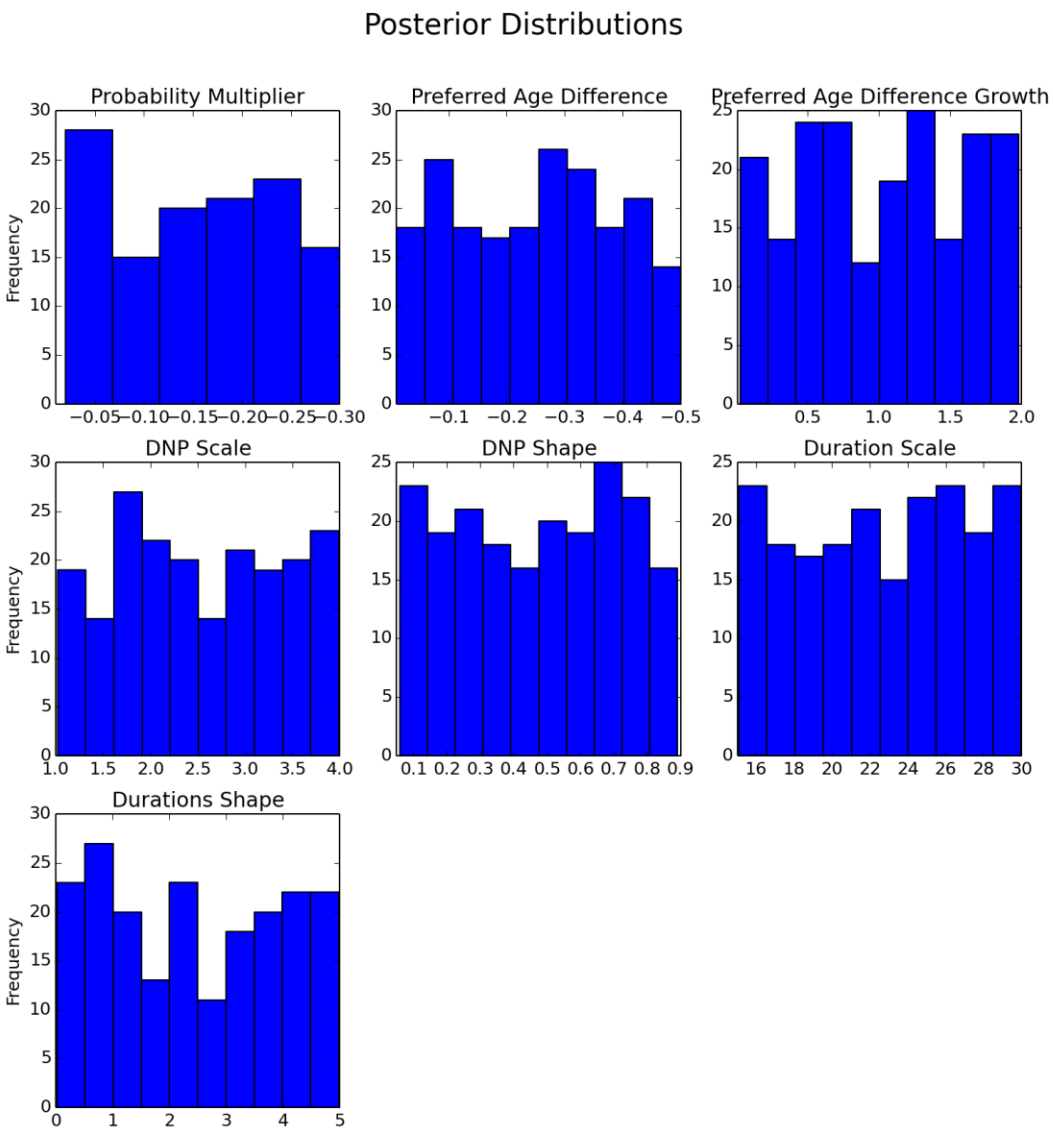

Figure A8: Posterior distribution for parameters if quality of simulation is not considered. As is expected the posterior distributions appear to be uniform between their bounds.

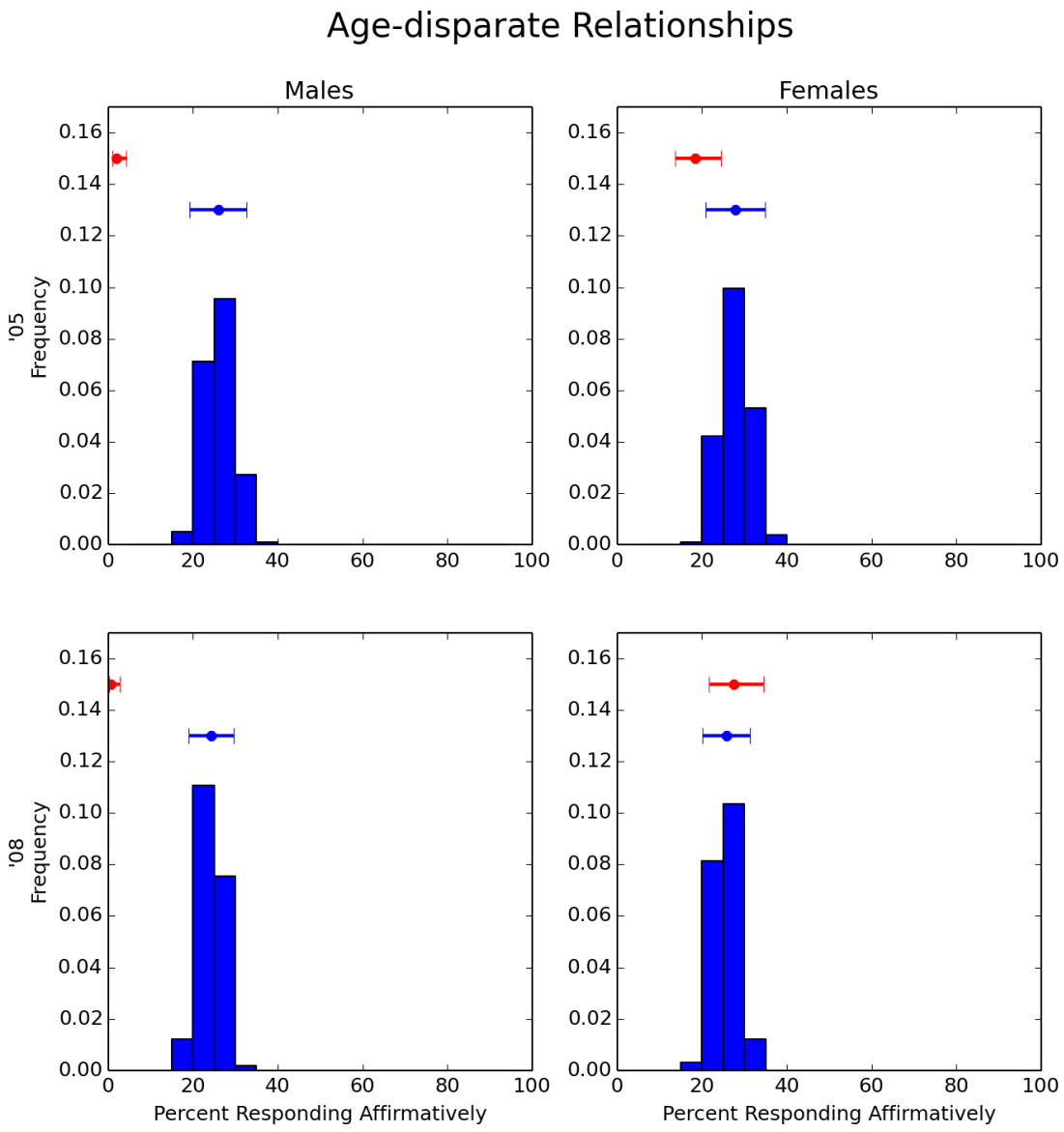

Figure A9: The distribution of 15-24 year old agents that had age-disparate and non-age-disparate relationships in the past year (blue bars) for different sexes and survey years.

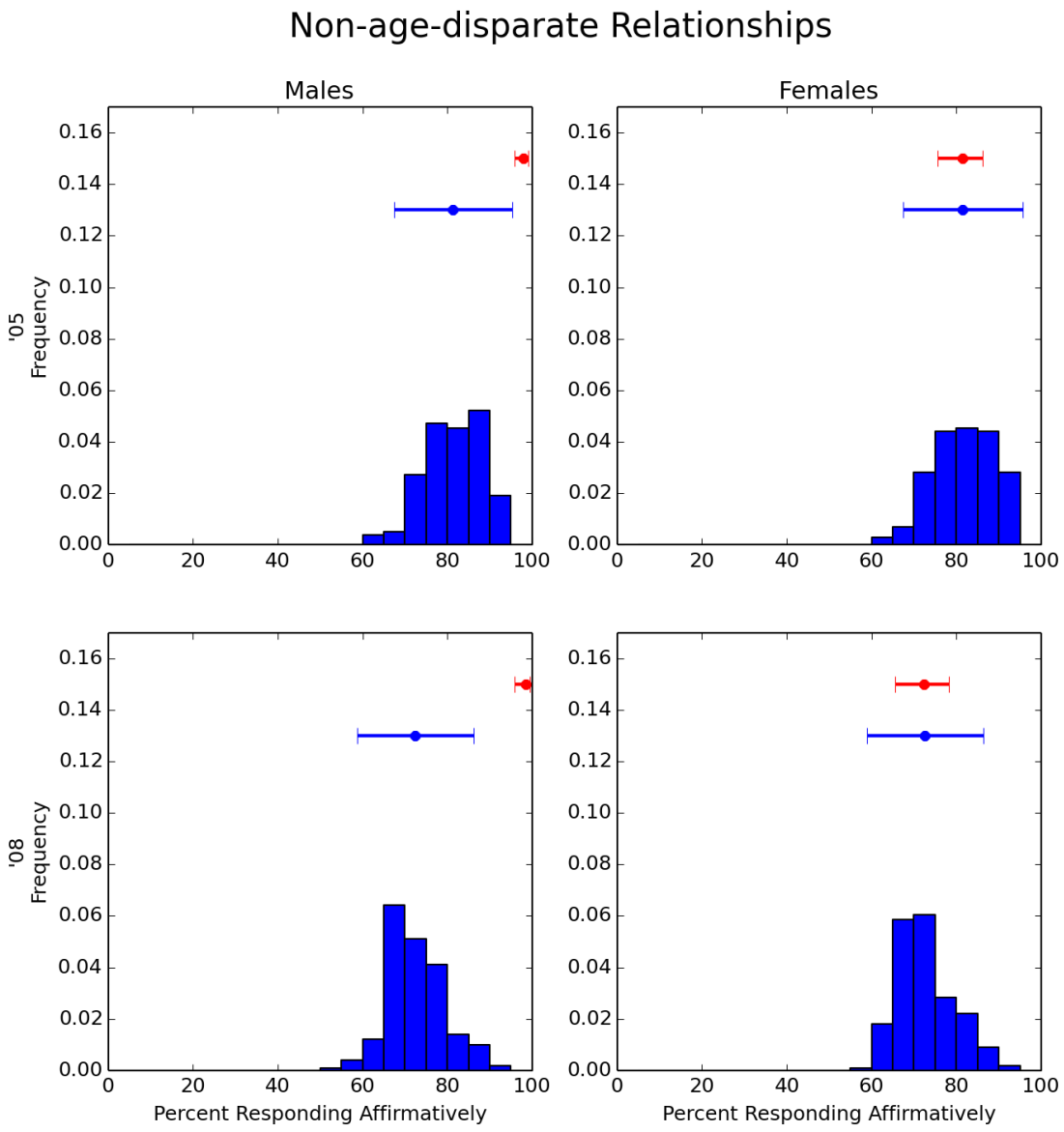

Figure A10: The distribution of 15-24 year old agents that had non age-disparate relationships in the past year (blue bars) for different sexes and survey years.

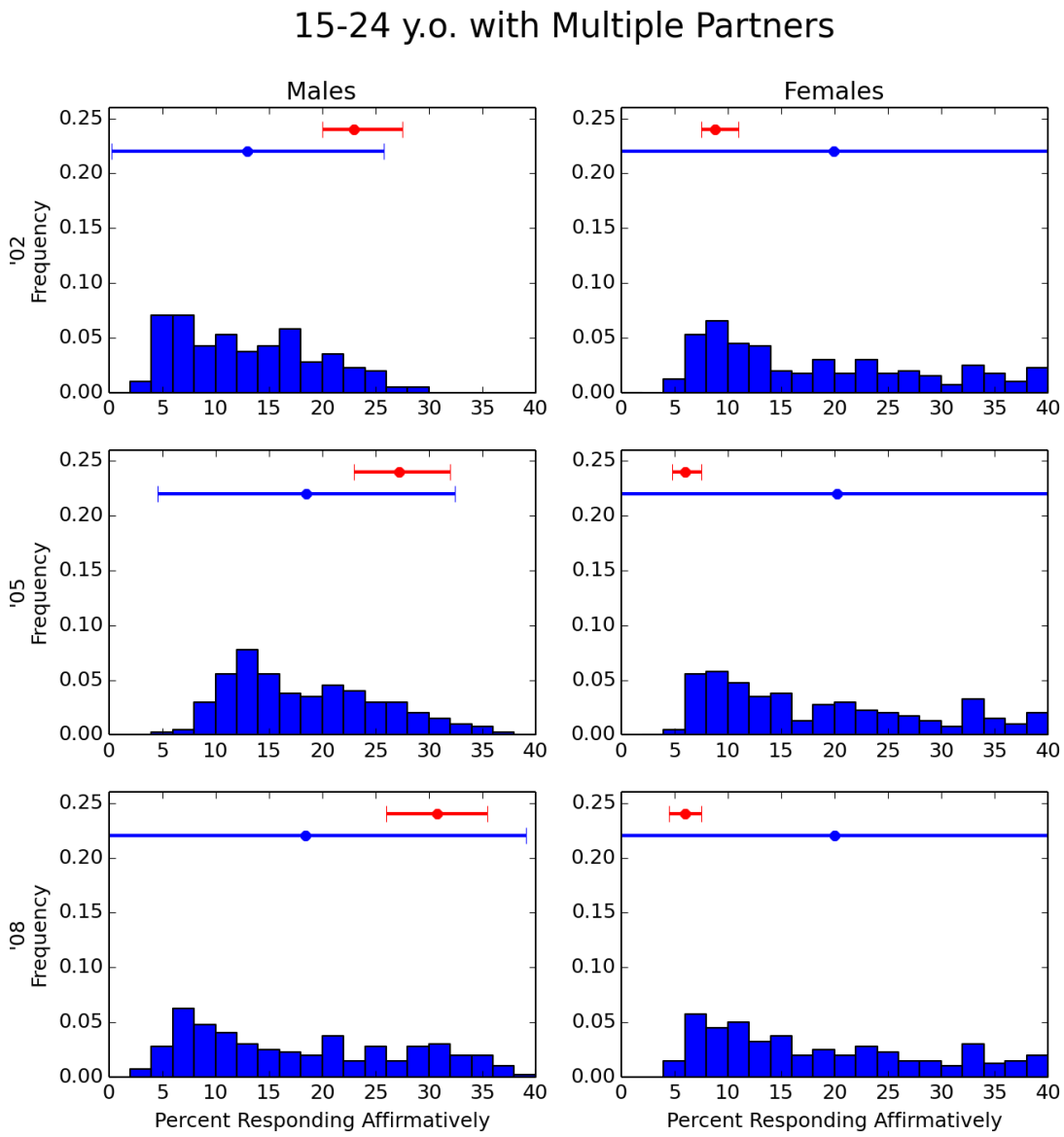

Figure A11: The distribution of 15-24 year old agents that had multiple partners in the past year (blue bars) for different sexes and survey years using the random sample of simulation runs.

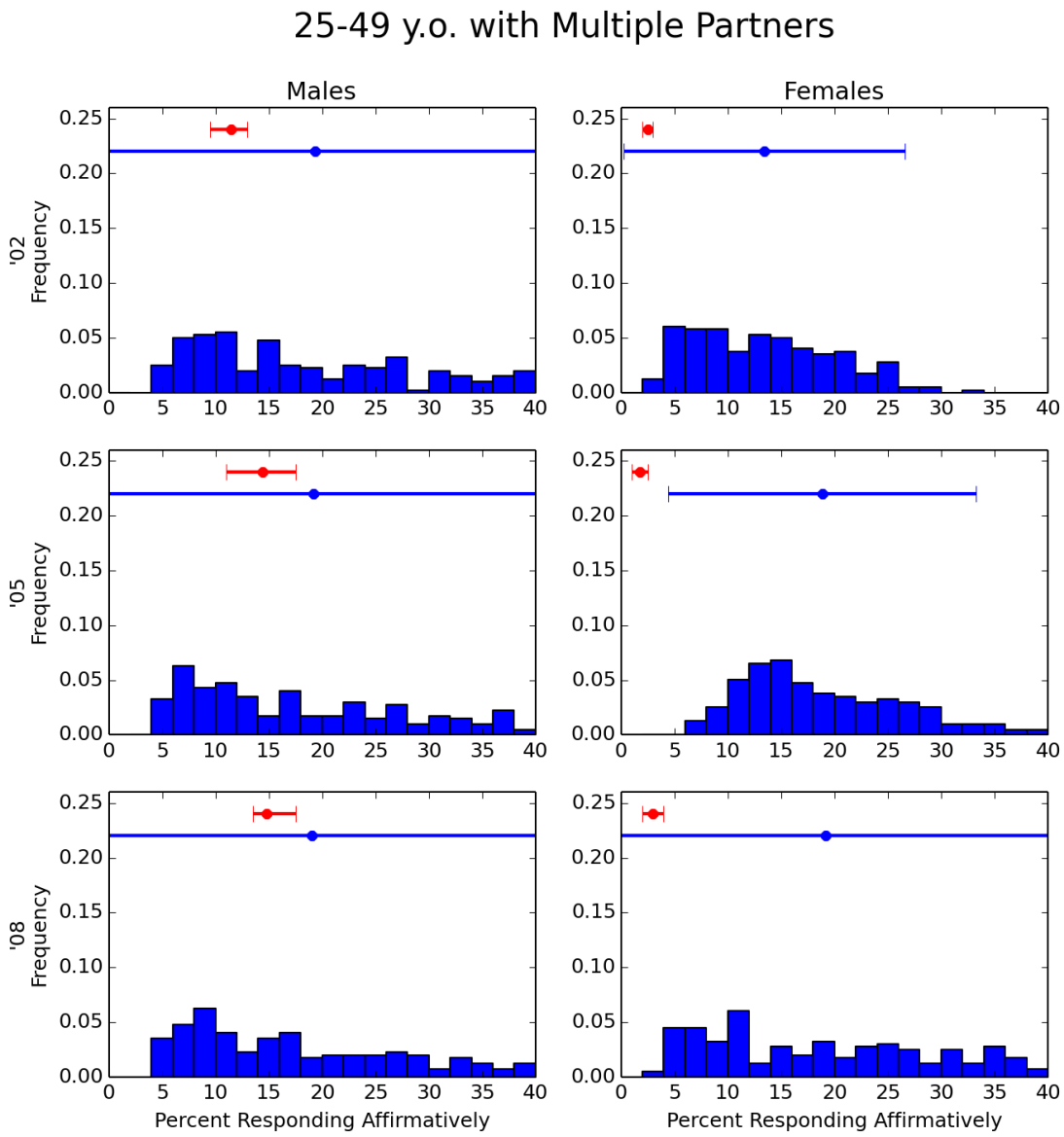

Figure A12: The distribution of 25-49 year old agents that had multiple partners in the past year (blue bars) for different sexes and survey years using the random sample of simulation runs.

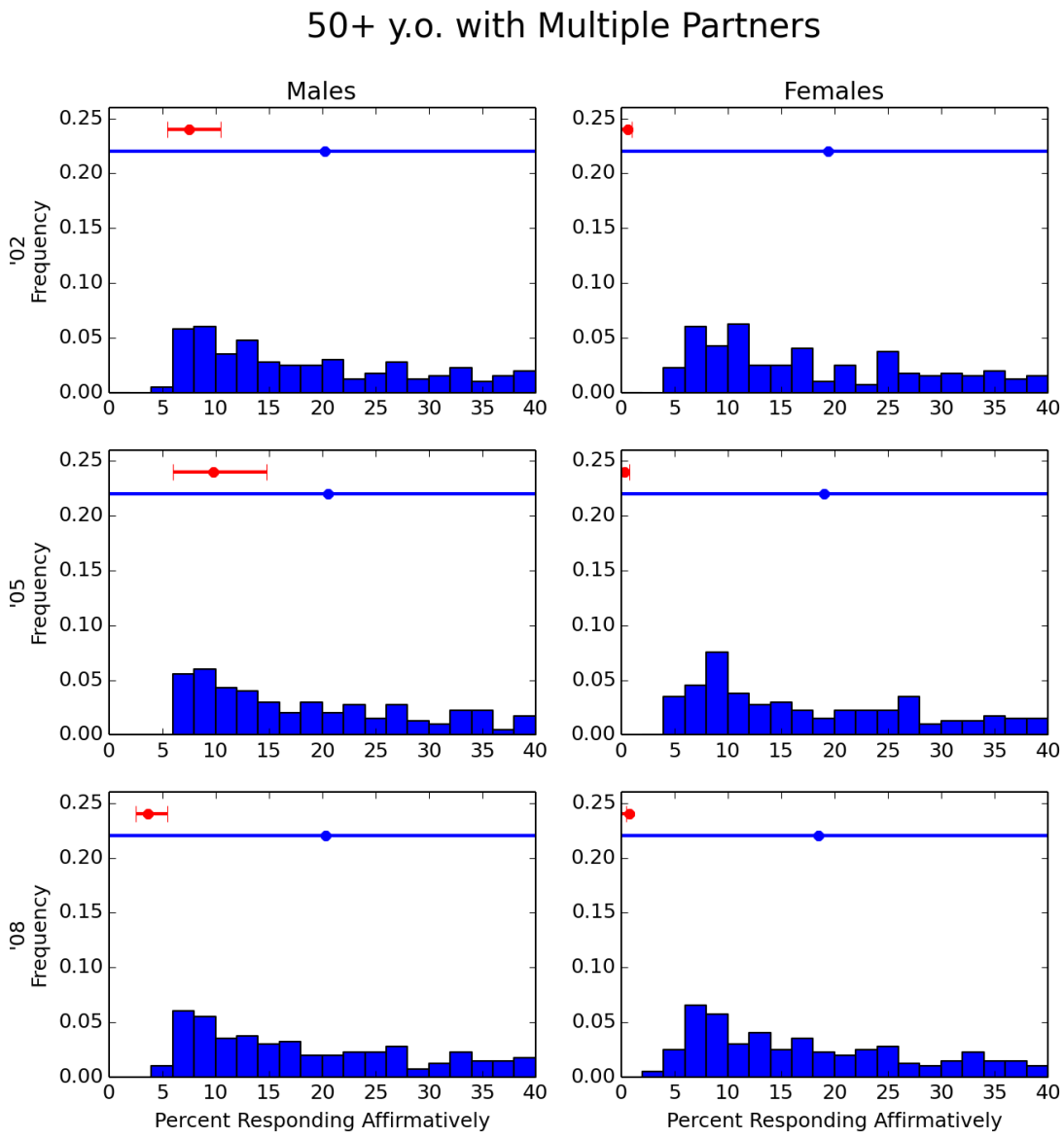

Figure A13: The distribution of 50+ year old agents that had multiple partners in the past year (blue bars) for different sexes and survey years using the random sample of simulation runs.

APPENDIX B. VALIDATION

Unfortunately, a model with a large population size is, by itself, insufficient to be a useful model. Once a model is "complete" (i.e. decisions have been made as to how many agents will be in the simulation, the events that can happen to the agents, the laws that govern these events, and the time horizon over which we want to simulate) we need to show that it is valid. This is done through a process that is aptly named *validation*. This can be hard because validation, in part, means showing that any change in the model world, and the consequences of those changes, would play out in the real world system that the model is supposed to represent. It is also the other way around where real-world changes should be seen in the model world. The conundrum is that the real world system is often too complex to test changes and their consequences – if it weren't too difficult we likely wouldn't spend time trying to model it!

An additional challenging aspect of validation is that the real world and the data derived from the real world are the result of many components and their subcomponents, and all of their interactions. The result is a complex system with many dimensions and begs several questions: how many of these components and interactions must be represented in the model? How "true" are the data collected for all of these dimensions? How does one test and confirm that the model is in line with the data across all these dimensions?

There are nonetheless a plethora of methods for validating models – and a large number of academic articles and books have been written describing how to do it. However, techniques like Cross Validation (the model is calibrated with a subset of the available data and the model is then tested on its ability to reproduce the remaining data) and Predictive Validity (the model makes a prediction about the future and is tested on whether the prediction comes to fruition) are often not applicable to complex long-term models like those studying HIV epidemiology. This

does not mean that models of HIV cannot be validated – it means that the stamp of validation will likely be more subjective and not involve a formal p-value from a goodness-of-fit test. Modellers must decide which dimensions are most likely driving the processes and determine the best way to show that their model captures those dimensions. For example, a model interested in the effect of age-mixing on HIV incidence will need to show that it is able to reasonably reproduce metrics like age-specific sexual activity and HIV prevalence. However, it would not be unreasonable to omit processes related to random biological variation in HIV infectiousness that is not associated with age or gender. This means that it's important to clearly link research question, model design, and validity checks to achieve high quality, meaningful models.

In our dynamic sexual network models we claim validity by showing that they can produce a sexual network that is approximately similar to the real-world sexual network: we compare prevalence of age-disparate relationships across different age groups and sexes; we compare the frequency with which individuals form multiple concurrent relationships; we compare the duration of relationships and the time between relationships. In short, we compare our simulated sexual network to a real world sexual network with statistics that are known to be important in the epidemiology of HIV. Hence our simulation is able to produce a facsimile of a real world sexual network.

APPENDIX C. RECRUITING STRATGIES SENSITIVITY ANALYSIS

In order to understand the simulation's sensitivity to different recruiting strategies we ran 100 simulations of 4 different scenarios with default parameter values: (1) optimized recruiting which recruits those agents that have been waiting the longest, and has queues that do not resort when a similar suitor (from the same queue as the previous suitor) is being matched; (2) random agent recruiting, which pulls agents randomly from their queue (as opposed to pulling the agent that has been waiting the longest); (3) constant resorting, which resorts the queue for ever suitor (as opposed to caching a suitor and recycling accept/reject decisions; (4) queue length recruiting, which recruits from queues probabilistically based on their length. We compare summary statistics of the simulation runs to summary statistics from South Africa's Sexual Behavioural Survey [2] in A12. The figure shows that none of the different recruiting strategies produces significantly different summary statistics about the underlying network.

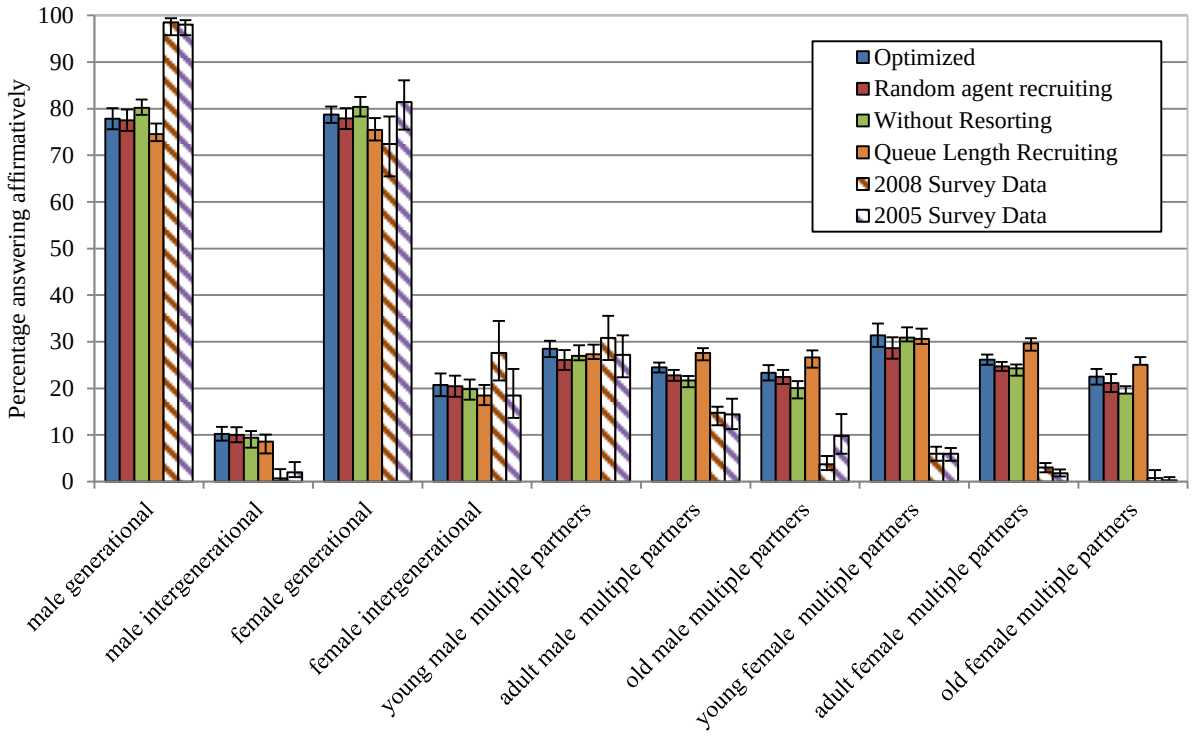

Figure C1: A comparison of simulation output metrics to survey data under four different scenarios: (blue) the default optimized algorithm which does not resort if a suitor is similar to the previous suitor and recruits agents from queues with a first-in-first-out (FIFO) strategy; (red) modified algorithm which recruits agents randomly from queues instead of FIFO; (green) modified algorithm which resorts the queue with every suitor; (orange) modified algorithm in which queues with more agents are more likely to be recruited from. The simulation output is similar to the survey data for each of the algorithms, but the optimized version runs significantly faster than the others.

APPENDIX D. COMMUNICATION OVERHEAD ANALYSIS

We investigated whether additional speed-up could be obtained by packing highly connected MPI processes (in terms of amount of communication) onto the same node. The experiment used four MPI processes with each sending out 100 messages (each message a random float) each time step to either an off-node partner or an on-node partner. The input variable *ratio* determined the probability of sending a single message to the off-node partner or on-node partner (for a ratio of 0, all messages are on-node; for a ratio of 1, all messages are off-node). Simulation run for 1500 time steps (the approximate number of time steps in our simulations). We recorded the amount of time the simulations required for different ratio, with each ratio repeated 10 times for consistency. The simulations were run on both the Milano and Helium clusters.

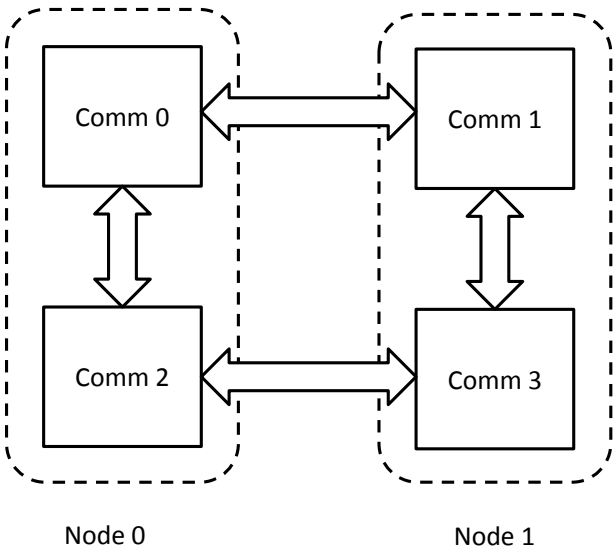

Figure D1: The set-up for an experiment to determine the necessity of packing highly communicative MPI processes on the same node.

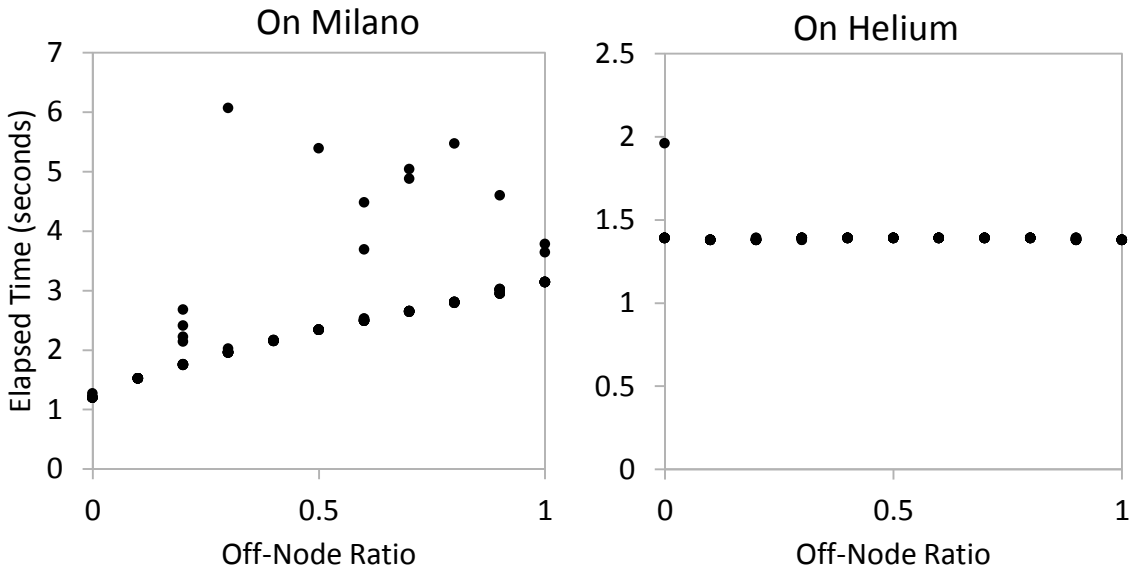

Figure D2: The amount of time required to run the simulation with different ratio values for the Milano and Helium cluster. For Milano, as the amount of off-node communication increases (goes to 1) the amount of time required to run increases linearly. There is a significant amount of noise in these values however as there are many background processes running on Milano. The amount of time required to run simulations on Helium were consistently low for all values of ratio – communication between nodes is indistinguishable from communication on nodes.

REFERENCES

1. **UNICEF - Avian and Pandemic Influenza Communication Resources - Bird flu : Communicating the risk** [http://www.unicef.org/dump/index_38356.html]

2. Shisana O, Rehle T, Simbayi LC, Zuma K, Jooste S, Pillay-van-Wyk V, Mbelle N, Van Zyl J, Parker W, Zungu N: *South African National HIV Prevalence, Incidence, Behaviour and Communication Survey, 2008: A Turning Tide among Teenagers?*. HSRC Press Cape Town; 2009.

3. UNAIDS: *Global Report 2013: UNAIDS Report on the Global AIDS Epidemic*. ebookpartnership. com; 2013.

4. Kent ME, Romanelli F: **Reexamining syphilis: an update on epidemiology, clinical manifestations, and management**. *Ann Pharmacother* 2008, **42**:226–236.

5. Woods CR: **Congenital syphilis-persisting pestilence**. *Pediatr Infect Dis J* 2009, **28**:536–537.

6. Wawer MJ, Gray RH, Sewankambo NK, Serwadda D, Li X, Laeyendecker O, Kiwanuka N, Kigozi G, Kiddugavu M, Lutalo T, Nalugoda F, Wabwire-Mangen F, Meehan MP, Quinn TC: **Rates of HIV-1 transmission per coital act, by stage of HIV-1 infection, in Rakai, Uganda**. *J Infect Dis* 2005, **191**:1403–1409.

7. Engel J: *Epidemic, The*. 1 edition. New York: Smithsonia; 2006.

8. Richman DD, Little SJ, Smith DM, Wrin T, Petropoulos C, Wong JK: **HIV evolution and escape**. *Trans Am Clin Climatol Assoc* 2004, **115**:289–303.

9. Iliffe J: *The African AIDS Epidemic: A History*. 1 edition. Athens : Oxford : Cape Town, South Africa: Ohio University Press; 2006.

10. Gould P: *The Slow Plague: A Geography of the AIDS Pandemic*. Oxford, UK; Cambridge, USA: Blackwell Publishers; 1993.

11. Quammen D: *Spillover: Animal Infections and the Next Human Pandemic*. 1 edition. W. W. Norton & Company; 2012.

12. Lewin SR, Rouzioux C: **HIV cure and eradication: how will we get from the laboratory to effective clinical trials?:**. *AIDS* 2011, **25**:885–897.

13. Shehu-Xhilaga M, Rhodes D, Wightman F, Liu HB, Solomon A, Saleh S, Dear AE, Cameron PU, Lewin SR: **The novel histone deacetylase inhibitors metacept-1 and metacept-3 potently increase HIV-1 transcription in latently infected cells:**. *AIDS* 2009, **23**:2047–2050.

14. Oxman GL, Smolkowski K, Noell J: **Mathematical modeling of epidemic syphilis transmission. Implications for syphilis control programs**. *Sex Transm Dis* 1996, **23**:30–39.

15. **Why is Syphilis Still Sensitive to Penicillin? | Clinical Correlations**. .

16. Donnell D, Baeten JM, Kiarie J, Thomas KK, Stevens W, Cohen CR, McIntyre J, Lingappa JR, Celum C: **Heterosexual HIV-1 transmission after initiation of antiretroviral therapy: a prospective cohort analysis**. *The Lancet* 2010, **375**:2092–2098.

17. Fleming DT, Wasserheit JN: **From epidemiological synergy to public health policy and practice: the contribution of other sexually transmitted diseases to sexual transmission of HIV infection**. *Sex Transm Infect* 1999, **75**:3–17.

18. Padayatchi N, Naidoo K, Dawood H, Kharsany ABM, Abdool Karim Q: **A review of progress on HIV, AIDS and Tuberculosis.** 2010.

19. Buvé A, Weiss HA, Laga M, Van Dyck E, Musonda R, Zekeng L, Kahindo M, Anagonou S, Morison L, Robinson NJ, Hayes RJ, Study Group on Heterogeneity of HIV Epidemics in African Cities: **The epidemiology of gonorrhoea, chlamydial infection and syphilis in four African cities**. *AIDS Lond Engl* 2001, **15 Suppl 4**:S79–88.

20. Mahy M, Stover J, Kiragu K, Hayashi C, Akwara P, Luo C, Stanecki K, Ekpini R, Shaffer N: **What will it take to achieve virtual elimination of mother-to-child transmission of HIV? An assessment of current progress and future needs**. *Sex Transm Infect* 2010, **86**(Suppl 2):ii48–ii55.

21. Fenton L: **Preventing HIV/AIDS through poverty reduction: the only sustainable solution?**. *The Lancet* 2004, **364**:1186–1187.

22. Lurie MN, Williams BG, Zuma K, Mkaya-Mwamburi D, Garnett GP, Sturm AW, Sweat MD, Gittelsohn J, Abdool Karim SS: **The Impact of Migration on HIV-1 Transmission in South Africa: A Study of Migrant and Nonmigrant Men and Their Partners**. *Sex Transm Dis* 2003, **30**:149–156.

23. Williams B: *Spaces of Vulnerability: Migration and HIV/AIDS in South Africa*. Idasa; 2002.

24. Watt MH, Aunon FM, Skinner D, Sikkema KJ, Kalichman SC, Pieterse D: **"Because he has bought for her, he wants to sleep with her": alcohol as a currency for sexual exchange in South African drinking venues**. *Soc Sci Med 1982* 2012, **74**:1005–1012.

25. Townsend L, Ragnarsson A, Mathews C, Johnston LG, Ekström AM, Thorson A, Chopra M: **"Taking care of business": alcohol as currency in transactional sexual relationships among players in Cape Town, South Africa**. *Qual Health Res* 2011, **21**:41–50.

26. Kristof ND, WuDunn S: *Half the Sky: Turning Oppression into Opportunity for Women Worldwide*. Reprint edition. New York: Vintage; 2010.

27. Gloyd S, Chai S, Mercer MA: **Antenatal syphilis in sub-Saharan Africa: missed opportunities for mortality reduction**. *Health Policy Plan* 2001, **16**:29–34.

28. **WHO | The use of rapid syphilis tests**
[http://www.who.int/reproductivehealth/publications/rtis/TDR_SDI_06_1/en/]

29. **The World Factbook 2013-14** [https://www.cia.gov/library/publications/the-world-factbook/index.html]

30. West B, Walraven G, Morison L, Brouwers J, Bailey R: **Performance of the rapid plasma reagin and the rapid syphilis screening tests in the diagnosis of syphilis in field conditions in rural Africa**. *Sex Transm Infect* 2002, **78**:282–285.

31. Lawrence J, Miner E, McInroy M: **Maps of Syphilis in Africa**. 2011.

32. Kleutsch L, Harvey S, Rennie W: **Rapid syphilis tests in Tanzania: A long road to adoption**. 2009.

33. **WHO | The global elimination of congenital syphilis: rationale and strategy for action** [http://www.who.int/reproductivehealth/publications/rtis/9789241595858/en/]

34. Vickerman P, Peeling RW, Terris-Prestholt F, Changalucha J, Mabey D, Watson-Jones D, Watts C: **Modelling the cost-effectiveness of introducing rapid syphilis tests into an antenatal syphilis screening programme in Mwanza, Tanzania**. *Sex Transm Infect* 2006, **82 Suppl 5**:v38–43.

35. Mabey D: **Interactions between HIV infection and other sexually transmitted diseases**. *Trop Med Int Health TM IH* 2000, **5**:A32–36.

36. Orubuloye IO, Caldwell P, Caldwell JC: **The Role of High-Risk Occupations in the Spread of AIDS: Truck Drivers and Itinerant Market Women in Nigeria**. *Int Fam Plan Perspect* 1993, **19**:43–71.

37. Bwayo JJ, Omari AM, Mutere AN, Jaoko W, Sekkade-Kigondu C, Kreiss J, Plummer FA: **Long distance truck-drivers: 1. Prevalence of sexually transmitted diseases (STDs)**. *East Afr Med J* 1991, **68**:425–429.

38. Livia Montana, Melissa Neuman, Vinod Mishra: *Spatial Modeling of HIV Prevalence in Kenya*. 2007.

39. Aron JL, Schwartz IB: **Seasonality and period-doubling bifurcations in an epidemic model**. *J Theor Biol* 1984, **110**:665–679.

40. Ludkovski M, Niemi J: **Optimal disease outbreak decisions using stochastic simulation**. In *Simul Conf WSC Proc 2011 Winter*; 2011:3844–3853.

41. Cohen MS, Chen YQ, McCauley M, Gamble T, Hosseinipour MC, Kumarasamy N, Hakim JG, Kumwenda J, Grinsztejn B, Pilotto JH: **Prevention of HIV-1 infection with early antiretroviral therapy**. *N Engl J Med* 2011, **365**:493–505.

42. Brandeau ML, Zaric GS: **Optimal investment in HIV prevention programs: more is not always better**. *Health Care Manag Sci* 2009, **12**:27–37.

43. Zaric GS, Brandeau ML: **Optimal investment in a portfolio of HIV prevention programs**. *Med Decis Mak Int J Soc Med Decis Mak* 2001, **21**:391–408.

44. Kleinberg J: *Algorithm Design*. 1 edition. Boston: Addison-Wesley; 2005.

45. Halloran ME, Ferguson NM, Eubank S, Longini IM, Cummings DAT, Lewis B, Xu S, Fraser C, Vullikanti A, Germann TC, Wagener D, Beckman R, Kadau K, Barrett C, Macken CA, Burke DS, Cooley P: **Modeling targeted layered containment of an influenza pandemic in the United States**. *Proc Natl Acad Sci* 2008, **105**:4639–4644.

46. Ferguson NM, Cummings DAT, Cauchemez S, Fraser C, Riley S, Meeyai A, Iamsirithaworn S, Burke DS: **Strategies for containing an emerging influenza pandemic in Southeast Asia**. *Nature* 2005, **437**:209–214.

47. Bisset KR, Chen J, Feng X, Kumar VSA, Marathe MV: **EpiFast: a fast algorithm for large scale realistic epidemic simulations on distributed memory systems**. In *Proc 23rd Int Conf Supercomput*. New York, NY, USA: ACM; 2009:430–439. [*ICS '09*]

48. Barrett CL, Bisset KR, Eubank SG, Feng X, Marathe MV: **EpiSimdemics: an efficient algorithm for simulating the spread of infectious disease over large realistic social networks**. In *Proc 2008 ACMIEEE Conf Supercomput*; 2008:37.

49. Grefenstette JJ, Brown ST, Rosenfeld R, DePasse J, Stone NT, Cooley PC, Wheaton WD, Fyshe A, Galloway DD, Sriram A, Guclu H, Abraham T, Burke DS: **FRED (A Framework for Reconstructing Epidemic Dynamics): an open-source software system for modeling infectious diseases and control strategies using census-based populations**. *BMC Public Health* 2013, **13**:940.

50. Van der Ploeg CP, Van Vliet C, De Vlas SJ, Ndinya-Achola JO, Fransen L, Van Oortmarssen GJ, Habbema JDF: **STDSIM: A microsimulation model for decision support in STD control**. *Interfaces* 1998, **28**:84–100.

51. Korenromp EL, Van Vliet C, Grosskurth H, Gavyole A, Van der Ploeg CP, Fransen L, Hayes RJ, Habbema JDF: **Model-based evaluation of single-round mass treatment of sexually transmitted diseases for HIV control in a rural African population**. *Aids* 2000, **14**:573–593.

52. Korenromp EL, van Vliet C, Bakker R, de Vlas SJ, Habbema JDF: **HIV spread and partnership reduction for different patterns of sexual behaviour - a study with the microsimulation model *STDSIM***. *Math Popul Stud* 2000, **8**:135–173.

53. Van Vliet C, Meester EI, Korenromp EL, Singer B, Bakker R, Habbema JDF: **Focusing strategies of condom use against HIV in different behavioural settings: an evaluation based on a simulation model**. *Bull World Health Organ* 2001, **79**:442–454.

54. Clark SJ, Eaton JW, Elmquist MM, Ottenweiller NR, Snavely JK: **Demographic consequences of HIV epidemics and effects of different male circumcision intervention designs: Suggestive findings from microsimulation**. *Cent Stat Soc Sci* 2008.

55. **Simulating the Control of a Heterosexual HIV Epidemic in a Severely Affected East African City** [http://pubsonline.informs.org/doi/abs/10.1287/inte.28.3.101]

56. Mei S, Sloot PMA, Quax R, Zhu Y, Wang W: **Complex agent networks explaining the HIV epidemic among homosexual men in Amsterdam**. *Math Comput Simul* 2010, **80**:1018–1030.

57. Beauclair R, Kassanjee R, Temmerman M, Welte A, Delva W: **Age-disparate relationships and implications for STI transmission among young adults in Cape Town, South Africa**. *Eur J Contracept Reprod Health Care* 2012, **17**:30–39.

58. Hawkins K, Price N, Mussá F: **Milking the cow: Young women's construction of identity and risk in age-disparate transactional sexual relationships in Maputo, Mozambique**. *Glob Public Health* 2009, **4**:169–182.

59. Leclerc-Madlala S: **Age-disparate and intergenerational sex in southern Africa: the dynamics of hypervulnerability**. *Aids* 2008, **22**:S17–S25.

60. **Concurrent partnerships and the spread of HIV : AIDS** [http://journals.lww.com/aidsonline/Fulltext/1997/05000/Concurrent_partnerships_and_the_spread_of_HIV.12.aspx]

61. W P, B M, P N, C C: **Concurrent sexual partnerships amongst young adults in South Africa. Challenges for HIV prevention communication.** .

62. Jewkes R, Sikweyiya Y, Morrell R, Dunkle K: **The Relationship between Intimate Partner Violence, Rape and HIV amongst South African Men: A Cross-Sectional Study**. *PLoS ONE* 2011, **6**:e24256.

63. Luke S, Cioffi-Revilla C, Panait L, Sullivan K: **Mason: A new multi-agent simulation toolkit**. In *Proc 2004 SwarmFest Workshop*. *Volume 8*; 2004.

64. Pitpitan EV, Kalichman SC, Eaton LA, Cain D, Sikkema KJ, Skinner D, Watt MH, Pieterse D: **Gender-based violence, alcohol use, and sexual risk among female patrons of drinking venues in Cape Town, South Africa**. *J Behav Med* 2013, **36**:295–304.

65. Delva W, Beauclair R, Welte A, Vansteelandt S, Hens N, Aerts M, Toit E du, Beyers N, Temmerman M: **Age-disparity, sexual connectedness and HIV infection in disadvantaged communities around Cape Town, South Africa: a study protocol**. *BMC Public Health* 2011, **11**:616.

66. Holmes KK, Levine R, Weaver M: **Effectiveness of condoms in preventing sexually transmitted infections**. *Bull World Health Organ* 2004, **82**:454–461.

67. Weller S, Davis K: **Condom effectiveness in reducing heterosexual HIV transmission**. *Cochrane Database Syst Rev* 2002, **1**.

68. Kurth AE, Celum C, Baeten JM, Vermund SH, Wasserheit JN: **Combination HIV prevention: significance, challenges, and opportunities**. *Curr HIV/AIDS Rep* 2011, **8**:62–72.

69. Van Dijk D, Sloot PMA, Tay JC, Schut MC: **Individual-based simulation of sexual selection: A quantitative genetic approach**. *Procedia Comput Sci* 2010, **1**:2003–2011. [*ICCS 2010*]

70. Anderson DF: **A modified next reaction method for simulating chemical systems with time dependent propensities and delays**. *J Chem Phys* 2007, **127**:214107.

71. Gillespie DT: **Exact stochastic simulation of coupled chemical reactions**. *J Phys Chem* 1977, **81**:2340–2361.

72. Delva W, Wilson DP, Abu-Raddad L, Gorgens M, Wilson D, Hallett TB, Welte A: **HIV Treatment as Prevention: Principles of Good HIV Epidemiology Modelling for Public Health Decision-Making in All Modes of Prevention and Evaluation**. *PLoS Med* 2012, **9**:e1001239.

73. Grimm V, Berger U, Bastiansen F, Eliassen S, Ginot V, Giske J, Goss-Custard J, Grand T, Heinz SK, Huse G, Huth A, Jepsen JU, Jørgensen C, Mooij WM, Müller B, Pe'er G, Piou C, Railsback SF, Robbins AM, Robbins MM, Rossmanith E, Rüger N, Strand E, Souissi S, Stillman RA, Vabø R, Visser U, DeAngelis DL: **A standard protocol for describing individual-based and agent-based models**. *Ecol Model* 2006, **198**:115–126.

74. Weiss HA, Quigley MA, Hayes RJ: **Male circumcision and risk of HIV infection in sub-Saharan Africa: a systematic review and meta-analysis**. *AIDS Lond Engl* 2000, **14**:2361–2370.

75. Kahn JG, Marseille E, Auvert B: **Cost-Effectiveness of Male Circumcision for HIV Prevention in a South African Setting**. *PLoS Med* 2006, **3**:e517.

76. Rosen S, Long L, Sanne I: **The outcomes and outpatient costs of different models of antiretroviral treatment delivery in South Africa**. *Trop Med Int Health TM IH* 2008, **13**:1005–1015.

77. **Statistics South Africa | The South Africa I Know, The Home I Understand**. .

78. Bedimo AL, Pinkerton SD, Cohen DA, Gray B, Farley TA: **Condom distribution: a cost-utility analysis**. *Int J STD AIDS* 2002, **13**:384–392.

79. Gray RH, Kiwanuka N, Quinn TC, Sewankambo NK, Serwadda D, Mangen FW, Lutalo T, Nalugoda F, Kelly R, Meehan M, Chen MZ, Li C, Wawer MJ: **Male circumcision and HIV acquisition and transmission: cohort studies in Rakai, Uganda. Rakai Project Team**. *AIDS Lond Engl* 2000, **14**:2371–2381.

80. Abuelezam NN, Rough K, Seage III GR: **Individual-Based Simulation Models of HIV Transmission: Reporting Quality and Recommendations**. *PLoS ONE* 2013, **8**:e75624.

81. Ghani AC, Garnett GP: **Risks of acquiring and transmitting sexually transmitted diseases in sexual partner networks**. *Sex Transm Dis* 2000, **27**:579–587.

82. Ghani AC, Ison CA, Ward H, Garnett GP, Bell G, Kinghorn GR, Weber J, Day S: **Sexual partner networks in the transmission of sexually transmitted diseases. An analysis of gonorrhea cases in Sheffield, UK.** *Sex Transm Dis* 1996, **23**:498.

83. Boily M-C, Baggaley RF, Wang L, Masse B, White RG, Hayes RJ, Alary M: **Heterosexual risk of HIV-1 infection per sexual act: systematic review and meta-analysis of observational studies**. *Lancet Infect Dis* 2009, **9**:118–129.

84. Jones E, Oliphant T, Peterson P: *SciPy: Open Source Scientific Tools for Python*. 2001.

85. Hagberg A, Schult D, Swart P: *NetworkX*. 2004.

86. Hunter JD: **Matplotlib: A 2D graphics environment**. *Comput Sci Eng* 2007, **9**:90–95.

87. Rubin DB: **Bayesianly Justifiable and Relevant Frequency Calculations for the Applies Statistician**. *Ann Stat* 1984, **12**:1151–1172.

88. Diggle PJ, Gratton RJ: **Monte Carlo Methods of Inference for Implicit Statistical Models**. *J R Stat Soc Ser B Methodol* 1984, **46**:193–227.

89. Wertheim JO, Leigh Brown AJ, Hepler NL, Mehta SR, Richman DD, Smith DM, Kosakovsky Pond SL: **The global transmission network of HIV-1**. *J Infect Dis* 2014, **209**:304–313.

90. Pennings PS, Holmes SP, Shafer RW: **HIV-1 Transmission Networks in a Small World**. *J Infect Dis* 2013:jit525.

91. Tolentino SL, Meng F, Delva W: **A Simulation-based Method for Efficient Resource Allocation of Combination HIV Prevention**. In *Proc 6th Int ICST Conf Simul Tools Tech*. ICST, Brussels, Belgium, Belgium: ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering); 2013:31–40. [*SimuTools '13*]

92. Bershteyn A, Klein DJ, Wenger E, Eckhoff PA: **Description of the EMOD-HIV Model v0.7**. *ArXiv Prepr ArXiv12063720* 2012.

93. McCormick AW, Abuelezam NN, Rhode ER, Hou T, Walensky RP, Pei PP, Becker JE, DiLorenzo MA, Losina E, Freedberg KA, Lipsitch M, Seage GR III: **Development, Calibration and Performance of an HIV Transmission Model Incorporating Natural History and Behavioral Patterns: Application in South Africa**. *PLoS ONE* 2014, **9**:e98272.

94. Butler AR, Hallett TB: **Migration and the Transmission of STIs**. In *New Public Health STDHIV Prev*. Edited by Aral SO, Fenton KA, Lipshutz JA. Springer New York; 2013:65–75.

95. Burton J, Billings L, Cummings DAT, Schwartz IB: **Disease persistence in epidemiological models: The interplay between vaccination and migration**. *Math Biosci* 2012, **239**:91–96.

96. Magis-Rodríguez C, Gayet C, Negroni M, Leyva R, Bravo-García E, Uribe P, Bronfman M: **Migration and AIDS in Mexico: an overview based on recent evidence**. *J Acquir Immune Defic Syndr 1999* 2004, **37 Suppl 4**:S215–226.

97. Hirsch JS: **Labor migration, externalities and ethics: Theorizing the meso-level determinants of HIV vulnerability**. *Soc Sci Med* 2014, **100**:38–45.

98. Coffee M, Lurie MN, Garnett GP: **Modelling the impact of migration on the HIV epidemic in South Africa**. *AIDS Lond Engl* 2007, **21**:343–350.

99. Lurie M, Harrison A, Wilkinson D, Karim SA: **Circular migration and sexual networking in rural KwaZulu/Natal: implications for the spread of HIV and other sexually transmitted diseases**. *Health Transit Rev* 1997, **7**:17–27.

100. Lurie MN, Williams BG, Zuma K, Mkaya-Mwamburi D, Garnett GP, Sweat MD, Gittelsohn J, Karim SSA: **Who infects whom? HIV-1 concordance and discordance among migrant and non-migrant couples in South Africa**. *AIDS Lond Engl* 2003, **17**:2245–2252.

101. Kniveton D, Smith C, Wood S: **Agent-based model simulations of future changes in migration flows for Burkina Faso**. *Glob Environ Change* 2011, **21, Supplement 1**:S34–S40. [*Migration and Global Environmental Change – Review of Drivers of Migration*]

102. Silveira JJ, Espíndola AL, Penna TJP: **Agent-based model to rural–urban migration analysis**. *Phys Stat Mech Its Appl* 2006, **364**(C):445–456.