Sexually Transmitted Infections, Sexual Behavior Change and the HIV/AIDS Epidemic

Emily Oster *
Harvard University
Draft: Preliminary and Incomplete

June 8, 2004

Abstract

Forty million people are infected with HIV; twenty-nine million of them are in Sub-Saharan Africa. Although the HIV pandemic has touched every country worldwide, there is enormous and largely unexplained variation in HIV rate across continents, and across countries within Africa. I first present existing evidence from the medical literature that untreated sexually transmitted infections and their associated genital ulcers dramatically increase HIV transmission rates. I then embed this into a model of sexual behavior that predicts HIV rates across countries using survey data. The model provides a remarkably good fit to the cross country picture of the epidemic. It indicates that differences across continents are driven by differences in transmission rates of the virus, but that differences across countries within Africa can be fully attributed to differences in risky sexual behavior and epidemic timing. The model is used to simulate interventions and evaluate their cost-effectiveness. I find interventions that treat non-HIV sexually transmitted infections are more cost effective than those designed to change sexual behavior, and cost as little $3.80 per life year. The results suggest that had 2002 expenditure on HIV interventions been optimally spent, over 40% of new infections could have been prevented.

*I am grateful to Nava Ashraf, Connie Celum, Edward Glaeser, Larry Katz, Michael Kremer, Ulla Larson, Steve Levitt, Ajay Mahal, Jesse Shapiro, Andrei Shleifer, Larry Summers, Rebecca Thornton and participants in the Harvard Development Workshop for helpful comments and discussion

1 Introduction

In 2002, 2.4 million people in Sub-Saharan Africa died of AIDS. There were 3.5 million new infections, bringing the total number of infections to 29.4 million, a prevalence rate of around 10%. The literature on the HIV epidemic, however, has not been able to explain why there is so much variation in HIV rate across countries and continents. There exist a number of possible explanations – differences in sexual behavior (either number of partners or type of partnerships), differences in transmission rates (either due to untreated sexually transmitted infections (STIs) or differences in circumcision rates), differences in safety of blood transfusions and others. However, virtually no attempt has been made to quantify differences in these parameters or to understand whether these differences are sufficient to explain the variation in HIV rates. Understanding which of these possibilities – if any – explain the variation across countries and continents is key to identifying what type of policy interventions will curb the epidemic.

This paper addresses this puzzle, and offers a model that provides a remarkable fit to the cross country variation in HIV rates, suggesting the parameters of the model are sufficient to explain this variation. I then use this model to identify the ideal policy interventions.

The paper first presents existing evidence from experiments in the medical literature showing that the transmission rate of HIV is higher for individuals who have other untreated STIs, particularly those that cause open genital sores. Given that treatment levels for STIs are much lower in the developing world, this may explain much of the variation across continents. In fact, in section 3 I show in a very simple, stylized model of sexual behavior that relatively minor differences in transmission rates can produce large differences in HIV rate.

Cross-continent transmission rate estimates from the literature are then embedded in a more complex model of sexual behavior that incorporates differences in behavior across ages, gender and marital status, and introduces the presence of female commercial sex workers (FSWs). This model is calibrated with cross-country survey data on sexual behavior from 14 countries in Sub-Saharan Africa, and the United States. I find that the model is an extremely good fit to the data on HIV rates and can explain both the cross-continent and cross-country variation. It is worth noting here that the model is *not* fitted to the existing HIV rates;

rather, sexual behavior parameters from survey data are taken as inputs and the predicted HIV rates from the model are compared with actual HIV prevalence.

The results indicate that the cross-continent variation – differences between Sub-Saharan Africa and the developed world – can be fully attributed to differences in viral transmission rates per unprotected sexual partnership. However, cross-country variation within Africa can be attributed to differences in risky sexual behavior and the timing of the epidemic.

This result has two primary implications. First, the cross-country results from within Africa may be helpful for understanding the effect of other demographics on the HIV epidemic. Other literature has discussed the effect of variables such as income, education and inequality on the HIV rate(Bonnel(2000), World Bank(1997), World Bank (2000)). The results here demonstrate that any variation arising from these variables must be due to their effect on sexual behavior, which makes calibrating the effect of changes in these underlying variables more tractable.

The second implication is that huge strides can be made in prevention of the HIV epidemic using existing, off-patent drugs that are readily available. Section 5 uses the model in the paper to simulate the effect of a number of interventions and evaluate their effectiveness and cost-effectiveness. The results indicate that treating untreated bacterial STIs (the most cost effective intervention) could prevent as many as 24% of new infections over the next decade, at a cost of less than $75 per infection, or around $3.80 per life year. Even greater strides could be made (albeit at higher cost) by incorporating herpes suppressive therapy as well. Interventions to decrease sexual behavior are less effective, although they may have a role as well.

All of these interventions are dramatically more cost-effective than anti-retroviral therapy, which currently costs around $1100 per life year. Even if anti-retroviral drugs were taken off patent it is difficult to imagine they could be anywhere near as cost-effective as treatment of bacterial STIs, simply because the marginal cost is not zero, and the drugs must be taken every day for the rest of the individual's life. Although some policy-makers have focused on anti-retroviral provision as the duty of rich countries and pharmaceutical companies (see, for example, Sachs(2000, 2004)) the results here indicate that this type of

treatment should be far from the first line of defense.[1]

These results have enormous implications for HIV policy in Africa. The drugs required to treat bacterial STIs, and even to treat herpes, are extremely inexpensive. This type of intervention is attainable without any concession from the pharmaceutical companies, without any lobbying and without requiring individuals to be tested for HIV. Further, it is unlikely to be difficult to convince people to have their STIs treated, because the experience of having an untreated STI is unpleasant even without the increased risk of HIV transmission.

The rest of the paper is structured as follows. Section two provides some background on HIV and the HIV/AIDS epidemic and discusses the evidence on untreated STIs and HIV transmission rates. Section three presents a very simple model of sexual interactions that demonstrates the importance of different per-partnership transmission rates, and suggests intuition about the differences between HIV rates in the United States and Sub-Saharan Africa. Section four presents the more complex model of sexual interactions and calibrates it using actual data on sexual behavior and epidemic timing across countries in Sub-Saharan Africa and the United States. Section 5 uses the model for forecasting and evaluates two types of interventions – those designed to treat STIs and those designed to alter sexual behavior. Section six discusses robustness issues. Section seven concludes and discusses optimal policy interventions.

2 Background on HIV

The Human Immunodeficiency Virus (HIV) is a retro-virus that attacks the immune system of the host individual. The virus slowly invades and kills T-cells. As the disease progresses, individuals become increasingly susceptible to other illnesses. Eventually, the compromised immune system will lead to death through another proximate cause. An individual is said to have Acquired Immune Deficiency Syndrome (AIDS) once their immune system has been severely compromised.

Immediately after infection, there is a period of three weeks to a month when the

[1]The work on this often cites the fact that some antiretrovirals can now be had for less than $1 per day in drug costs. It is worth noting that even the most expensive intervention evaluated here – suppression of herpes and treatment of bacterial STIs – would result in less than 10 *cents* per day in drug costs.

patient will experience a brief flu-like illness (tiredness, fever, etc) after which they will be asymptomatic for, typically, seven to ten years. For the last year of life, when the individual has AIDS, they will normally be in and out of the hospital with illnesses that have taken advantage of the compromised immune system.

HIV is spread through a number of channels. The most common are sexual. HIV can be spread through heterosexual or homosexual encounters. The other major type of transmission is vertical – from mother to child – either in the womb, during birth, or while breastfeeding. HIV can also be spread through sharing needles (either by intravenous drug users, or poor hygiene in hospitals) and through transfusions with infected blood. The efficiency of these transmission mechanisms varies. Infection rates are higher for anal than vaginal sex, higher still for mother-to-child transmission, and extremely high (essentially 100%) for transmission with infected blood.

HIV was first identified in the gay community in the United States in the early 1980s when doctors noticed an increase in Kaposi's Sarcoma (an otherwise quite rare form of cancer) among young, gay men. It subsequently became clear that the disease was being spread around the homosexual bathhouses in San Francisco. The origin of the virus, however, is generally thought to be in Africa, probably from the region around Lake Victoria. HIV appears to have evolved from the Simian form of the disease, Simian Immunodeficiency Virus (SIV-I) between 120 and 160 years ago. Researchers have not, however, been able to identify when the virus jumped to humans, how, and why it appears to have taken so long for it to become an epidemic. Current consensus suggests that the virus probably jumped species through human consumption of raw, infected monkey meat, although a number of other methods have been suggested (a fascinating discussion of this issue is contained in Hooper(1999)).

Although the virus was first identified in the United States, the developing world, and in particular Africa, has unquestionably borne the majority of the infection burden. In the U.S., the virus has largely been limited to the homosexual community and intravenous drug users, and condom promotion and needle exchange have kept the infection rate relatively low in those groups, as well. The only group with increasing incidence in the U.S. is minority women. Europe has been even less affected than the United States.

The story is quite different in the worst-affected regions, particularly Sub-Saharan Africa. The vast majority of the 40 million people with HIV live in Sub-Saharan Africa (about 29.4 million of them). The continent has an average infection rate of 10%, although this masks a huge amount of variance among countries. Table 1 shows the five most infected countries in the region and the five least affected. In general, infection rates are highest in Eastern and Southern Africa, and lower in the West and the Sahara region. It is worth noting that estimates of HIV prevalence are not perfect, and are primarily drawn from testing of women at antenatal clinics. For obvious reasons (pregnant women are definitely having unprotected sex) some have argued that these rates may be overestimated, and there has been some recent work arguing that the true population rates are lower. For the purpose of this paper, we focus only on pregnant women and match the predicted rates for that group with the actual rates from the data.

Drugs that dramatically slow the progression of HIV have become available in recent years. These drugs have prolonged countless lives in the developed world, but use of these regimens in the developing world is rare, due both to the cost of the (on-patent) drugs and the difficulty of administering the drugs on a continent with so few doctors. Interventions in Africa have focused more on prevention and treatment of opportunistic infections. One of the most successful of these has been the prevention of mother-to-child transmission by treating infected women with antiretrovirals during pregnancy, doing elective caesarian sections and avoiding breastfeeding. Evidence suggests that these protocols could reduce the rate of mother-to-child transmission to 2% (from a baseline of around 30%), although they have only been partially adopted in the developing world (Thorne and Newell, 2003). Preventing horizontal (sexual) transmission has proved less tractable. Some have looked to the success of Uganda and argued for interventions designed to decrease risky sexual behavior. Others have pointed to evidence that having untreated STIs makes the chance of infection much higher and have argued that treating these diseases is a better form of intervention.

It is clear why decreasing the share of the population engaging in risky sexual behavior could slow or halt the spread of the HIV epidemic. The relationship between untreated STIs and transmission of HIV, however, may be less obvious. For the purposes of this paper heterosexual transmission rates – and how they differ across continents – are a vital

issue.

The first thing to note in a discussion of heterosexual HIV transmission rates is that there are no widely accepted estimates of these rates, particularly in the developing world. In the developed world, the most widely cited numbers are a probability of 1 in 1000 per sexual contact male-to-female and 1 in 300 female-to-male. This, taken along with other studies in the developed world, suggests a per-partnership transmission rate of about 10% male-to-female and 3% female-to-male (Ragni et al, 1989; Rockstroh et al, 1995; Kamradt et al, 1990; Smiley et al, 1988; Di Vincenzi, 1994). Throughout this paper we use per-partnership rates rather than per-sexual contact rates of transmission, both because per-partnership rates are easier to estimate and because per-contact rates may be a less accurate measure.[2]

In the developing world, particularly in Sub-Saharan Africa, a number of studies have, either directly or indirectly, attempted to estimate transmission rates, and these estimates vary widely. Table 2 reports estimates of transmission rates for four studies (Appendix A details how these were calculated). It is clear that these rates appear to be somewhat higher in Sub-Saharan Africa than the 10%/3% in the United States and elsewhere in the developed world. Throughout the rest of the paper we use a weighted average of these transmission rates for Sub-Saharan Africa: 27% male-to-female and 12% female-to-male. We will use a slightly lower transmission rate for male contact with FSWs, because these partnerships are of much more limited frequency (Cameron et al, 1989).

While informative and vital to calibration exercises, these empirically estimated differences do not provide insight into how transmission rates might be affected by interventions. This requires accounting for the underlying cause of different transmission rates. There are a number of reasons why these transmission differences exist. The most widely supported, argument is that untreated STIs (other than HIV) increase the chance of

[2]There exists good evidence that the relationship between HIV transmission in a partnership and number of sexual contacts with the partner may be highly nonlinear (Kaplan, 1990). For example, studies of infected hemophiliacs and their uninfected partners suggested that the length of partnership had no affect on probability of infection (Ragni et al, 1989; Downs and Di Vincenzi, 1996). In addition, there is evidence that reporting using condoms "sometimes" provides no more protection than reporting no use, suggesting that lowering the number of unprotected sexual acts has little effect, assuming that there are at least some unprotected acts (Di Vincenzi, 1994). Both of these pieces of evidence support a "lock-and-key" interpretation of HIV transmission – infection either happens within a partnership or not. This supports the use of per-partnership rather than per-act transmission rates

HIV transmission (for a general review of this, see Kapiga and Aitken(2003)).

Untreated STIs, particularly herpes and syphilis, cause open genital sores which dramatically increase the probability of blood transmission during sex, and therefore increase the probability of HIV transmission. Most of the evidence on the STI-transmission relationship is non-experimental. Genital ulcers have been shown in a number of studies (for example, Gray et al (2001); Di Vincenzi (1994)) to be associated with higher rates of HIV transmission. Finally, non-ulcerative STI infection is also associated with a greater HIV viral load in the semen of infected men (Cohen et al, 1997). Increased viral load is associated with increased chance of infecting one's partner (Gray et al (2001); Di Vincenzi (1994)), implying that untreated STIs may increase not only the chance of any individual contracting HIV, but also the chance of infecting someone else.

There are two examples of randomized controlled trials to address this question. Grosskurth et al (1995) use a randomized controlled trial and find large decreases in HIV incidence in Tanzania when STI treatment is offered. In contrast, Wawer et al(1999) finds no effect of STI treatment on HIV in Uganda. It is difficult to understand this disparity, but it is worth noting that Uganda is in many ways a special case in that the government has instituted a number of HIV prevention initiatives. In addition, the STI treatment was provided relatively infrequently (only every 10 months) and this left time for reinfection between treatments (see Kapiga and Aitken(2003) for a more detailed discussion of why the Ugandan trial might have failed to find an effect.)

Despite the results from Uganda, the Mwanza results and the non-experimental evidence mean that the vast majority of evidence points towards an effect of STIs on HIV transmission. It is clear, also, that Sub-Saharan Africa has much higher rates of untreated STIs than the developed world. Some estimates suggest that the prevalence of untreated curable STIs in Africa is around 11.9%, while the comparable figure in the U.S. is 1.9%, and in Western Europe around 2.0%.[3] In addition, untreated ulcerative STI infections (in particular HSV-2) are much more common in Sub-Saharan Africa and other developing countries.

Although the issue of untreated STIs is clearly the most prominent in terms of explaining differences in transmission rates, there are other reasons that the transmission

[3]http://www.avert.org/STIstatistiFSWorldwide.htm

rates might be higher in the developing world. For example, infection with malaria appears to increase viral load, which would, in turn, increase the chance of infecting a partner (Corbett et al, 2002; Quinn et al, 2000; Kapiga et al, 2002).

Overall, this evidence suggests that there are large differences in transmission rates between the developed and the developing world. In addition, the results on STIs suggest a way that these differences might be greatly reduced. This type of intervention is among those considered in Section 5.

3 Analytic Model

This section presents an extremely simple model of the heterosexual HIV epidemic. The goal of this section is not to accurately estimate HIV rates, or to address the subtleties of modelling sexual behavior. Rather, this is intended to provide some intuition about the importance of transmission rates in driving the epidemic. In addition, this will give a sense of the structure of the more complex simulation model presented in Section 4. This model is similar to the type of model considered frequently in mathematical epidemiology (see, for example, Anderson and May (1987), Anderson, Gupta and Ng (1990) , Garnett and Anderson (1995)).

The world consists of a continuum of measure one of individuals, evenly split between men and women. We observe the infection status of individuals at the end of each period, and each period will have a length of one year. In each period individuals choose partner(s) at random from the population. There are no long-term partnerships – all partnerships last for a single year. We denote the infection rate for men in the population in period t as m_t and women w_t. In period 0, the infection rate in the population seeded at 1% for both men and women.

We assume that individuals choose to have either one or two sexual partners each period. In addition, we assume that individual type is not serially correlated, so having had one partner last period does not make the individual more likely to have only one this period. A share λ of men choose to have two partners, and $1 - \lambda$ have only one. Likewise, a share γ of women choose to have two partners, and $1 - \gamma$ choose to have one.

The chance of infection in a partnership is denoted β_m for men having sex with an

infected woman and β_w for women with an infected man (as discussed previously, all rates in the paper are per-partnership, rather than per sexual contact). This implies that the overall chance of infection for an uninfected woman having sex with a random man in the population is $\beta_w m_t$. For an uninfected man with a random woman the chance is $\beta_m w_t$.

The model assumes that infection is Bernoulli in partnerships (for a defense of this assumption, see Kaplan (1990)), implying that individuals who have sex with two partners in each period get two random, uncorrelated draws. The chance of infection for a woman having sex with two partners is $1-(1-\beta_w m_t)^2$, and symmetrically for men having sex with two partners. The model assumes that a share μ of the individuals of both genders who are currently infected die in each period, and that the population grows in each period at rate α through the introduction of uninfected people (for example, because younger individuals enter the sexually active population and they are not infected yet).

The parameters discussed above imply the following equations of motion for the disease in the population.

$$m_t = \frac{m_{t-1} + (1-m_{t-1})[\lambda(1-(1-\beta_m w_{t-1})^2) + (1-\lambda)(\beta_m w_{t-1})] - \mu m_{t-1}}{1-\mu m_{t-1}+\alpha} \tag{1}$$

$$w_t = \frac{w_{t-1} + (1-w_{t-1})[\gamma(1-(1-\beta_w m_{t-1})^2) + (1-\gamma)(\beta_w m_{t-1})] - \mu w_{t-1}}{1-\mu m_{t-1}+\alpha} \tag{2}$$

We are ultimately interested in two results of this model – the steady state prevalence rate and the 20-year rate for men and women at a particular β_m and β_w. Solving for the steady state yields a pair of non-linear equations that can be solved together numerically for given values of β_m, β_w, α and μ. Our primary interest here lies with the relationship between β_m, β_w and the prevalence rate, so we assume $\alpha = .03$ and $\mu = .1$ from here forward. In addition, for simplicity, we will assume a specific ratio of β_m to β_w going forward. The transmission rates for the developed world and Sub-Saharan Africa discussed previously suggest a $\beta_w : \beta_m$ ratio of 3:1 at low transmission rates and 2:1 at higher rates. These are used as endpoints and intermediate ratios are imputed.

In order to solve for the steady state it is necessary to make assumptions about λ and γ. First, in a closed society with equal shares of each gender, we must have $\lambda = \gamma$.[4] Further,

[4]Obviously, this assumption of identical behavior across genders is unrealistic. In the simulation model in

we will assume for the simulations that $\lambda = 0.25$, which roughly reflects the average share of individuals having non-marital sex in the data form Sub-Saharan Africa. This implies that 25% of individuals are having sex with more than one partner.

Figure 1 shows the relationship between male-to-female transmission rate (β_w) and both steady state prevalence and prevalence twenty years into the epidemic. Below a transmission probability of about 15% the steady state (and, obviously, the twenty-year infection rate) are zero. However, the graph climbs sharply after that. A transmission rate of 17% produces a twenty-year infection rate of about 2% for women, and a steady state of 25%, whereas with a transmission rate of 30% these numbers are 12% and 85%, respectively.

This model and figure, while obviously quite stylized, give a sense of the structure of the model in the next section. In addition, they illustrate the importance of transmission rates in the epidemic – the transmission rates estimated for the U.S. suggest a 20-year infection rate of zero, whereas those for Sub-Saharan Africa suggest an infection rate close to 11% – an enormous difference that only grows as the model moves to steady state.

4 Simulation Model

The model in section three assumes constant behavior across individuals, without differentiating behavior across age groups, or across countries. A more realistic model is necessary to explore differences across countries with different sexual behavior, epidemic timing, etc. This section presents a substantially more complex behavioral model. This model can be calibrated using actual reported data on sexual behavior across countries, as well as information on epidemic timing, condom usage, etc. The goal of this model is to use data on behavior to predict HIV rates at the current stage of the epidemic in different countries and explore how closely these match to actual HIV rates. As mentioned in the introduction, the model is not fitted to the HIV data explicitly. Sexual behavior parameters and other measures are taken as inputs, and the model predicts HIV rates. These are then compared to actual HIV rates, but neither the model or the input parameters are altered to fit the model to the

section 4 we allow for a wide variety of heterogeneity across genders, and the system is closed by the presence of female sex workers (FSWs)

HIV data.[5]

The first sub-section below presents the details of the model. The second discusses the data used in calibration and presents some simple correlations. The final sub-section presents results.

4.1 Model

The world consists of three types of individuals – men, women and female sex workers (FSWs). Men and non-FSW women are tracked through the model by age cohort. An age cohort enters the sexually active population at age 15, and exits at age 60. These (roughly) reflect median reported age at first sexual contact and life expectancy in the region (without HIV). Within each gender-age cohort, individuals are divided by marital status. Prevalence for the age cohort will be a weighted average of the prevalence for single and married individuals, based on the share that are married in that cohort.

Choices about sexual behavior differ by gender and marital status. Women do not have contact with FSWs, so their behavior is less varied than their male counterparts. Married individuals must partner with their spouse, but otherwise have the same choices as single individuals. The rest of this section describes the evolution of infection for single women, married women, single men and married men. The equations presented are for a single age cohort. The full simulated model will use the same type of equations for each age group.

Single women have either no partners, or casual partners. We assume that all women who choose to have casual partners have the mean number of partners, which we denoted $P_{c,sw}$.[6] As before, the male-to-female transmission rate is denoted β_w. In addition, we allow for the possibility of condom use in non-spousal partnerships. We denote the share of condom usage in period t as $1 - c_t$, so the share *without* condom use is c_t. Finally, casual partners are assumed to be drawn from the overall male population, which has a rate of m_{t-1}. This implies that the incidence in the two groups of single women in period t is as follows:

[5]This distinguishes the work here from simulation models in epidemiology (for example,Robinson, Mulder, Auvert and Hayes (1995)).

[6]The model would ideally take into account the distribution of number of partners rather than simply the average. However, the mean will be a good approximation and substantially decrease the complexity of the calculations required.

No partners	$Inc_1 = 0$	(5.1)
Casual Partners	$Inc_2 = (1 - (1 - \beta_w c_t m_{t-1})^{P_{c,sw}})$	(5.2)

We can then denote the overall prevalence among single women in period t as a weighted average of the population of the two groups. We denote the share in group i as $share_i$ and d_t as the death rate in period t (in general, we will assume that people live for 10 years with the disease).

$$sw_t = \frac{sw_{t-1} + (1 - sw_{t-1}) \sum_{i=1}^{2} (share_i)(inc_i) - d_t}{1 - d_t} \tag{3}$$

The logic is very similar for married women, except that all married individuals partner with their spouse. We assume condom usage does not occur in spousal partnerships. Further, we assume that women are married to a man between 1 and 10 years older than them (uniformly distributed), while men will be married to a woman between 1 and 10 years younger than them, which is consistent with actual marriage patterns in the DHS data from the countries used here. Denote $P_{c,mw}$ as the mean number of casual partners for married women, and S_t as the chance that one's spouse is infected in period t. We have the following incidence equations for married women in period t:

Spouse only	$Inc_1 = (\beta_w S_t)$	(6.1)
Spouse and Casual Partners	$Inc_2 = (1 - (1 - \beta_w c_t m_{t-1})^{P_{c,mw}} (1 - \beta_w S_t))$	(6.2)

Similarly, the overall prevalence for married women is:

$$mw_t = \frac{mw_{t-1} + (1 - mw_{t-1}) \sum_{i=1}^{2} (share_i)(inc_i) - d_t}{1 - d_t} \tag{4}$$

As discussed, the overall prevalence in this age group in time t will be a weighted average of sw_t and mw_t based on the share in that age group that is married. Obviously, among older ages the behavior of single individuals becomes less important as their share diminishes.

The evolution of infection rate for a one age group of single women and one age group of married women can be seen in Figures 2a and 2b, respectively. These are designed to give a

visual representation of the equations above; in both cases the figures (although not the model) assume that those who have casual partners have only one casual partner.

The behavior of men follows a similar pattern, but they have more choices about sexual behavior. Specifically, they may have sex in any given period with a spouse (if married), casual partners and/or FSWs. For single men, this implies four groups – those with no partners, those with casual partners only, those with FSW partners only and those with both casual and FSW partners.

Denote the female-to-male transmission rate as β_m and the FSW-to-male transmission rate β_{mp}. In addition, we denote the number of casual partners $P_{c,sm}$ and the number of FSW partners $P_{f,sm}$. We assume that individuals who have partnerships with both FSWs and casual partners have the same number of each as those who have partnerships with only one type, as can be seen in the equations below. Finally, the rate for women in period $t-1$ is w_{t-1} and for FSWs it is f_{t-1}. Condom usage is the same as for women. The following equations determine incidence across groups:

None	$Inc_1 = 0$	(7.1)
Casual only	$Inc_2 = (1 - (1 - \beta_m c_t w_{t-1})^{P_{c,sm}})$	(7.2)
FSW only	$Inc_3 = (1 - (1 - \beta_{mp} c_t (f_{t-1}))^{P_{f,sm}})$	(7.3)
Casual + FSW	$Inc_4 = (1 - ((1 - \beta_{mp} c_t (f_{t-1}))^{P_{f,sm}})((1 - \beta_m c_t w_{t-1})^{P_{c,sm}}))$	(7.4)

As with women, the overall prevalence among single men is a weighted average of the groups:

$$sm_t = \frac{sm_{t-1} + (1 - sm_{t-1}) \sum_{i=1}^{4} (share_i)(inc_i) - d_t}{1 - d_t} \tag{5}$$

Finally, married men have the same behavioral choices as single men, but must also partner with a spouse. Here, $P_{c,mm}$ is the mean number of casual partners for married men and $P_{f,mm}$ is the mean number of FSW parters for married men. The following equations determine incidence:

Spouse only	$Inc_1 = (\beta_m S_t)$
Spouse, Casual	$Inc_2 = (1 - ((1 - \beta_m c_t w_{t-1})^{P_{c,mm}})(1 - \beta_m S_t))$
Spouse, FSW	$Inc_3 = (1 - ((1 - \beta_{mp} c_t (f_{t-1}))^{P_{f,mm}})(1 - \beta_m S_t))$
Spouse,Casual,FSW	$Inc_4 = (1 - ((1 - \beta_{mp} c_t (f_{t-1}))^{P_{f,mm}})((1 - \beta_m c_t w_{t-1})^{P_{c,mm}})(1 - \beta_m S_t))$

The overall prevalence for married men in period t is:

$$mm_t = \frac{mm_{t-1} + (1 - mm_{t-1}) \sum_{i=1}^{4} (share_i)(inc_i) - d_t}{1 - d_t} \tag{6}$$

Again, the overall rate for men in this group will be determined by a weighted average of the rates for single and married.

The final element of the model is the behavior of female sex workers. Unlike the behavior of men and women, the behavior of this group is not age-graded, and it will not differ across countries. However, we do assume some heterogeneity across FSWs. Specifically, consistent with anthropological information from Sub-Saharan Africa, we assume that most "FSWs" are women who work in bars and, while they take money for sex, have relatively few partners per year. The remainder of FSWs work in brothels, and have many partners (Wojcicki, 2002; Nagot et al, 2002).

We denote the rate among the non-brothel FSWs in period t as lf_t and the rate for FSWs in brothels in period t as hf_t. We assume that, since relationships with non-brothel FSWs tend to be long-term, that the transmission rate is the same as the per-partnership rate in the overall population (β_w). However, since partnerships in brothels tend to be more one-off, we denote this transmission rate β_{wp}, where $\beta_{wp} < \beta_w$.[7] . The number of partners for each group is $P_{l,f}$ and $P_{h,f}$, respectively. The incidence in each group is determined by the following equations:

Non-Brothel	$Inc_1 = (1 - (1 - \beta_w c_t m_{t-1})^{P_{l,f}})$	(9.1)
Brothel	$Inc_2 = (1 - (1 - \beta_{wp} c_t m_{t-1})^{P_{h,f}})$	(9.2)

[7]Although, as noted previously, a lock-and-key interpretation of transmission is generally more accurate than a constant per-contact transmission rate, the literature generally suggests that single contact partnerships will have lower transmission rates (Kaplan, 1990). However, in the simulation, this population will reach extremely high HIV levels almost immediately so the assumption about β_{wp} is not crucial.

The overall rate is a weighted average of the two rates:

$$cw_t = \frac{cw_{t-1} + (1 - cw_{t-1}) \sum_{i=1}^{2} (share_i)(inc_i) - d_t}{1 - d_t} \tag{7}$$

The age-specific equations for men and women, and the equations for FSWs, will be simulated together to produce the results in section 4.3. See Appendix B for a presentation of all the equations estimated.

The same model here will also be estimated for the U.S., with the slight adjustment that FSWs play a smaller role and we assume all FSWs fall into the brothel category.

4.2 Data and Correlations

The model above is designed to be calibrated with actual data on sexual behavior, as well as other parameters of the epidemic. The paper is concerned with cross country variation in HIV rate, so all data will be country-specific. A total of 14 countries in Sub-Saharan Africa will be used in the analysis: Benin, Burkina Faso, Cameroon, Chad, Cote d'Ivoire, Ethiopia, Guinea, Kenya, Malawi, Mali, Namibia, Niger, Tanzania and Zambia.[8]

There are six types of data required by the model. The first three are parameters of sexual behavior – the share of people in each behavior group (as divided above), the number of partners for those having non-marital sex and information on condom usage. The fourth parameter is the share of individuals who are married. The fifth is data on the "start date" of the epidemic, and the sixth is data on HIV rates. Details on the source of the data for each parameter are discussed below.

Data on the share of individuals having casual sex are calculated from Demographic and Health Surveys (DHS). These are household-level surveys run in many developing countries. All individuals in the household are asked detailed questions focusing on aspects of health and behavior, including information about their sexual activity. Relevant for this paper, in a subset of the surveys people are asked about their extramarital sexual behavior. In particular, married men and women in the household are asked either: "When was the last

[8]Ideally, we would use all the countries in the region, but we are limited by the availability of data on sexual behavior. However, it should be noted that the data we have covers the major regions of Sub-Saharan Africa that have had very different HIV experiences (the range of HIV rates at the end of the decade is 1.6% to 22%).

time you had sex with someone other than your husband (wife)?", or "How many people other than your husband (wife) have you had sex with in the last twelve months?". Single individuals are asked about the last time they had sex with anyone. An indicator function was created from the responses that is equal to one if the individual reports casual sex (either premarital or extramarital) in the last year, and zero otherwise. This data are tabulated by gender, age and marital status in each country.[9]

Data on the number of sexual partners for those having sex outside of marriage are also drawn primarily from the DHS household survey data. These data, however, are available for a somewhat more limited number of countries. Specifically, data was available for Benin, Ethiopia, Mali, Namibia and Tanzania. In addition, another individual-level survey (the CAPS survey) provided data for Kenya. Rather than heavily limit the sample to only the countries for which these data were available, countries without data were assumed to be similar to those close to them. Although this is obviously far from ideal, the data suggest there is substantially less variance in this than in the share of individuals having extramarital and premarital sex. This indicates this assumption may not be biasing the results very much.

The data on share of individuals having casual sex and number of partners are sufficient to calibrate the model above for women. However, it is insufficient for the men. Namely, we assume men fall into four groups, and the data discussed above allows us only to distinguish between those who have no non-marital partners and those who have some. It does not allow us to differentiate between those who have only casual partners, casual and FSW partners and only FSW partners. We use the assumption of a closed system and data from the literature about the likely shares in each of the non-marital sex groups to fully calibrate the model. Details of the calculation are in Appendix C.

Data on condom use by country are collected from the World Bank Millennium Indicators. This provides information both on the share of individuals reporting condom use and the first year of "condom social marketing" in each country. I assume that condom use

[9]There are obvious potential problems with underreporting in these data. They are, in many way, unavoidable. Nearly all surveys on sexual behavior, particularly of this type, are subject to downward biases. These data do not seem to be worse than other data – for example, the DHS results for Kenya and Tanzania were checked against results of the CAPS survey, another survey focused exclusively on sexual behavior – and found to be similar. In addition, as long as the underreporting is similar across countries, this will not affect the ability of the model to explain variation.

was at zero in the first year of social marketing, and increased linearly until the present level.[10]

The last type of individual-level data necessary is marriage rates by gender and age across countries. As with the data on sexual behavior, these are collected from the DHS household survey data.

Although it is not explicit in the model above, in order to run this simulation it is necessary to know the year the HIV epidemic began in each country. This varies across countries in the sample, reflecting differences in the time of introduction of the epidemic into each country. Information on "year zero" was calculated in two ways. First, information from a number of sources, primarily UNAIDS, the CDC and the World Bank was collected and used to suggest starting dates for each country. Second, data on HIV rates across countries were used. Data for each country were searched and the first year in which any study reported a rate of higher than 1% for pregnant women was reported was marked.

It becomes clear from both methods that the virus introduction date was much earlier in East Africa than in West Africa. There is a substantial censoring issue because testing for the virus was not available until 1982. In a number of East African countries the HIV rate was already quite high by this point, suggesting it was introduced substantially earlier. In West Africa, the rate appears to be zero until the mid-1980s, well after testing is available. This is reflected in the "year zero" data, in which most of the East African countries have starting dates around 1980, whereas those in West Africa are around 1985. I assume that the epidemic was introduced to each country by prostitutes, who are assumed to a have an HIV level of 5% in year zero. This appears to well reflect the data.

Finally, in order to test the fit of the model, it is necessary to calculate HIV rates for these countries over time. The data for this come from the U.S. Census HIV/AIDS Surveillance Database. This database collects all studies of HIV/AIDS prevalence since the early 1980s and extracts them into information on number of subjects, prevalence, population, etc. From this, an average prevalence rate for pregnant women by country-year is calculated. This will then be matched with the predicted rate for women of child-bearing age from the simulation model.[11] In order to avoid noise in the data as much as possible, only years in

[10]This appears to be roughly consistent with the pattern of condom sales in Kenya and Cameroon (Hearst and Chen, 2000)

[11]By using the actual and predicted rates for pregnant women, we avoid many of the issues that have been

which 1000 or more individuals are reported on are used in the analysis.

Before turning to the main results, it is useful to look at some simple graphical comparisons across countries to explore the source of the identification in the simulation model. There are two primary sources of identification: differences in the share of individuals having premarital or extramarital sex, and differences in epidemic timing. Figure 3 illustrates a relationship between sexual behavior and HIV prevalence rate across countries. HIV prevalence rate is graphed against the estimated share of men in the population having sex with a non-spouse (a weighted average of single men having premarital sex and married men having extramarital sex). The corresponding graph for women looks similar. Although the actual data used in the simulation is somewhat more complex (in particular, the share varies by age), this gives some picture of what is driving the model. Within West Africa and East Africa, there is a positive relationship between HIV rate and sexual behavior – more people engaging in risky sexual behavior implies a higher HIV rate. Further, although on average the share of individuals having risky sexual behavior is higher in East Africa, it does not fully explain the differences across regions.

Figure 4 illustrates a similar relationship between the "start year" of the epidemic and the HIV rate. Here, there are significant differences between West and East Africa but relatively little difference within region. The graphs together suggest that nearly all of the within-region identification will come off of differences in sexual behavior. However, identification across region depends in part on differences in sexual behavior and in part on differences in timing.

It is worth noting, in addition, that the parameters for the United States are also included on this graph. It should be clear that neither timing or risky sexual behavior is driving the differences between the U.S. and Africa. The frequency of risky sexual behavior in the U.S. is similar to that of Malawi, and the epidemic timing also suggests similarity to East Africa. The actual HIV rate among pregnant women in the U.S. is extremely low, implying that neither of these facts is driving the cross-continent differences.

Finally, we consider the possibility that differences in condom use between the

recently noted in surveillance data in which the rate for the entire adult population is estimated from that for pregnant women.

developed and the developing world are driving these results. Figure 5 shows the relationship between HIV rate and share of individuals reporting condom usage in 2000. This graph suggests that there is little relationship between condom use and HIV rate within Africa. The U.S. has higher condom usage, as would be expected, but not substantially higher. We will see later that this is not nearly enough to explain the differences between the U.S. and Sub-Saharan Africa.

4.3 Simulation Results

The main simulation results are presented in Figure 6, which shows the graph of the predicted HIV rate for pregnant women from the model against actual HIV rate (the line is a 45 degree line). Both are averaged over the period 1998-2000. The model is an extremely good fit to these data. It is able to differentiate both between East and West Africa, and between countries within those regions. The largest errors are in Benin and Ethiopia. This may be, at least in part, an issue with the estimates of HIV rates from the Census Surveillance Database. In the case of Ethiopia, in particular, there are relatively few sites sampled even in later years, and they are virtually all in the capital city, which may well have higher prevalence than outlying areas.[12]

Given that HIV/AIDS surveillance in Africa improved dramatically over the 1990s, the comparison between end-of-decade predicted and actual rates may be the most apt. However, it is worth considering whether the model is a good fit to the time series of the data as well as the cross section. Figure 7, therefore, includes estimates for each country from the beginning (1990-1992), middle (1994-1996) and end (1998-2000) of the decade. The fit of the model clearly improves over time (see the mean absolute errors), although in many ways it fits the time series well. Most of the errors in the earlier time periods reflect the simulation underestimating relative to the actual estimated rate. This may be due in part to errors in the actual rate during these time periods – early in the epidemic testing was more likely to be done in higher-risk areas, particularly in countries where the epidemic was becoming a

[12]The data for Namibia was not released until the model in the paper was finalized, providing an out of sample test. The model is an extremely good fit for Namibia, suggesting it is likely to be successful even out of sample.

problem such as Tanzania and Zambia. It is encouraging that Kenya, which has quite good surveillance over the whole decade, is a good fit to the model in all three time periods.

The preceding subsection argued that the differences between the U.S. and Africa did not appear to be due to differences in sexual behavior or timing. In the section on transmission rates it was argued that these differences are likely to be due to differences in transmission rates. There are two ways to illustrate this – we can consider the predicted values for the African countries under U.S. transmission rates, or the predicted values for the U.S. under African transmission rates. Both can be seen in Figure 8. When African transmission rates are used in the model, the U.S. is an extreme outlier – a predicted rate of over 20%, and an actual rate of 0.15%. When transmission rates from the U.S. are used, the predicted value for the U.S. is quite close to the actual value (predicted around 0.30%), but the predicted values for Africa are all dramatically lower then the actual values. These graphs take into account differences in condom usage rates, sexual behavior and timing across countries and continents. They therefore reinforce the earlier point that condom use is not driving differences across continents; rather, they lend significant support to the theory that differences across continents are driven by differences in transmission rates.

Having demonstrated that this simulation model is a good fit to the HIV epidemic in Africa, we turn to using the model to explore the cost-effectiveness of interventions designed to curb the spread of AIDS.

5 Cost Effectiveness of Interventions

One of the primary advantages of the simulation model presented above is that it allows us to explore the effect of several types of interventions. By assuming that parameters of the model are altered (or not) in the current year, and simulating the model over the next several years, it is possible to produce time paths for infection with no intervention and with interventions designed to slow the epidemic growth. I will consider two types of interventions, which are discussed in more detail below. The first type are interventions designed to decrease the probability of infection per sexual contact with an infected partner. This is achieved through treatment of STI infections, either treatment of bacterial infections, or treatment of bacterial

infections and herpes suppression. The second type of intervention is designed to bring about behavioral change, as has been seen in Uganda over the past fifteen years. Obviously these are not the only types of interventions, nor are they the only types designed to affect the spread of the heterosexual epidemic. However, these do encompass the two major types of interventions focused on the heterosexual spread of the disease. In addition, it is possible to do some cost-effectiveness analysis, which is not possible with most other interventions. At the end of this section, I compare the cost-effectiveness results for these simulations with two different avenues – treatment of HIV and prevention of mother-to-child transmission.

This is not the first work to consider these questions of cost-effectiveness. Good summaries of the cost-effectiveness of specific HIV interventions are contained in Creese, Floyd, Alban and Guinness (2002) and Kumaranayake and Watts (2001). The primary difference between the work here and earlier work is that I attempt to evaluate the future cost-effectiveness of a very large-scale intervention covering the entire continent. The results here are likely to give a much better sense of how these interventions could affect the overall epidemic. In addition, previous studies do not use this type of simulation approach.

For all interventions, I consider several parallel measures of their effectiveness – life years saved, disability adjusted life years (DALYs)[13] saved and total infections averted. These calculations take into account the different effects of each intervention on individuals of different ages, which is particulary necessary for calculations of life years and DALYs, where it is important to note whether someone acquired the disease at a young age, or close to when they might have passed away from something other than HIV/AIDS.

In each case, the model is simulated in the no-intervention scenario, and under the intervention. The output for each is a time path of infection for each gender-age cohort in each country over the next decade. Data on population for each gender-age-country cohort was used to calculate the number of new infections in the case of non-intervention, and the case of intervention. The difference is the "effectiveness" – the number of infections averted. In the case of life years, I assume a life expectancy of 60 (about average for this region without HIV infection) and a length of infection of 10 years. Then, for example, a fifteen-year-old who

[13]A DALY is a life year adjusted for changes in quality of life. In general, if individual quality of life is half as good, this represents a loss of 0.5 DALYs.

becomes infected has lost 35 life years. Finally, DALYs are calculated in much the same way as life years, except we assume some loss of mobility and productivity while the individual has the disease. Here, I follow Gilson et al. (1997) and assume that individuals lose 0.1 "life years" during the first nine years of infection, and 0.9 during the final year before death.

5.1 STI Treatment Interventions

I consider two interventions that would treat STI infections, both of which would lead to decreases in the transmission rate of HIV per sexual partnership. The first mimics the Mwanza, Tanzania, intervention from 1995 in which bacterial STIs only were treated, and resulted in decreases in transmission rate relative to a control group. Specifically, the difference between estimated transmission rate per sexual contact for treatment and control women was 25%, and for treatment and control men, 36%. I apply this decrease to the transmission rates used in the model, to obtain transmission rates of 0.20 male-to-female, and 0.075(0.0533) female-to-male (non-FSW and FSW transmission). The intervention is simulated by assuming that in the next year after the results in section 4 (which is 2002, since the HIV data were limited to 2001 and before) the intervention was introduced and the transmission rate immediately decreased to its new levels, and remained there for the following decade.

The cost of this intervention is calculated using information on the Mwanza trial. In that trial, they provided services to 150,000 individuals at a cost of $59,060 per year (in US$1993). To get the cost for the entire population considered in the intervention, in current dollars, we scale up by population. The total population of the 13 countries[14] in which we consider the intervention is 235 million, which implies that the yearly cost in US$2000 is around $103.6 million, with a ten-year cost of $1.04 billion.

The second STI intervention (herpes suppression therapy along with bacterial STI treatment) is more difficult to calibrate because, while a trial of this type is currently underway, none has been done thus far. Nevertheless, it is important to consider because eliminating virtually all genital ulcers could have enormous effects on transmission rates.

[14]Although 14 countries were simulated in the model above, we exclude Benin, because it is a significant outlier and might therefore bias the results.

Unlike in the case of the bacterial STIs, there is no easy way to estimate the decrease in transmission rate with treatment of all genital ulcers. I use information from within specific studies on individuals who do and do not have genital ulcers. Specifically, I use data from the Gray et al. (2001) study, which suggests a difference in per-partnership transmission rate of 70% with and without genital ulcers. De Vincenzi (1994) suggests a difference of 78% between the two groups. The more cautious of these suggests using transmission rates in the model to 0.0787 male-to-female, and 0.0349(0.0233) female-to-male (non-FSW and FSW transmission). These are lower (in the case of the female-to-male) than the transmission rates calculated for the U.S. It seems unrealistic to expect treatment better than that available in the U.S., so we assume that the lowest transmission rates possible are those calculated for the U.S. and we use those in the simulation.

It is clear that there are extreme limitations to this approach to evaluating the effect of this type of treatment. However, given the huge differences between those with genital ulcers and those without, and the enormous impact of herpes in Sub-Saharan Africa, it seems worthwhile to explore this intervention, even while noting important limitations.

The cost of the intervention is also difficult to evaluate. In this case, I assume that the same infrastructure that was used in the Mwanza trial could be used again, and the only additional cost would be the herpes suppression drugs. This seems to be a reasonable assumption, as it is not difficult to diagnose an additional STI when testing is already being done, and distribution of drugs could be done for many treatments simultaneously. Herpes suppressive therapy is achieved with a drug called acyclovir, which is given in a suppressive dose of 400mg twice a day, and must be taken constantly. The cost per dose is around \$0.06, implying a yearly cost of \$43 per person.[15] It is also necessary to know how many people would likely require this therapy. Data suggest that around 44% of women and 20% of men in this region are infected with herpes, implying that of the 235.0 million individuals in the regions that the intervention covers, approximately 76.3 million would need suppressive therapy. This results in a total yearly cost of \$3.4 billion, or a ten-year cost of \$33.8 billion, in US\$2000.

[15]This information was provided by the University of Washington researchers currently exploring the effects of this therapy on HIV transmission in Africa, and reflects their drug costs.

5.2 Behavioral Change Interventions

A second class of interventions considered here seeks to change the share of people having extramarital or premarital sex, and the number of partners these people have. Essentially, this is the Ugandan approach to HIV/AIDS, and by most accounts it has been quite successful at changing behavior and affected HIV incidence. Using data on sexual behavior at the beginning, middle and end of the 1990s, it is possible to calculate the decrease in percent of individuals with extramarital or premarital partners, and the decrease in the number of such partners.

Specifically, data from the WHO/GPA and DHS surveys in 1989 and 1995 allow us to calculate the decrease in percent of women having premarital sex (35% to 22%), women having extramarital sex (6% to 3%) and men having extramarital sex (23% to 16%). Unfortunately, data limitations prevent us from doing a similar calculation for single men. We therefore make the assumption that their percentage decrease was the same as for single women. Using data from 1995 and 2000, we can also calculate decreases in the number of partners for those having casual partners (either premarital or extramarital). We find an estimated decrease of 10% for single men, 13.5% for single women, 10% for married women, and no apparent decrease for married men.

Given these data, we model the effect of the intervention in two ways. The first is to assume that the intervention achieves a constant percentage decrease in sexual behavior each year. So, for example, the data on the share of single women having premarital sex suggests a 7.4% decrease in this variable in each year of the intervention (obviously, this would eventually be capped, but we assume that the decrease would continue at least through the decade intervention explored here). In this first sexual behavior intervention we assume that all of the parameters of sexual behavior decrease in this smooth way. The second possibility is that the intervention immediately achieves the entire decrease in sexual behavior in the first year, and the behavior is then constant at this lower level over the rest of the intervention. It is worth noting that the truth probably lies somewhere in between these two extremes. It is probably the case that the first year of the intervention archives larger-than-average results, but that the total change is increasing over time. Nevertheless, by exploring both

interventions we will get a sense of the bounds on the effects.

Regardless of which model of the intervention is used, the costs will be the same. We simply assume that the costs reflect a scaling up of the Ugandan experience. According to Hogle (2002), the total ten year cost from 1989-1998 of the intervention was $180 million in donor financing, and about $77 million in national spending, for a total ten-year cost of $257 million. Uganda has a population of around 24 million, so to scale up the intervention to the 235.0 million people affected in the countries considered here implies a total ten year cost of around $2.8 billion, in US$2000.

5.3 Results

Before using the data on cost to address cost effectiveness issues, it is worthwhile to consider which of the interventions above is the most effective. Figure 9 shows the path of infection for the countries simulated (a population-weighted average), under no intervention, the bacterial STI intervention, bacterial STI treatment and herpes suppression, continuous behavioral change and abrupt behavioral change. The figure demonstrates a number of features of the interventions. First, all of the interventions are successful in the sense of at least slowing the epidemic growth. Second, both of the STI treatment interventions appear to be more effective than either of the behavioral interventions. Finally, treatment of herpes along with bacterial STIs makes a huge difference in the time path of infection, and the herpes suppression intervention is by far the most effective. However, it is also the most expensive, so given a limited budget, this may not be the best use of finances.

Calculations of cost effectiveness of the interventions demonstrates that the treatment of bacterial STIs only is by far the most cost effective intervention, at a cost of $3.81 per life year and $72.48 per infection averted, on average over the decade-long intervention. This calculation, and the calculations for the other intervention, are shown in Table 3. Despite the fact that it is the most effective, the herpes suppression intervention is also the most expensive, at an average cost of over $1100 per infection averted over the entire period. The behavioral interventions are intermediate in their cost effectiveness, with the abrupt behavioral change more cost effective than the gradual change.

As would be expected, each of these interventions becomes more cost effective over

time, relative to doing nothing. Costs are constant over time (if anything, the drug interventions may get less expensive over time as fewer people are passing around STIs), but as the intervention matures, the difference between doing nothing and the intervention grows (as can be seen clearly from Figure 9). Interestingly, however, the time path of cost effectiveness is not constant across interventions. This is demonstrated in Figure 10, in which cost effectiveness for each intervention is graphed against year. The bacterial STI treatment intervention is always the most cost effective in this analysis, but the cost effectiveness of the behavioral interventions converge with the bacterial STI treatment over time. In addition, the gradual behavioral change is actually less cost effective than herpes treatment initially, but by the second year it is already more cost effective.

5.4 Combined Interventions

In addition to the single-issue interventions discussed above, it is interesting to consider the possibility of combining them. It is obviously possible both to treat STIs and to change sexual behavior at the same time. We will therefore consider two additional possibilities – sexual behavior change combined with each of the STI interventions. In both cases we assume gradual behavior change, which is likely to be a closer approximation to reality.

The results suggest that there is relatively little additional gained by the combination, at least in the case of the combination with herpes suppressive therapy. Table 4 shows the results of these interventions, and their cost effectiveness calculations. The cost effectiveness of the first combined intervention falls somewhere between the effectiveness of the two interventions, but is still higher than the STI treatment alone. The number of infections averted is increased by around 2 million over the decade. The second combined intervention (behavioral change and all STI treatment) is less cost effective than either intervention alone, reflecting the fact that there is little additional effect gained. However, the number of infections averted is increased slightly, by around 1.1 million over the decade.

5.5 Alternative Interventions

The model above is designed to model the heterosexual HIV epidemic, and the interventions considered are therefore limited to those intended to decrease heterosexual transmission. However, it may be interesting to consider how these results compare to cost-effectiveness of other popular prevention/treament options. In particular, we may be interested to compare these with cost-effectiveness results from mother-to-child prevention and treatment with antiretrovirals.

Creese et al (2002) provide a comprehensive review of the cost-effectiveness of a number of interventions of this type. They consider two types of interventions to prevent mother-to-child transmission: drug treatment during pregnancy and discouraging breastfeeding. The cost-effectiveness of drug interventions (treatment during pregnancy with nevaprivine) vary from a low of $140 per infection averted to a high of over $2000 per infection. Breastfeeding interventions are less cost-effective, at a cost of between $3000 and $21,000 per infection averted. On average, these are somewhat higher than the cost-effectiveness numbers presented for the bacterial STI intervention above, although it is important to note that in the case of mother-to-child prevention the infection averted is in a baby, so there are more life-years saved per infection.

Creese et al (2002) also present some results from treatment with anti-retrovirals in Africa. These results generally suggest that this type of treatment is extremely expensive relative to prevention. The two studies they cite suggest a cost of between $1100 and $1800 per life year (obviously infections averted is not a reasonable metric). This is quite a bit more expensive than the interventions discussed above. In addition, it is difficult to imagine how this type of intervention could be more cost-effective than the drug interventions discussed above, even if the price of anti-retrovirals decreases, given that the infrastructure needed to provide the drugs is quite large.

Although less often discussed, simple treatment for tuberculosis can extend life for AIDS patients dramatically. Tuberculosis is quite a common opportunistic infection, and the treatment is very simple (in many way, simple in the same spirit as treatment of bacterial STIs). Creese et al (2002) discuss this treatment as well, and their review suggests the cost

per DALY ranges from $2 to $16, dramatically lower than other treatment options. It is worth noting that this type of intervention could be complementary with the STI drug interventions discussed above. These STI interventions require setting up testing and treatment clinics, which might also be used to dispense antibiotics to those with tuberculosis. In general, this simply reiterates the fact that inexpensive drugs may be better – for both treatment *and* prevention – than anti-retrovirals, even if they are off patent.

6 Robustness and Extensions

The analysis above suggests that the most cost-effective response to the HIV epidemic is to treat non-HIV STIs in Africa. This would reduce the transmission rate from infected to uninfected individuals, and slow the epidemic spread. These results are obviously quite reliant on the model, and on the assumption that each intervention acts independently on the population. It is worth considering the effects of relaxing these assumptions on the results. The first two subsections below discuss the introduction of individual heterogeneity into the model. The third discusses the general equilibrium issues.

6.1 Individual Behavioral Heterogeneity

Thus far, the simulation model has followed age-gender cohorts through time, but has not followed individual actors. That is, the model simulates incidence rates based on the aggregate group behavior; it does not simulate the time path of infection for a given individual with a certain set of behavioral characteristics. This methodology is not problematic as long as we can assume that individual behavior is uncorrelated from one period to another. That is, as long as we assume no individual-specific heterogeneity within an age-gender cohort this methodology will produce accurate results. Of course, it is probably unrealistic to make this assumption. It may be more realistic to imagine that, for example, men who engage in premarital sex with FSWs are more likely to continue to have this type of relationship after marriage. It is worth considering, therefore, how significantly the results are affected if we allow for individual heterogeneity.

Allowing for individual heterogeneity requires the ability to "follow" individuals over

time. That is, to fully specify this type of model it would be necessary to know what the correlation structure is between individual behavior at age 15, 16, 17 and so on. We could then create individuals with these characteristics and compare their simulation results with the output for the homogeneity case. Given data limitations, however, we do not have this type of rich data about individuals. Instead, we introduce heterogeneity into a slightly simpler version of the model presented above and compare the results with the same simpler model with a homogeneity assumption.

In particular, we modify the model in section 4 so we have individuals of only four types: single men, married men, single women and married women and we have behavior for each type. We assume that all women get married at age 20 and all men get married at age 25. We then run the same simulation equations discussed in Section 4 with the same type of cross country parameters, with the modification that the parameters are now gender-marital status specific and not age-gender-marital status specific. Although this model yields less precise estimates than those produced by the full simulation model, the results are not enormously different, and the cross-country variation remains well-explained in the simulation results (See Appendix Figure A).

By simplifying the model in this way we are able to introduce an extreme form of heterogeneity. In particular, we assume that individuals come in a continuum of types and we use the data on share of individuals in each risky sexual behavior group to calibrate the measure of each type. This is easiest to see for women. We group women into three types: those who have premarital sex and extramarital sex, those who have premarital sex only and those who have neither form of casual sex. This implies that, instead of assuming that women who have risky sex in the last period are equally likely to have or not have it in this period, there is a certain set of women who always have risky sex, a set who are slightly less risky and some who never engage in risky sex.

For men, the situation is more complex because there are more "types", but the same logic applies. There are some men who have both casual partners and FSW partners when they are single, and they continue to have them when they are married. There are others who have casual partners and FSW partners when single, but only FSW partners (or only casual partners) when married, and so on. It is worth noting that this is an extreme form of

heterogeneity in which people's behavior is (in a sense) perfectly correlated over time. The reality is probably somewhere between this heterogeneity assumption and the assumption of totally uncorrelated behavior that we implicitly make in the original model.

The simulations required for these models are structured very similarly to those for the model in section 4, with the caveat that the model in section 4 follows groups over time, and the models created for this robustness check actually simulate each individual and tracks them over the period.

The result we are concerned with here is the deviation of the heterogeneity model from the results of the homogeneity model. In general, we expect the predicted values in the case of heterogeneity to be lower than those for the homogenous case, because those who have more sex are more likely to die out and therefore less likely to infect others. The results can be seen in Figure 11. As hypothesized, the predicted values in the case of heterogeneity are generally slightly lower. This is primarily true when the epidemic is more advanced. However, the difference between the two models is not substantial, even in the countries in which the error is the largest (the largest error is only 5 percentage points).

This result suggests that this issue, while important, is probably not biasing the results in a significant way. In addition, it is worth recalling that this is quite an extreme version of heterogeneity, and the true heterogeneity in the population is likely to be producing even less bias than demonstrated in the results here.

6.2 Transmission Rate Heterogeneity

Another type of heterogeneity that may affect these results is heterogeneity in transmission rates per partnership within Africa (we have obviously discussed at length the issue of transmission rate heterogeneity across continents). There are two types of heterogeneity that we might be concerned about – first, based on STI status of the partners, and second based on circumcision status of partners.

The issue of partner STI status is the most straightforward. As the model is currently structured, we assume that all partnerships (within continent) have the same per partnership transmission rate. However, given that much of this analysis rests on the idea that having an STI makes infection more efficient it is reasonable to attempt to model this more accurately.

Specifically, we might imagine modifying the model such that partnerships between two individuals that both have active STIs results in infection much more often than partnerships between two people neither of whom is infected.

It is possible to build this into the model, if the infection rate for each type of partnership is know. That is, if we know the transmission rate for a partnership in which both partners have STIs, and the rate for partnerships in which only the woman does, only the man does and neither do then it is relatively simple to build this into the model (assuming non-selective matching). With a modified model of this type, *if the weighted average transmission rate remains the same* the predicted infection rate is almost identical to what is calculated in the overall model. This is fairly straightforward: if the average infection rate is just a weighted average of the infection rates for each type of partnership then we should not find any differences.

This is, however, as far as this robustness check can go because we simply do not have information on per-partnership transmission rates for different types of partnerships. Even if we assume that two people without STIs have the same transmission rate as partnerships in the developed world, there are still too many degrees of freedom. The literature is simply not thick enough to provide good estimates for these parameters.

A second issue that has come up increasingly in the literature is the issue of male circumcision and efficiency of HIV transmission. A number of studies have suggested that circumcised males are less likely to become infected with the virus than uncircumcised males, and they have attributed differences in infection rates across countries in Africa to this issue (for reviews, see Weiss et al (2000), Siegfried et al (2003), Halperin et al (1999), Moses et al (1994)).

Although there appears to be some evidence for this theory, I have argued elsewhere that none of the evidence that exists thus far is sufficient to conclude that circumcision status actually affects transmission rates (Oster, 2004). In addition, if circumcision status is truly responsible for much of the transmission rate heterogeneity, it is difficult to understand why HIV rates in Europe (primarily uncircumcised) are lower than the U.S. (primarily circumcised). Given that difference in sexual behavior appear to be sufficient to fully explain differences across countries within Africa, I would argue there is little need to appeal to this

explanation.

6.3 General Equilibrium Effects

The results in this paper argue that bacterial STI treatment is the most cost-effective available to curb the spread of HIV. It is worth noting, however, that this is a partial equilibrium analysis. That is, if treating STIs leads people to have more sexual partners then the effect of this intervention would be (at the least) tempered and (at the most) reversed.

There are a few ways to address this concern, and to get a sense of its importance. The first thing to note is that the Mwanza study (from which effectiveness and cost data was drawn) did attempt to explore the effect of the intervention on sexual behavior. The results presented suggest little or no behavioral change. The average number of partners for men decreased slightly in both the intervention and comparison group, and the average number of partners for women decreased in the intervention group and increased slightly in the comparison group. The percent reporting a casual partner is reported only for the follow-up survey, and it appears to be lower in the intervention than the comparison group implying, if anything, a decrease in risky sexual behavior with the intervention. This result indicates that, perhaps, the previous analysis would be changed very little, if at all, by the possibility of changes in sexual behavior.

The possibility of altered sexual behavior, however, remains an issue despite the Mwanza results. To get a sense of the importance of the issue, we can use the simulation model in section 4 to explore how much sexual behavior would have to change such that the behavioral intervention would be more cost effective than the STI intervention. We consider a modification of the STI intervention in which all sexual behavior (percent of individuals engaging in risky sexual contact and the number of sexual partners) increases by a constant fraction when the STI treatment is introduced. The results suggest that the cost per life year would exceed that for the behavioral intervention at around a 35% increase in sexual behavior, whereas the cost per infection would exceed the behavioral intervention only at an increase of about 42%.

This leaves us with the question of whether it is possible that we could see this type of behavioral response. Certainly it is vastly more than suggested by the Mwanza trial itself.

On the other hand, the behavioral change seen (in the other direction) in Uganda was not much less than this, which implies that at least this magnitude of behavioral change is not impossible.

In a final attempt to calibrate the magnitude of this problem, we turn to one of the few instances where the perceived threat of HIV has decreased over the last decade: among gay men in the United States. There has been substantial evidence within the last seven or eight years that the perceived threat of HIV among the gay community has decreased, at least in large part because of the introduction of antiretrovirals. Although this is not the same as a change in the perceived change of acquiring the virus, it is worth considering the behavioral response. Ekstrand et al (1999) report an increase in the percent of gay men having unprotected anal intercourse from 37% in 1992 to 50% in 1996. This result suggests that it would certainly be possible to have a behavioral change large enough such that the STI intervention was less cost effective. On the other hand, the differences between the case of gay men in the U.S. and heterosexual behavior in Africa are large, and ultimately it seems unlikely that we could possibly see behavioral change large enough to reverse the cost effectiveness analysis.

It is also worth noting, as a final point, that the Mwanza trial may have seen low behavioral change because it also included some counselling on risky sexual behavior. As this is included in the cost of the trial, it is taken into account in the calculations of cost, and we could expect similar behavioral responses in a larger-scale trial. In general, is important to note that failing to take into account the possibility of endogenous behavioral change will bias the results in favor of drug interventions, and away from behavioral change intervention. This is important to consider when ultimately making decisions about policy interventions in this type of situation.

7 Discussion and Conclusions

This paper is motivated by the observation that there is a finite amount of money available to deal with the HIV/AIDS epidemic. The goal is to understand the differences across countries and use this understanding to identify the appropriate use of the limited funds. I find that the

"best" intervention is treatment of bacterial STIs, which is by far the most cost effective solution. With unlimited amounts of money the ideal is obviously to do all the possible interventions – treatment all types of STIs, including herpes, *and* intervene to change behavior. What is not yet clear from the simple analysis above, however, is the best intervention given the resources available, and whether the ideal (budget-constrained) intervention would combine more than one of these together.

To attempt to address this issue I consider a counterfactual: what would the best use of the funds spent in 2002 have been, and what share of infections might have been prevented? Actual 2002 expenditure on HIV/AIDS in low-and-middle income countries was around US$3.9 billion. Although it is difficult to estimate exactly how much went to Africa, we can assume it is a large share, given their disproportionate infection burden. Rather than simply assuming that all of this money could be spent on the optimal intervention, we explore how many infections could have been prevented with a number of expenditure levels. One thing that is important to note is that each of the interventions explored must be affected on at least a relatively large region all at once. The model does not allow for the idea that half of the people in a country were treated for STIs, or that half of the population was told to limit their risky sexual behavior. When we consider the interventions, therefore, we assume that each intervention must be affected in an entire country all at once. Although this is a limitation, and the model could be changed to allow for it, it does not seem unreasonable for policy exploration. In addition, in this case we are considering only the cost effectiveness of the first year of interventions. We assume, as when interventions were combined, that gradual sexual behavior change is the most accurate model.

By far the most cost effective intervention is treatment of bacterial STIs alone, implying that the first money should be spent on this intervention. To determine the next best intervention, it is necessary to consider not the cost effectiveness from the intervention alone, but the cost effectiveness of the incremental lives saved. This requires considering not only the number of incremental lives, but the incremental cost. For example, the incremental cost of the herpes treatment over bacterial STI treatment is only the cost of the herpes suppression drugs, not the fixed costs. Taking this into account implies that the second intervention that would ideally be undertaken is herpes suppression. Only after this has been

fully financed should money be spent on sexual behavior change.

Table 5 formalizes the paragraph above, and explores the cost and the effectiveness of each intervention. Note that the costs differ from those in Table 4 because we are considering scaling the intervention up to the entire continent. The total new infections in Sub-Saharan Africa in 2002 was 3.5 million, and Table 5 reports the number of these that could have been averted with each expenditure level, with the share averted coming from the results of the simulation compared with the total new infections predicted. The results suggest that for about $300 million, we could have prevented 24% of the new infections, around 835,000 in Sub-Saharan Africa. This was clearly feasible at the time, given the total level of expenditures was about 13 times as high. The next level of the intervention is substantially more expensive, costing over $9 billion to prevent slightly over half of the new infections in the year; the addition of the sexual behavior intervention provides only small increases in both cost and infections prevented. It is clear, then, that at the 2002 expenditure level it would not have been possible to achieve this type of reduction in infection levels. The optimal strategy, however, is to treat herpes in as many countries as possible, starting with those with the highest level of infection. If we imagine approximately $3 billion of the total $3.9 billion was spent in Sub-Saharan Africa, this would have been enough to use herpes suppression therapy in 18 countries, including the most heavily affected – Botswana, Zimbabwe, Zambia, etc. In total, estimates suggest roughly 40% of the new infections in the region could have been prevented by this use of funds. It is difficult to say for certain how many new infections were actually prevented with this 2002 expenditure, but it should be very clear that it was nowhere near this many.

To this point, the argument for interventions to treat STIs has focused on their cost-effectiveness. It is worth noting, however, that this is one of the few interventions that people are likely to *want* to participate in. For example, it may be difficult to convince people to have fewer sexual partners or use condoms even with dire warnings about the consequences. However, it is likely to be substantially less difficult to convince someone to come in and have their genital ulcers treated. By providing people with something they *want* it may be easier to affect this change than it would be with other interventions.

The HIV/AIDS epidemic threatens to cripple Africa – to slow economic growth. This

paper suggests that the epidemic can be understood in a relatively straightforward way as largely the result of sexual behavior and a transmission probability enhanced by untreated STIs. Relatively inexpensive interventions could dramatically decrease the incidence of the disease over the coming decade, particularly if the funds are focused on STI treatment.

References

Anderson, RM and G Garnett, "Strategies for limiting the spread of HIV in developing countries: conclusions based on studies of the transmission dynamics of the virus," *Journal of Aquired Immune Deficiency Syndrom Human Retorvirus*, 1995, *9* (5), 500–513.

___ , S Gupta, and W Ng, "The significance of sexual partner contact networks for the transmission dynamics of HIV," *Journal of Aquired Immune Deficiency*, 1990, *3* (4), 417–429.

Bonnel, Rene, "Economic Analysis of HIV/AIDS," Technical Report, World Bank 2000.

___ , "HIV/AIDS: Does It Increase or Decrease Growth in Africa," *World Bank Working Paper*, 2000.

Cameron, DW, JN Simonsen, LJ DCosta, AR Ronald, GM Maitha, MN Gakinya, M Cheang, JO Ndinya-Achola, P Piot, RC Brunham et al., "Female to male transmission of human immunodeficiency virus type 1: risk factors for seroconversion in men," *The Lancet*, 1989, *2* (8660), 403–407.

Carael, M, J Cleland, JC Deheneffe, B Ferry, and R Ingham, "Sexual behaviour in developing countries: implications for HIV control," *AIDS*, 1995, *9* (10), 1171–1175.

Cohen, M, I Hofman, R Royce, P Kazembe et al., "Reduction of Concentration of HIV-1 in Semen After Tratment of Urethritis: Implicaitons for Prevention of Sexual Transmission of HIV-1," *The Lancet*, 1997, *349*, 1868–1873.

Corbett, E, R Steketee, F O ter Kuile et al., "HIV-1/AIDS and the Control of Other Infectious Diseases in Africa," *The Lancet*, 2002, *359*, 2177–2187.

Creese, A, K Floyd, A Alban, and L Guinness, "Cost-Effectivness of HIV/AIDS Interventions in Africa: A Systematic Review of the Evidence," *The Lancet*, 2002, *359*, 1645–1652.

De Vincenzi, I, "A longitudinal study of human immunodeficiency virus transmission by heterosexual partners," *New England Journal of Medicine*, 1994, *331* (6), 341–346.

Downs, AM and I De Vincenzi, "Probability of heterosexual transmission of HIV: relationship to the number of unprotected sexual contacts," *Journal of Acquired Immune Deficiency Syndrome Human Retrovirology*, 1996, *11* (4), 388–395.

Ekstrand, ML, RD Stall, JP Paul, DH Osmond, and TJ Coates, "Gay men report high rates of unprotected anal sex with partners of unknown or discordant HIV status," *AIDS*, 1999, *13* (12), 1525–1533.

Gilson, L, R Mkanje et al., "Cost-effectiveness of Improved Treatment Services for Sexually Transmitted Diseases in Preventing HIV-1 Infection in Mwanza Region, Tanzania," *The Lancet*, 1997, *350.*

Gray, R, M Wawer, R Brookmeyer et al., "Probability of HIV-1 Transmission Per Coital Act in Monogamous, Heterosexual, HIV-1 Discordant Couples in Rakai, Uganda," *The Lancet*, 2001, *357*, 1149–1153.

Grosskurth, H, F Mosha, J Todd, E Mwijarubi et al., "Impact of Improved Treatment of Sexually Transmitted Diseases of HIV Infection in Rural Tanzania: Randomized Controlled Trial," *The Lancet*, 1995, *346*, 530–536.

Halperin, DT and RC Bailey, "Male Circumcision and HIV Infection: 10 Years and Counting," *The Lancet*, 1999, *354*, 1813–1815.

Hogle, Janice, "What Happened in Uganda?," Technical Report, USAID 2002.

Hooper, Edward and Bill Hamilton, *The River*, Little, Brown and Company, 1999.

Kamradt, T, D Niese, HH Brackmann et al., "Heterosexual transmission of HIV in hemophiliacs," *Klin Wochenschr*, 1990, *68* (4), 1203–1207.

Kapiga, SH, H Bang, D Spiegelman, GI Msamanga et al., "Correlates of plasma HIV-1 RNA viral load among HIV-1-seropositive women in Dar es Salaam, Tanzania," *Journal of Acquired Immune Deficiency Syndrome*, 2002, *30* (3), 316–323.

Kaplan, Edward, "Modeling HIV infectivity: must sex acts be counted?," *Journal of Acquired Immune Deficiency Syndrome*, 1990, *3* (1), 55–61.

Kumaranayake, I and C Watts, "Resource Allocation and Priority Settion of HIV Intervention: Addressing the Generalized Epidemic in Sub-Saharan Africa," *Journal of International Development*, 2001, *13*, 451–466.

May, RM and RM Anderson, "Transmission dynamics of HIV infection," *Nature*, 1987, *326* (6109), 137–142.

Moses, S, FA Plummer, JE Bradley, JO Ndinya-Achola, NJ Nagelkerke, and AR Ronald, "The Association Between Lack of Male Circumcision and Risk for HIV Infection: a Review of the Epidemiological Data," *Sexually Transmitted Diseases*, 1994, *21* (4), 201–210.

Nagot, N, A Ouangre, A Ouedraogo, M Cartoux, P Huygens, MC Defer, T Zekiba, N Meda, and P Van de Perre, "Spectrum of commercial sex activity in Burkina Faso: classification model and risk of exposure to HIV," *Journal of Acquired Immune Deficiency Syndrome*, 2002, *29* (5), 517–521.

Oster, Emily, "HIV and Male Circumcision: Differences May Not Be Due to Different Transmission Rates." Mimeo, Harvard University.

Ragni, MV, LA Kingsley, P Nimorwicz, P Gupta, and CR Rinaldo, "HIV heterosexual transmission in hemophilia couples: lack of relation to T4 number, clinical diagnosis, or duration of HIV exposure," *Journal of Acquired Immune Deficiency Syndrome*, 1989, *2* (6), 557–563.

Robinson, NJ, DW Mulder, B Auvert, and RJ Hayes, "Modelling the impact of alternative HIV intervention strategies in rural Uganda," *AIDS*, 1995, *9* (11), 1263–1270.

Rockstroh, JK, S Ewig, T Bauer, G Luchters et al., "Male-to-female transmission of HIV in a cohort of hemophiliacs–frequency, risk factors and effect of sexual counseling," *Infection*, 1995, *23* (5), 29–32.

Sachs, Jeffrey, "Fighting the AIDS Pandemic," *Project Syndicate*, 2000.

___ , "Third World Disease," *Ob.Gyn. News*, 2004.

Siegfried, N, M Muller, J Volmink, J Deeks, M Egger, N Low, H Weiss, S Walker, and P Williamson, "Male Circumcision for Prevention of Heterosexual Acquisition of HIV in Men," *Cochrane Database System Review*, 2003, *3*, CD003362.

Smiley, ML, GC White, P Becherer, G Macik et al., "Transmission of human immunodeficiency virus to sexual partners of hemophiliacs," *American Journal of Hematology*, 1988, *28* (1), 27–32.

TC, TC Quinn, MJ Wawer, N Sewankambo, D Serwadda, C Li et al., "Viral load and heterosexual transmission of human immunodeficiency virus type 1," *New England Journal of Medicine*, 2000, *342* (13), 921–929.

Thorne, C and ML Newell, "Mother-to-child Transmission of HIV Infection and Its Prevention," *Current HIV Research*, 2003, *1* (4), 447–62.

Voeten, HA, OB Egesah, MY Ondiege MY, CM Varkevisser, and JD Habbema, "Clients of female sex workers in Nyanza province, Kenya: a core group in STD/HIV transmission," *Sexually Transmitted Disease*, 2002, *29* (8), 463–464.

Wawer, M, N Sewankambo, and others., "Control of Sexually Transmitted Diseases for AIDS Prevention in Uganda: A Randomized Community Trial," *The Lancet*, 1999, *353*, 525–535.

Weiss, HA, MA Quigley, and RJ Hayes, "Male Circumcision and Risk of HIV Infection in Sub-Saharan Africa: a Systematic Review and Meta-Analysis," *AIDS*, 2000, *14*, 2361–2370.

Wojcicki, Janet, "COMMERCIAL SEXWORK OR UKUPHANDA? SEX-FOR-MONEY EXCHANGE IN SOWETO AND HAMMANSKRAAL AREA, SOUTH AFRICA," *Culture, Medicine and Psychiatry*, 2002, *26*, 339–370.

WorldBank, *Confronting AIDS: Public Priorities in a Global Epidemic*, Oxford University Press, 1997.

Table 1
Highest and Lowest Infection Countries

Highest		Lowest	
Botswana	33.8%	Madagascar	0.3%
Zimbabwe	33.7%	Senegal	0.5%
Swaziland	33.4%	Somalia	1.0%
Lesotho	31.0%	Guinea	1.5%
Namibia	22.5%	The Gambia	1.7%

Source: UNAIDS Fact Sheets, 2001

Table 2
HIV Transmission Probability

Citation	# of Subjects	M-to-F	F-to-M	Location
Gray et al (2001)	174	17.50%	27.30%	Uganda
CAPS Data	750	27.50%	7.10%	Kenya and Tanzania
Grosskurth et al (1995)	523	29.30%	14.50%	Tanzania
Quinn et al (2000)	415	26.7%	17.5%	Uganda

Table 3
Cost-Effectiveness of Interventions

	Bacterial STI Treatment	Bacterial and Herpes	Continuous Behavioral Change	Abrupt Behavioral Change
Cost	$1,036,360,546	$33,827,131,482	$2,800,263,630	$2,800,263,630
Life Years	272,247,804	549,261,076	108,605,974	140,361,412
DALYs	292,539,667	590,773,772	114,940,762	148,640,583
Infections Averted	14,315,441	29,324,685	4,284,996	5,599,324
Cost per LY	$3.81	$61.59	$25.78	$19.95
Cost per DALY	$3.54	$57.26	$24.36	$18.84
Cost per Infection	$72.39	$1,153.54	$653.50	$500.11

Table 4
Cost-Effectiveness of Combined Interventions

	Bacterial STI and Behavior	All STI Treatment and Behavior
Cost	$3,836,624,176	$36,627,395,111
Life Years	339,501,379	576,022,766
DALYs	363,799,916	619,159,726
Infections Averted	17,032,039	30,427,366
Cost per LY	$11.30	$63.59
Cost per DALY	$10.55	$59.16
Cost per Infection	$225.26	$1,203.76

Table 5
Possible 2002 Expenditures and Infections Averted

Intervention	Cost	Share Averted	Number Averted
Bacterial Only	$295,903,656	23.9%	835,478
Bacterial and Herpes	$9,530,502,637	52.1%	1,824,846
All Interventions	$10,330,039,294	52.8%	1,847,095

Figure 1:
Transmission Probability and HIV Rates, Analytic Model

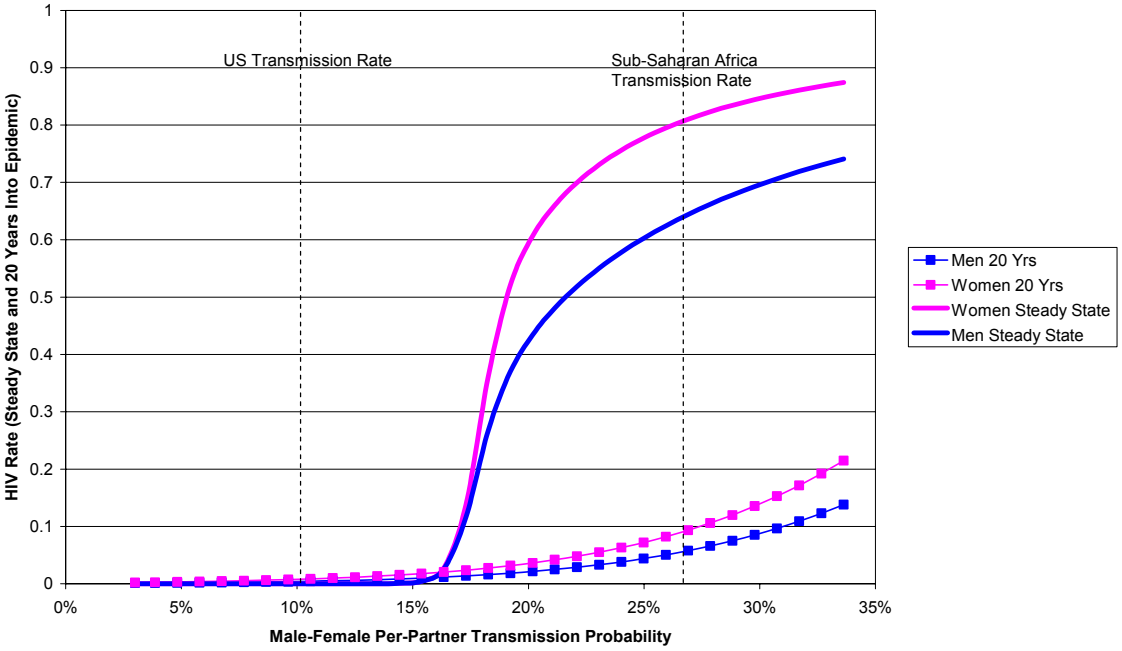


Figure 3: HIV Prevalence and Male Sexual Behavior

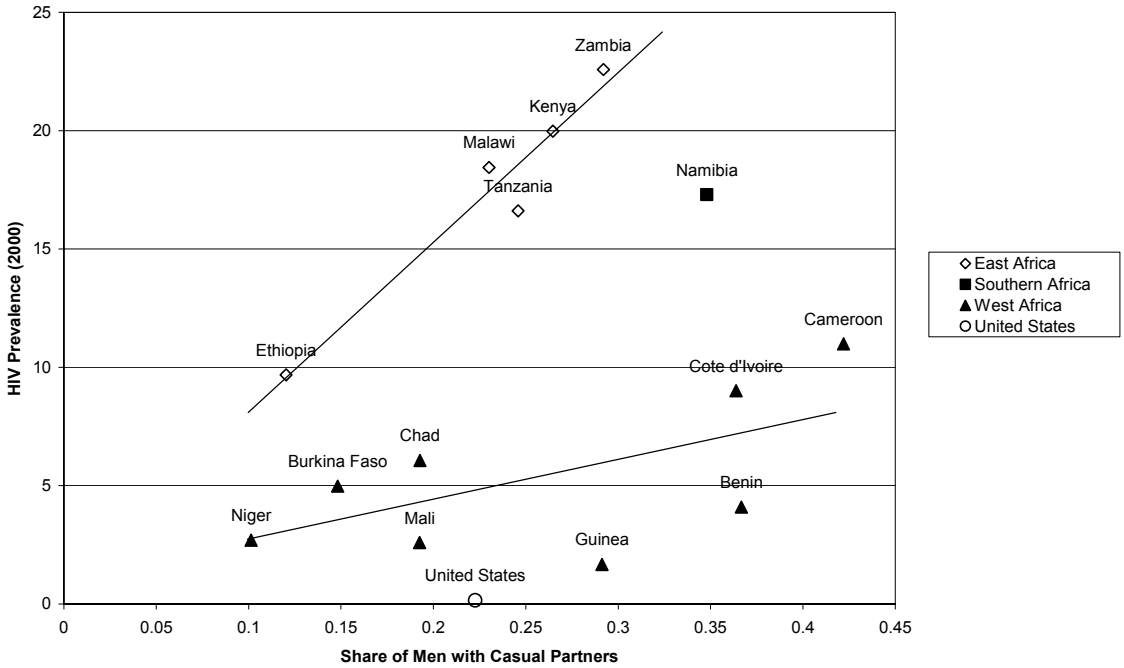
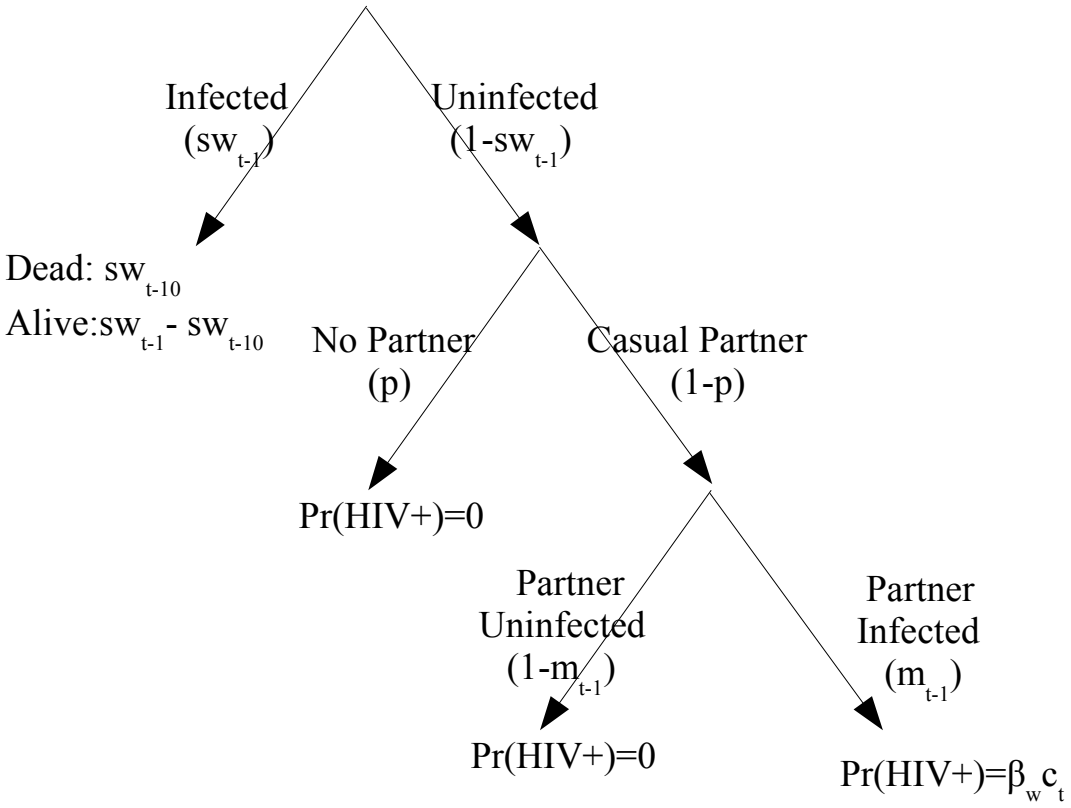


Figure 2a: Evolution of Infection Rates
Single Women with Zero or One Partner



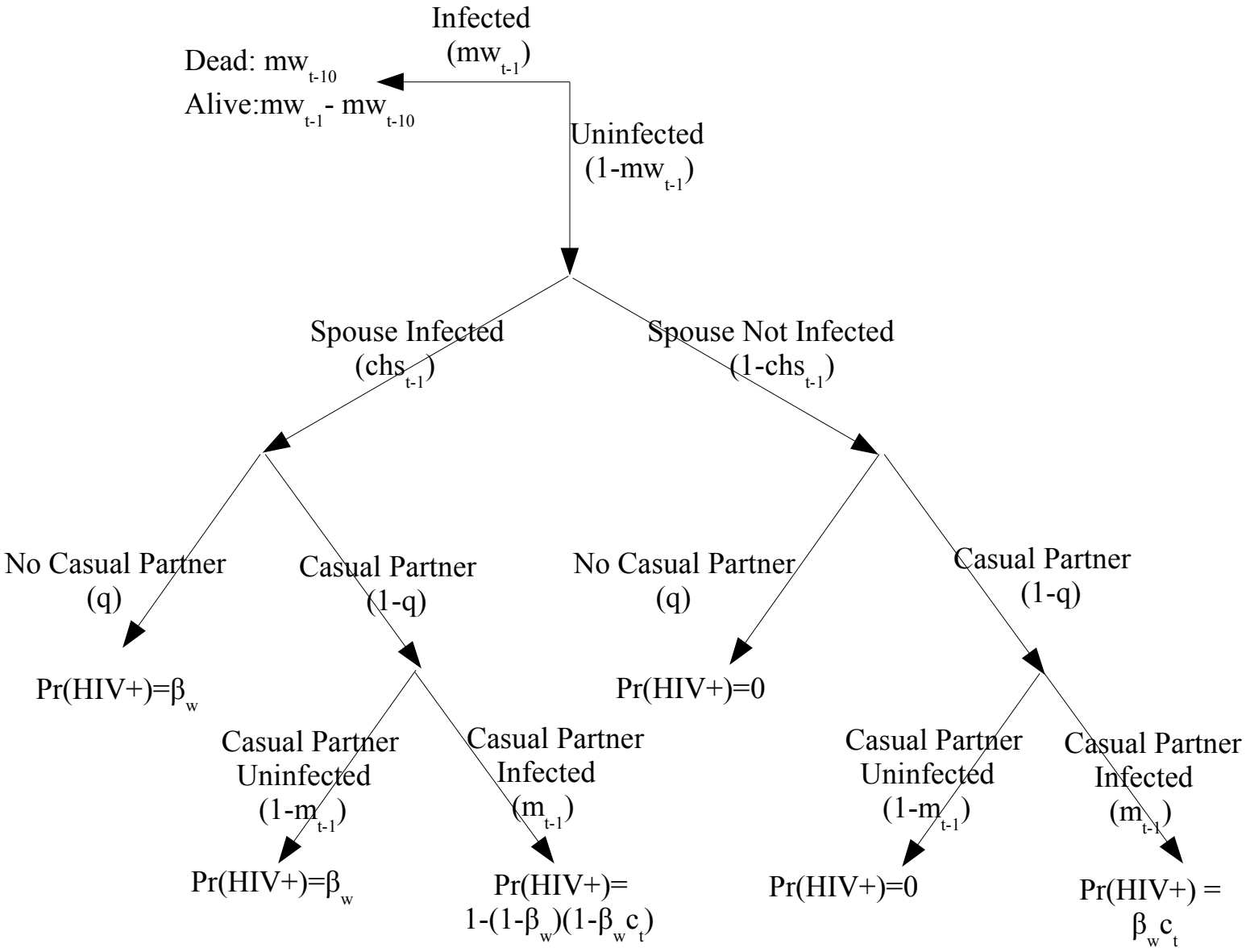
$Inc_t=(1-sw_{t-1})(1-p)\beta_w c_t m_{t-1}=(1-sw_{t-1})(1-p)(1-(1-\beta_w c_t m_{t-1})^1)$

$Prev_t=(1-sw_{t-10})^{-1}(sw_{t-1}-sw_{t-10}+(1-sw_{t-1})(1-p)(1-(1-\beta_w c_t m_{t-1})^1))$

Legend:

sw_t	infection rate for single women, time *t*
m_t	infection rate for men, time *t*
β_w	infection rate per unprotected partnership
c_t	share of partnerships without condoms, time *t*
p	share of individuals who have a casual partner
$Pr(HIV+)_t$	probability become HIV+ in period *t*
Inc_t	HIV incidence for this group in period *t*
$Prev_t$	HIV prevalence for this group in period *t*

Figure 2b: Evolution of Infection Rates
Married Women with Spouse and Zero or One Casual Partner



$Inc_t=(1-mw_{t-1})[(q)(chs_{t-1})(\beta_w)+(1-q)[1-(1-\beta_w chs_{t-1})(1-\beta_w c_t m_{t-1})]]$

$Prev_t=(1-mw_{t-10})^{-1}(mw_{t-1}-mw_{t-10}+(1-mw_{t-1})[(q)(chs_{t-1})(\beta_w)+(1-q)[1-(1-\beta_w chs_{t-1})(1-\beta_w c_t m_{t-1})]]$

Legend:

mw_t	infection rate for married women, time *t*
m_t	infection rate for men, time *t*
chs_t	chance spouse infected, time *t*
β_w	infection rate per unprotected partnership
c_t	share of partnerships without condoms, time *t*
q	share of individuals who have a casual partner
$Pr(HIV+)_t$	probability become HIV+ in period *t*
Inc_t	HIV incidence for this group in period *t*
$Prev_t$	HIV prevalence for this group in period *t*

Figure 4: HIV Prevalence Rate and Epidemic Start Date

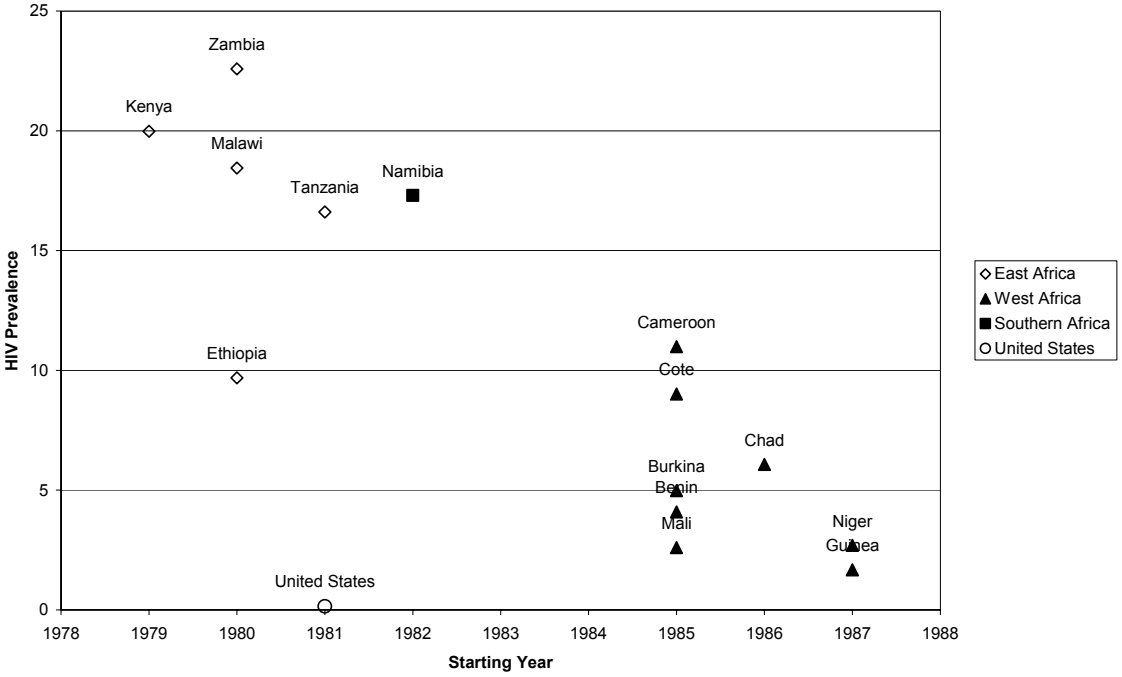


Figure 5: Condom Use and HIV Prevalence

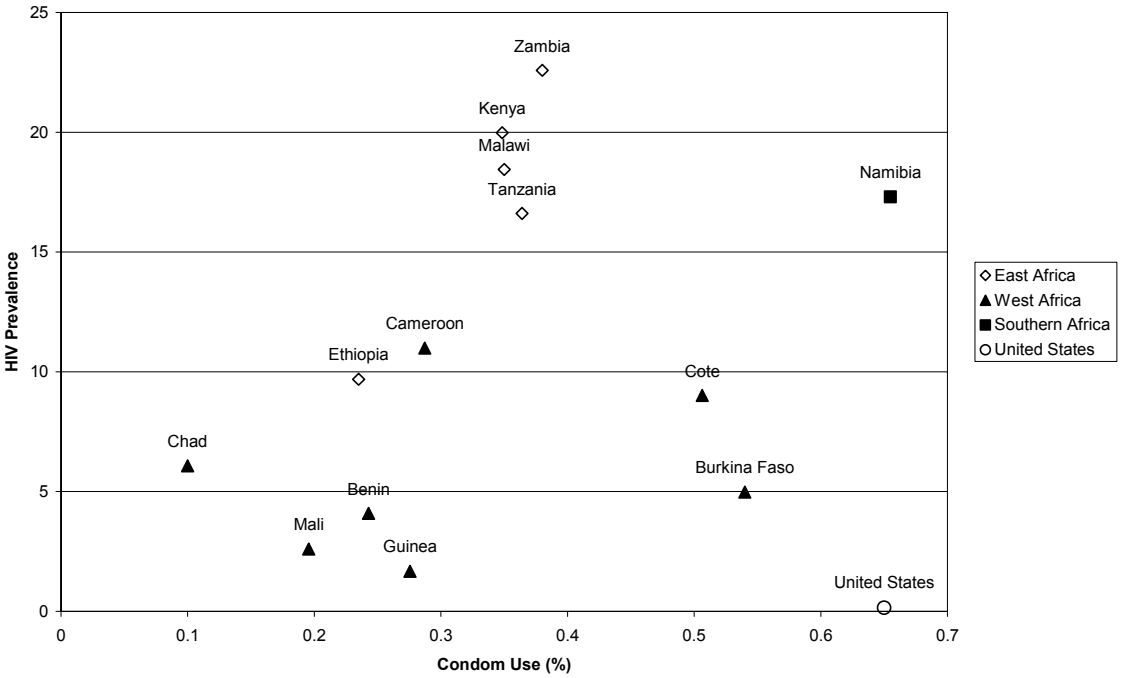


Figure 6: End of Decade Actual and Predicted HIV Rates

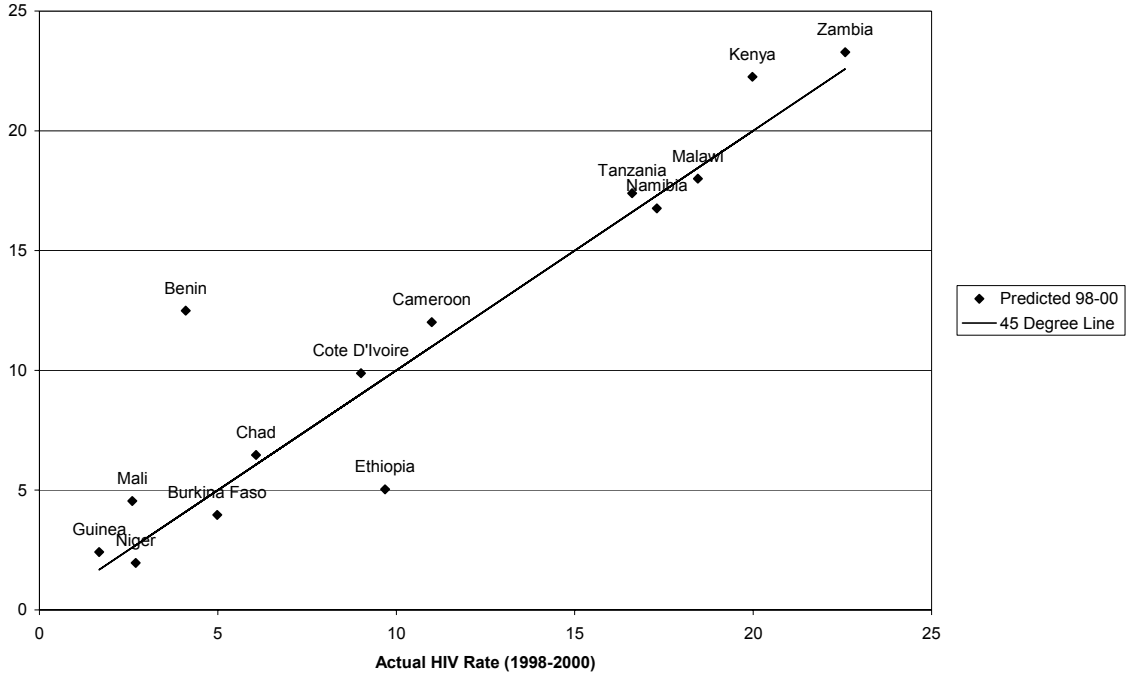


Figure 7: Actual and Predicted HIV Rates

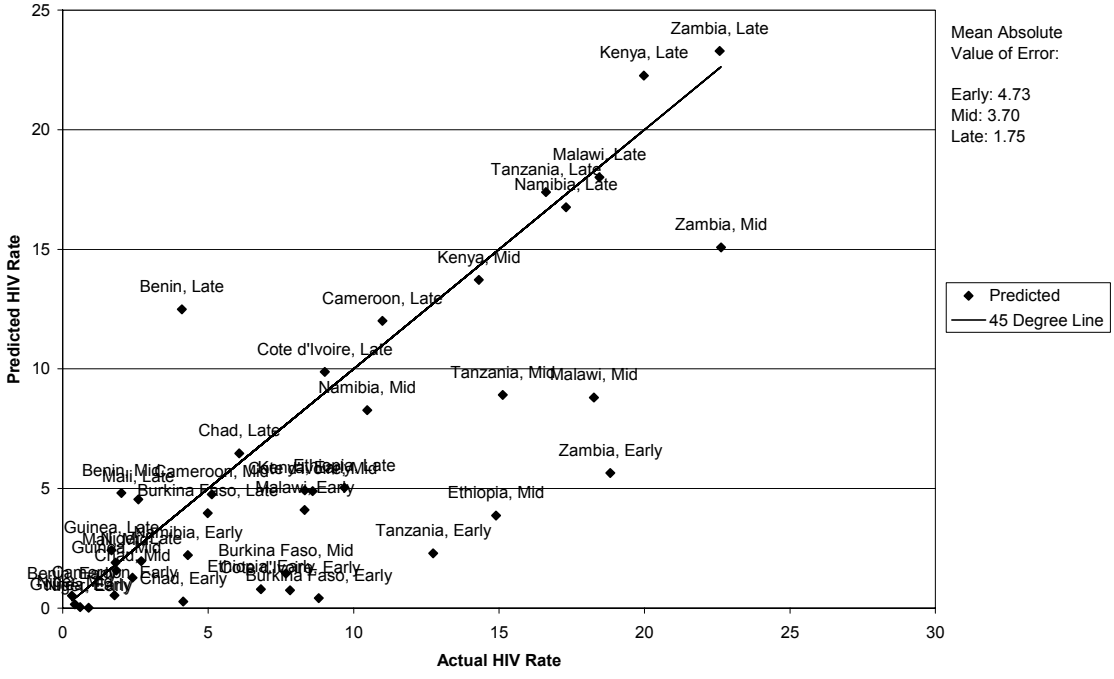


Figure 8: Predicted HIV Rate, Varying Transmission Rates

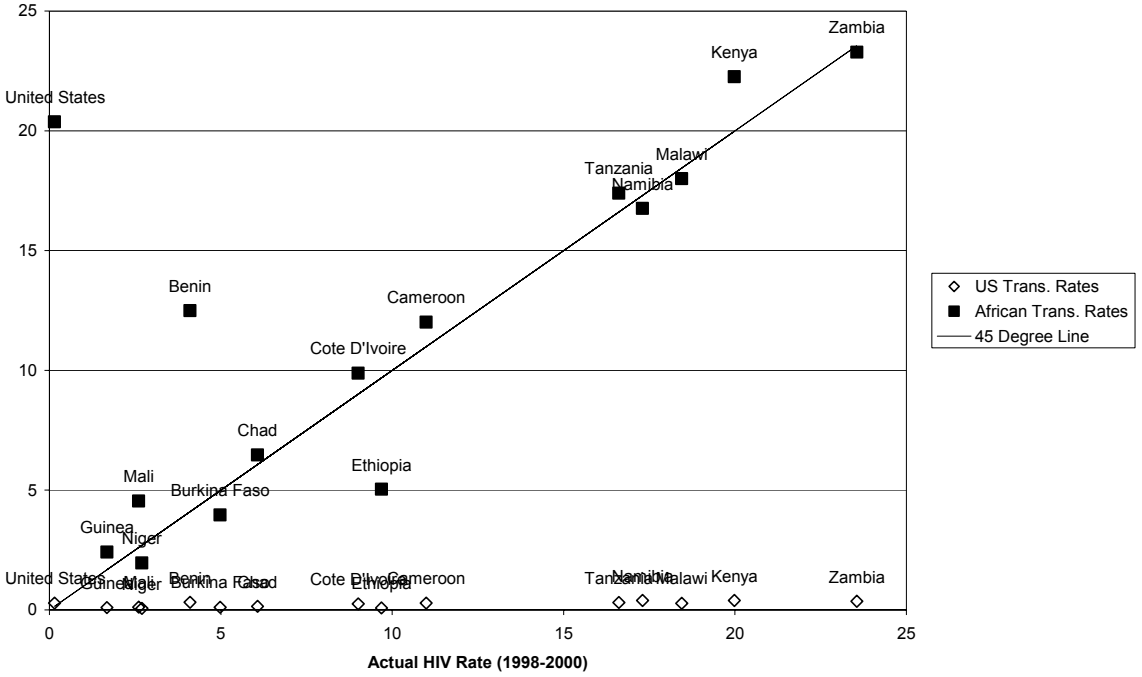


Figure 9: Average Infection in Sample Countries

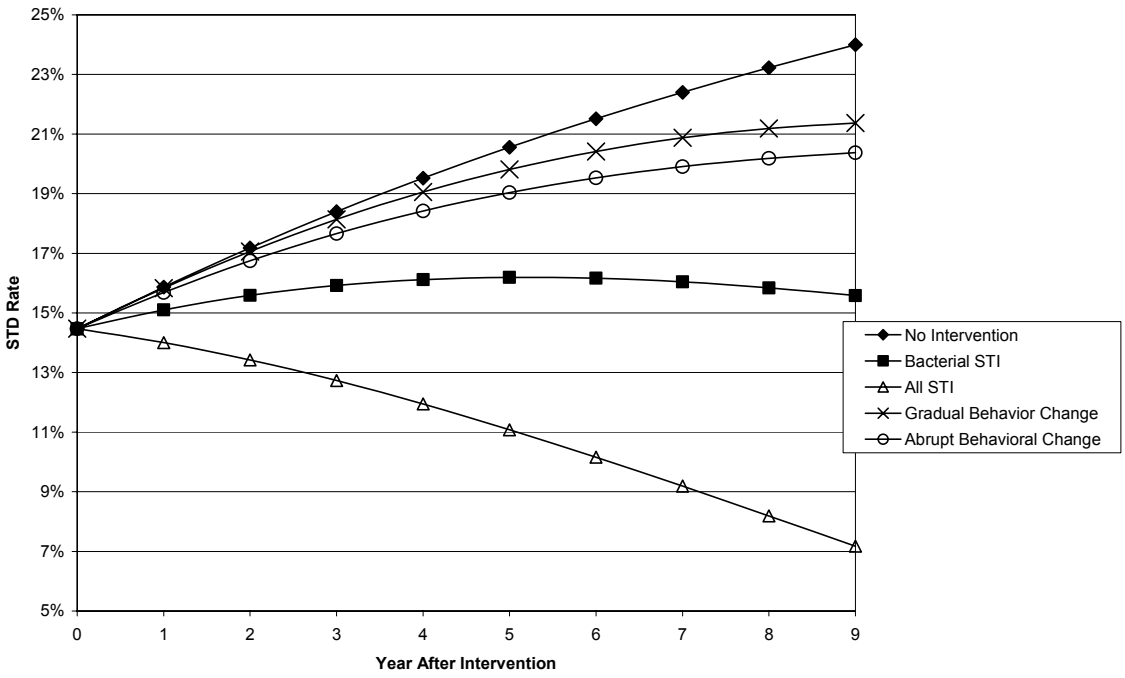


Figure 10: Time Path of Cost Effectivness

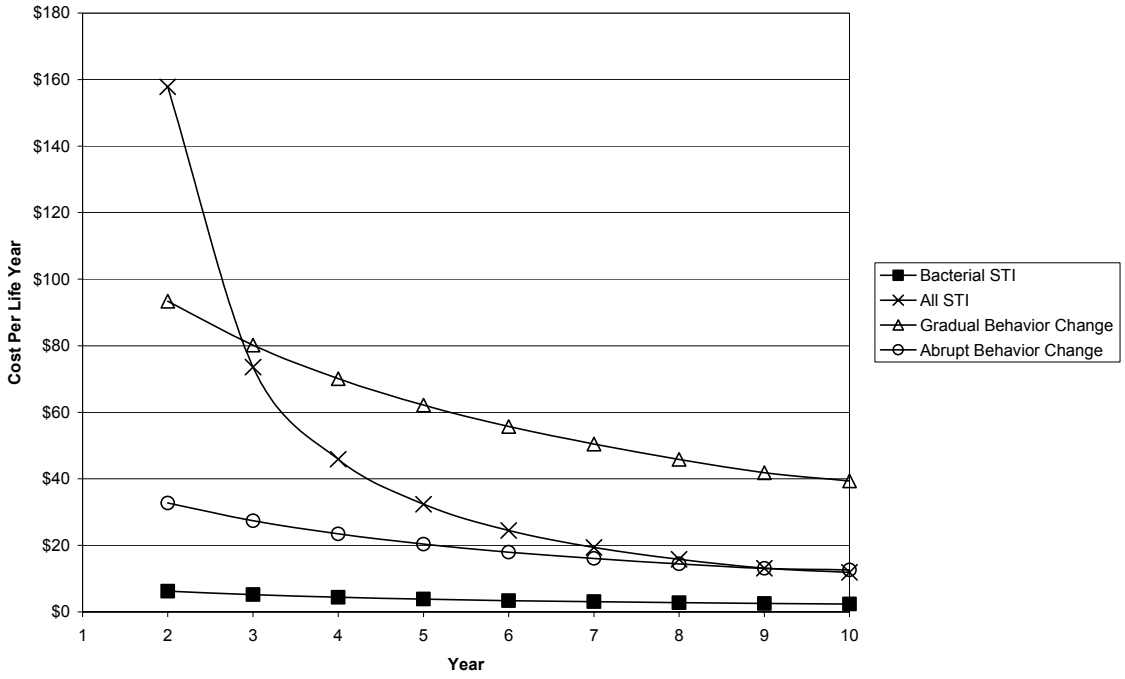
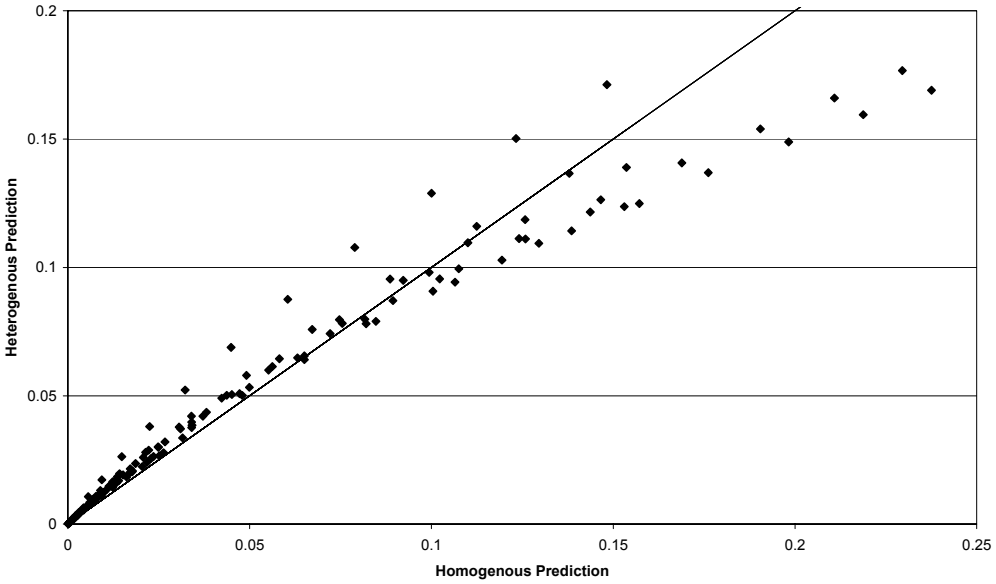
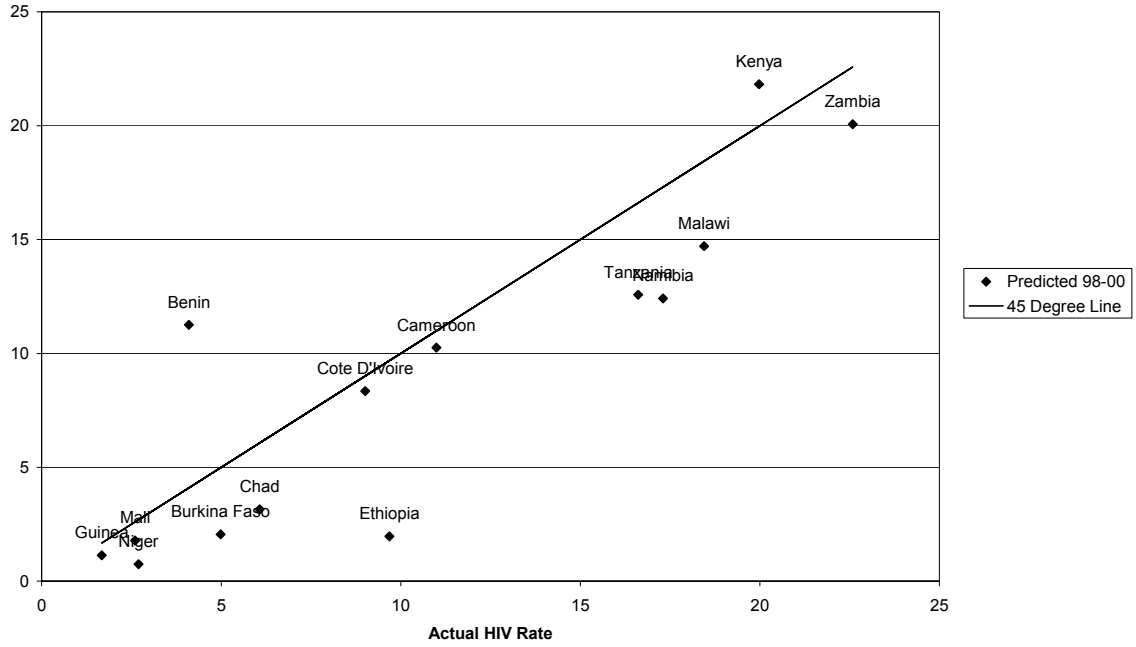


Figure 11:
Predicted Values from the Homogenaity and Heterogenaity Models



Appendix Figure A
Robustness Check: Simplified Homogenous Model



Appendix A: Calculation of Transmission Rates

Transmission rates are calculated from four prospective studies of HIV incidence. For three of them, data is taken from published work. In the case of the CAPS study, calculations are done with the data directly.

Information is recorded on the rate of sereoconversion ($SCR = \frac{numberofinfections}{originalnumberuninfected}$), the mean number of partners for individuals in the study ($PART$), the overall HIV rate in the population of the other gender (HIV) and rate of condom use ($COND$). The measure of interest is per-partnership transmission rate ($PPTR$). In all cases, we assume infection are Bernoulli in partnerships (Kaplan, 1990), yielding the following relationship:

$$SCR = (1 - (1 - PPTR \times COND \times HIV)^{PART})$$

The interpretation is that the rate of infection is equal to the chance that the individual with the median behavior becomes infected, adjusting for his or her chance of having protected sex, and the chance of having sex with an infected partner. The only unknown is the per-partnership transmission rate, implying that we can solve directly for that. In the case of the Gray et al (2001) study, the couples used are discordant and reported to be monogamous, so the calculation was somewhat more straightforward.

A weighted average of the estimates for the studies listed in Table 2 is calculated to get the overall estimates used in the analysis. The three studies other than the Gray et al (2001) study look quite similar. That study is somewhat different. One issue is that couples may not have been monogamous, and if men were less likely than their wives this could account for the discrepancy. It is also worth noting that the differences in that study are not statistically significant, and the study is quite small. These estimates do not impact the overall calculation very significantly.

Appendix B: Full Simulation Equation Specification

The model presented in section 4.1 can be summarized in the equations presented below. $SW_{a,t}$ denotes the HIV rate for single women of age a in time t, and $MW_{a,t}, SM_{a,t}, MM_{a,t}$ are similar monikers for married women, single men and married men; the HIV rate for female sex workers is denoted F_t. The share of individuals within each behavior group is denoted with $qi_{X,a}$, where i is the behavior group, X is either SW, MW, SM or MM and a is age. The number of partners for those with non-spousal partners is pj_X, where X is as above, and j is either c for casual partners or f for female sex worker partners (only for men). SPM_t and SPW_t are HIV rates for male and female spouses, respectively, in time t.

We note that in each year this entire set of equations will be simulate for each age group $a \in [15, 49]$.

$$SW_{a,t} = W_{a-1,t-1} + q1_{SW,a}(1 - W_{a-1,t-1})(1 - (1 - \beta_w M_t c_t)^{pc_{SW}}) \tag{5.1}$$

$$\begin{aligned} MW_{a,t} = {} & W_{a-1,t-1} + q1_{MW,a}(1 - W_{a-1,t-1})(1 - (1 - \beta_w M_t c_t)^{pc_{MW}}(1 - \beta_w SPM_t)) \\ & + (1 - q_{MWa})(1 - W_{a-1,t-1})(\beta_w SPM_t) \end{aligned} \tag{5.2}$$

$$W_{a,t} = \frac{\text{pct single}_a SW_{a,t} + \text{ pct married}_a MW_{a,t} - (W_{a-11,t-11} - W_{a-10,t-10})}{1 - W_{a-11,t-11} - W_{a-10,t-10}} \tag{5.3}$$

$$\begin{aligned} SM_{a,t} = {} & M_{a-1,t-1} + (1 - M_{a-1,t-1})[q1_{SM,a}(1 - (1 - \beta_m W_t)^{pc_{SM}}(1 - \beta_m pF_t c_t)^{pf_{SM}}) \\ & q2_{SM,a}(1 - (1 - \beta_m pF_t c_t)^{pf_{SM}}) + q3_{SM,a}(1 - (1 - \beta_m W_t c_t)^{pc_{SM}})] \end{aligned} \tag{6.1}$$

$$\begin{aligned} MM_{a,t} = {} & M_{a-1,t-1} + (1 - M_{a-1,t-1})[q1_{MM,a}(1 - (1 - \beta_m W_t c_t)^{pc_{MM}} \\ & (1 - \beta_m pF_t c_t)^{pf_{MM}}(1 - \beta_m SPW_t) + q2_{MM,a}(1 - (1 - \beta_m F_t c_t)^{pf_{MM}} \\ & (1 - \beta_m SPW_t) + q3_{MM,a}(1 - (1 - \beta_m W_t c_t)^{pc_{MM}}) \\ & (1 - q1_{MM,a} - q2_{MM,a} - q3_{MM,a})(\beta_m SPW_t) \end{aligned} \tag{6.2}$$

$$M_{a,t} = \frac{\text{pct single}_a SM_{a,t} + \text{ pct married}_a MM_{a,t} - (M_{a-11,t-11} - M_{a-10,t-10})}{1 - M_{a-11,t-11} - M_{a-10,t-10}} \tag{6.3}$$

Appendix C: Male Sexual Behavior Calculations

In general in the model, women can choose to partner with casual partners from the general population and with their spouse, if they are married. Men may choose to partner with a spouse if married, casual partners and female sex workers. The inputs to the model are the share of individuals choosing each set of partnerships and the average number of sexual partners for those in each non-marital sex grouping. The survey data, however, is limited and additional assumptions must be made to identify the parameters in the model. This appendix details how those parameters are identified. Here, we do the analysis for single individuals only. The analysis for married individuals is identical, but they must also partner with a spouse with certainty.

For women, the model requires us to identify the parameters in the following table. Obviously the values will vary by country.

Group	Share in Group	Average Non-Marital Partners for Group
None	$1-x$	0
Casual Partners	x	y

This is not an issue. Both x and y are observable – x is simply the share of women reporting sexual partners, and y is the average partners they report.

The analysis is somewhat more complex for men. The parameters that need to be identified appear in the table below.

Group	Share in Group	Average Non-Marital Partners for Group
None	$1-a-b-c$	0
Casual Partners	a	d
FSWs	b	e
Casual and FSWs	c	$d+e$

We make the assumption that people who partner with both casual partners and female sex workers partner with the same number of each as those who partner with only one group. What is observed in the data is $t = a+b+c$, the share of men reporting any casual sex. In addition, we observe the average number of casual partners for these men: $p = \frac{ad+be+c(d+e)}{a+b+c}$.

Following Carael et al (1995), Voeten et al (2002) and DHS survey data we estimate the share of men in each group. Consistent with these studies and data, we calculate that half the individuals who report non-marital sex are having sex only with casual partners, another 40% are having sex with casual partners and FSWs and the remaining 10% have sex with only FSWs.

In order to close the model, it must be the case that the number of male casual partners taken on by the women is the same as the number of female casual partners taken on by the men. Assuming equal sample sizes, this implies that $xy = ad + cd$. Solving for d and e yields:

$$d = \frac{xy}{a+c} = \frac{xy}{.5t}$$
$$e = \frac{pt-xy}{.5t}$$

The calculations above apply for all countries in Africa. For the U.S. we modify the parameters somewhat to reflect the lower use of FSWs. In particular, we assume that 70% of those with any non-marital partners are having sex with women from the overall population, 20% are having sex with women from the overall population and FSWs and 10% are having sex only with FSWs.