The Relationship between Health Care Provision and AIDS Prevalence in Sub-Saharan Africa

Craig McIntosh*
University of California at San Diego

PRELIMINARY DRAFT – PLEASE DO NOT CITE

April 2006

Abstract:

HIV/AIDS is a disease for which we ourselves are the vector. Consequently, a high prevalence of the disease in the population is likely to increase the rate at which new infections occur. This paper argues that in Sub-Saharan Africa, where the prevalence of other fatal diseases is high, there is a counter-intuitive effect of health care spending: such spending increases the life expectancy of the infected, and so drives up the prevalence of HIV/AIDS in the population. Because the growth rate of infection is found to be increasing in the prevalence, this implies that high-quality health care also increases the speed of spread of the disease, and is likely to drive up the peak prevalence that will be observed in the course of an epidemic. This depressing conclusion is derived from a simple theoretical model, and is confirmed using cross-country evidence on historical health spending and current HIV prevalence and rates of spread. The paper concludes by discussing the policy implications that flow from this result.

*Craig McIntosh: School of International Relations and Pacific Studies, U.C. San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0519, e-mail: *ctmcintosh@ucsd.edu*.
Thanks to Tara Ramanathan and Jacob Allen for excellent research assistance.

I. INTRODUCTION

The HIV/AIDS epidemic in Africa has been unique in several respects. The disease has moved through a different population (women) and a different transmission mechanism (heterosexual sex) than in most other regions of the world, and the peak prevalence rates seen there have been an order of magnitude higher than those observed elsewhere. Africa is also uniquely burdened by other endemic diseases; malaria, yellow fever, tuberculosis, and many others are constant scourges on the continent. Since it is virtually always secondary infections that ultimately prove fatal to those infected with HIV, the role of these other diseases in shaping the life expectancy of the infected is central. As we look across Sub-Saharan Africa, we see wide variation in the prevalence of these other diseases, and the incidence of these secondary diseases is highly correlated with low life expectancies and high mortality rates. HIV/AIDS, on the other hand, displays the reverse correlation: where mortality is highest and health care systems are at their worst, the AIDS epidemic has been substantially more muted in its progress. Indeed, the list of the highest prevalence countries in Africa is interchangeable with the list of countries that, as of 1990, did the most to promote the public health of their citizens: Botswana, Lesotho, Swaziland, Zimbabwe, Zambia, and South Africa.

Why should AIDS display a relationship with indicators of public health that is so different from other diseases? This paper suggests that the answer lies in a couple of simple facts about the disease. First, there may be limited direct efficacy of health care spending on the retardation of AIDS transmission. Numerous recent studies have shown that efforts to increase condom use, discourage sexual promiscuity, or to change behavior through testing have been only moderately successful (Wawer et al. (1999), Boozer & Phillipson (1999), Thornton (2006)). Only in mother-to-child transmission have interventions been efficacious. Hence we do not expect to see a better health care system delivering a dramatically lower rate of adult-to-adult transmission. Secondly, a simple epidemiological model shows that the rate of spread of a sexually transmitted disease is a function of the prevalence of the disease in the sexually active population. Because secondary diseases in Africa are likely to play a major role in increasing both morbidity and mortality from AIDS, they will have the effect of pushing the infected out of the sexually active population sooner than would have been the case in a disease-free environment. It is demonstrated below that the strongest correlate of peak female HIV prevalence, and the correlate most robust to the inclusion of other explanatory

variables, is female life expectancy as of 1982. Figure 1 plots the average prevalence for the longest- and shortest-lived quartiles in 1982, and we see an enormous difference between the subsequent trajectories of the disease. Hence our results suggest that good public health as of the beginning of the epidemic is a primary determinant of a severe subsequent AIDS epidemic.

This finding runs contrary to predictions from several strains of the economics literature on AIDS. Kremer (1996) argues that fatalism among sexually active individuals may cause them to reduce their activity by less than others, leading to an increase in the percentage of risky partners in the pool. He states that 'early public health efforts may allow societies to reach more favorable steady states'. Oster (2005) focuses on the role that sexually transmitted diseases play in the spread of HIV, and using simulations based on a model of sexual behavior and survey data from 14 African countries finds that 'differences across countries in Africa (in transmission rates) can be fully attributed to differences in risky sexual behavior and epidemic timing'. We fail to find any significant cross-sectional correlations between seroprevalence and condom use, and Figure 2 in the appendix shows the scatterplot of the maximum prevalence observed against the year in which prevalence first went above 5%; again no clear relationship is observable. While Young (2005) holds that there is indeed a 'gift of the dying', in his setup this benefit is realized through a posited decrease in fertility and the increase in resources per capita that accompany AIDS mortality.

The mechanism suggested here is more mechanical, simple, and macabre: the 'gift of the dying' is the removal of the potential infections that might be caused by that individual. While this is a result of most standard epidemiological models, I have found only a single direct reference to it; May & Anderson (1987) say that:

"The frequent assumption that the severity of the epidemic, in terms of cumulative mortality, will be greatest if all those infected eventually develop AIDS and subsequently die is not necessarily true. Mortality depends critically on the duration of infectiousness of both those infected who develop AIDS and those infected who do not. If the latter have a similar life expectancy to those not infected, but remain infectious for life, they may contribute more to the net transmission of the virus than those who die of AIDS".

By extension, if a policy intervention has the effect of extending the lifespan of the infected, this intervention will drive down mortality in the short run, but assuming any transmission of the disease to the uninfected, will increase the total number of people who eventually contract the disease. In what follows, we derive this result in 'closed-system' transitional model of epidemics as well as in a steady-state model of endemic disease with population

growth. The relationship is demonstrated using data from 32 African countries, and we conclude with a discussion of policy implications.

II. THEORETICAL MODEL

The effect that changes in mortality have on the prevalence and incidence of HIV/AIDS can be illustrated through the use of a model which describes the dynamic path of the disease. I use a simplified form of the model found in Palloni (1996). These dynamic equations, which are fit to East African data by Hueveline (2003), feature the age (a) and time (t) specific numbers of individuals who are uninfected, who are infected but have yet to develop the disease, and then those with full-blown AIDS. For parsimony I drop the intermediate category, and so we examine only the healthy $H(a,t)$ and those with AIDS $A(a,t)$. The differential equations which govern the motion of these two populations through time are:

$$\partial H(a+v,t+v)/\partial v = -(\mu_1(a,t)+\lambda(a,t))H(a,t)$$

$$\partial A(a+v,t+v)/\partial v = -\mu_2(a,t)A(a,t)+\lambda(a,t))H(a,t)$$

Here μ_1 is the mortality rate among the healthy, μ_2 is the mortality rate among those infected with AIDS, and λ is the rate at which the healthy contract the disease.

Because this model is primarily intended for estimation, there is no explicit modeling of the process through which infection occurs. For this, we modify the model used by Sweat et al (2000) and Bouey, Saidel and Rehle (1998). They give the probability of contracting AIDS over a discrete time period as:

$$1-\{P[1-R(1-F*E)*S]^N+(1-P)\}^M$$

where P is the prevalence rate among the sexual partners of healthy individuals, R is the probability of transmission per sexual act, F is the fraction of acts that use a condom, E is the efficacy of condoms, S is the likelihood of having sex, N is the number of sex acts within the period, and M is the number of sexual partners.[1]

In order to fit this model of infection into our dynamic equations, we assume that healthy individuals randomly partner with a member of their own age, and that a fraction

[1] To see values of these parameters estimated using data from Malawi, see Thornton (2006), and for per-act transmission probabilities in Rakai, Uganda, see Gray et al (2001).

$\theta(a,t)$ of the infected individuals in a given age cell remain sexually active.[2] Thus the prevalence rate P of a healthy individual's sexual partners will be

$$P(a,t) = \frac{\theta(a,t) * A(a,t)}{H(a,t) + A(a,t)}.$$

To simplify the problem we can drop the age/time cell notation, denote the effective probability of transmission per sex act as $\pi = R(1 - F * E) * S$, and assume that in each period each individual has one sex act with one random partner (so M=N=1), leaving us with

$$\lambda = 1 - \left\{ \frac{\theta A}{A+H}(1-\pi) + \left(1 - \frac{\theta A}{A+H}\right) \right\}$$

and so our equations of motion become

$$\partial H \Big/ \partial v = -(\mu_1 + \left(1 - \left\{ \frac{\theta A}{A+H}(1-\pi) + \left(1 - \frac{\theta A}{A+H}\right) \right\}\right))H$$

$$\partial A \Big/ \partial v = -\mu_2 A + \left(1 - \left\{ \frac{\theta A}{A+H}(1-\pi) + \left(1 - \frac{\theta A}{A+H}\right) \right\}\right)H$$

or, denoting the time derivative as $\dot{H}$ and collecting terms, we have

(1) $$\dot{H} = -\mu_1 H - \pi \frac{\theta A}{A+H} H$$

(2) $$\dot{A} = -\mu_2 A + \pi \frac{\theta A}{A+H} H.$$

All infection in this model comes from the sexual activity of those who are currently infected. Should θ fall to zero, meaning that there is no sexual activity among the infected, λ also becomes zero and transmission stops completely.[3] This is similar to the classic 'SIR' epidemiological model except that there is no recovered group when we discuss HIV.

So, what is the effect of the provision of health care in this environment? If high-quality health care were to alter the use or effectiveness of condoms, or decrease the

[2] It is a well-known feature of such models that random mixing leads to higher steady-state levels of infection than assortative matching models where individuals pair only with members of their own subgroup, see Garnet & Anderson, 1996.

[3] Note that this would require not only that all individuals who know they are infected remain totally abstinent, but that all individuals realize that they have become infected immediately. Given the very high infectiousness of individuals within their first few months of contracting HIV (Quinn et al, 2000), this is unlikely to be a realistic outcome with any level of testing and restraint.

likelihood of a sex act, then we would see a corresponding decrease in π, the probability of infection per sex act. In a randomized study, however, Thornton (2006) finds that the impacts of testing & incentives on condom use are minimal. In the same study, while some evidence is found that condom use increases slightly when individuals learn their HIV status, there is no evidence that their amount of sexual activity changes. Hence while some decreases in π or θ may be observed as a consequence of extensive counseling and testing, these effects are likely to be much smaller than self-selected samples would lead us to believe. The more immediate and concrete effect of good health care is likely to be a decrease in mortality.

It is of interest to study how the mortality rate among the infected, μ_2, influences the time path of prevalence. The instantaneous rate of change in the time derivative with respect to μ_2 is simply $-A$; meaning that as the rate of mortality increases, the change in the number of surviving infected individuals is proportional to the number of infected individuals. Once we examine changes over a discontinuous period of time, however, we see an additional effect enter the equation: a higher mortality rate leads to a lower prevalence rate, and assuming π and θ greater than zero, this leads to a lower *incidence* rate. This can be seen by examining the future change in prevalence in a discrete time model. In the first period, we would see only the direct effect of mortality on decreasing prevalence:

$$\frac{\Delta \dot{A}_t}{\Delta \mu_{t-1}} = -\Delta \mu_{t-1} A_{t-1}.$$

In the second period, however, we would begin to see the feedback effect of increased mortality:

$$(3) \quad \frac{\Delta \dot{A}_t}{\Delta \mu_{t-2}} = -\Delta \mu_{t-2} A_{t-1} - \theta \pi \frac{\Delta \mu_{t-2} A_{t-2}}{A_{t-2} + H_{t-2}} H_{t-2}$$

This tells us that the change in the prevalence (which equals incidence minus mortality) will be affected in two different ways by an increase in the mortality rate. The first is the direct effect through which a one unit increase in the mortality rate removes a number of individuals from the population of infected which is proportional to the population of infected. The second effect arises because as mortality among the infected increases, the number of sexually active people who are HIV positive also decreases, and hence the probability of new infections, or the incidence, decreases as well. This means that we expect more rapid increases (on the upward-sloping portion of the prevalence curve) and less rapid

decreases (on the downward-sloping portion) in countries that provide long life spans and low mortality for their infected populations, both because these individuals remain in the population to be counted for more years, and because the presence of any sexual activity among the infected ($\theta > 0$) caused a higher rate of infection among those who are currently healthy. This latter effect is solely due to sexual activity among the infected, and again should θ fall to zero the link between the prevalence and the incidence is broken.

Figure 1 in the appendix shows simulated results for a population of 1000 individuals where the initial prevalence rate is 10%, μ_1 is zero (meaning that population growth and mortality balance each other among the uninfected), $\theta = \pi = .5$, so the probability of contracting AIDS in any period is .25 times the prevalence rate. We see that when the mortality rate among the infected is 10%, the model converges rapidly to 100% infection. A mortality rate of 20% leads to a similar outcome but more slowly. Once the mortality rate has increased to 30% or 40%, however, die-off among the infected proceeds more rapidly than new infections, and so prevalence falls immediately from its initial level and asymptotes towards zero. The remaining number of the original 900 healthy individuals in the population after 40 periods in this simulation is 1, 77, 574, and 757 respectively, using the same mortality rate for the healthy across scenarios.

Because the preceding approach uses transition dynamics it is not tractable in terms of equilibrium conditions, and also is not informative about long-run equilibria. One way of writing down a model that generates a stable steady state is to introduce population growth into the healthy population, and then to make the 'mass action' or perfect mixing assumption; namely that the probability of contracting HIV is proportional to the *number* of interactions between the healthy and infected, which can be modeled as $\theta\pi AH$, with θ and π modified to take the appropriate values for this new formulation.[4] The rationale for this assumption would be that large numbers in both populations increase opportunities to match with new partners and hence the number of individuals who contract the disease. The value of this simplification is that it allows us to write down a dynamic system which is a special case of the Volterra-Lotka predator-prey model. Writing population growth among the healthy in the absence of the disease as $p - \mu_1$ (where *p* is the fertility rate), we have

[4] See Hethcote (2000) for an excellent summary of mathematical models of infectious diseases.

(4) $$\frac{\dot{H}}{H} = p - \mu_1 - \pi\theta A$$

(5) $$\frac{\dot{A}}{A} = \pi\theta H - \mu_2.$$

This is a standard formulation of the predator-prey model, except that because healthy and infected are measured in the same units, we impose that 'one wolf equals one sheep'. This problem has two solutions; an unstable equilibrium at (0,0) and a stable equilibrium at $(H, A) = \left(\frac{\mu_2}{\theta\pi}, \frac{p - \mu_1}{\theta\pi} \right)$. Figure XX shows the phase plane diagram of this system; the steady state is the intersection of the two dotted lines where $\dot{H} = \dot{A} = 0$, and healthy/infected numbers will move in a closed-orbit trajectory around this point.[5] From this picture it is clear that as μ_1 decreases, the steady state number of infected individuals increases. An increase in mortality among the infected has no effect on the steady state, and an increase in either π or θ leads to an initial increase in the number of infections but a decrease in the steady-state. Thus using two different classes of models we show that mortality decreases prevalence; in the first, mortality among the infected shortens duration of life and hence decreases the rate of infection, and the second model we see that the steady state level of infection falls as there is higher mortality among the healthy.

III. DATA.

In examining the relationship between health care systems and AIDS rates there is a natural bias which would emerge if we used prevalence estimates based on data collected by the national health care systems. Those health care systems which are the most extensive and well funded would likely detect the disease in a higher proportion of cases and hence appear to have a higher prevalence level. For this reason, it is crucial that we use an independent, objective data source which is collected in a consistent way across countries. The UNAIDS/WHO Working Group on Global HIV/AIDS and STI Surveillance provides such a source: this data conducts blood tests at maternity clinics across SSA and reports statistics for each testing site for each year.[6] Sophisticated epidemiological modeling tools exist which

[5] If we return to modelling the probability of infection as being a function of the share of infected in the population, this model generates a steady state to which the model converges, rather than featuring perpetual oscillation.

[6] http://www.who.int/GlobalAtlas/predefinedReports/EFS2004/EFS_PDFs/EFS2004_ET.pdf

allow for the simulation of incidence and prevalence rates. These tools, however, use substantial additional demographic information in forming these projections, and so regression of such simulated data on explanatory characteristics may lead to a circular statistical relationship.

For this reason, in this analysis I use the median seroprevalence level for rural and urban areas reported across the sentinel testing sites for each year. This data is somewhat noisy, and in some cases is missing intermediate values, and so in order to create as sound and continuous a set of outcomes as possible, I first linearly interpolate missing years for any year in which the preceding and following data is not missing, and then run a lowess smoother over the resulting data points to remove noise. These interpolated, smoothed data points are used for the analysis to follow, and in Figure 5 we see them, along with the original median levels, plotted for urban women for the countries used in the analysis. Any country that has fewer than 6 years of data is dropped from the analysis, leaving us with 32 countries total.

Table 1 gives the 1980 male and female life expectancies, the 1990 health care expenditures as a percentage of GDP, and the peak levels of rural and urban HIV prevalence observed for each country in the sample. Mortality and life expectancy are closely related, as we would expect, but the relationship with health care spending appears surprisingly weak. From this table it is visually clear that countries with longer life expectancies before the outbreak of the disease have higher maximum prevalence; the average urban HIV prevalence rate among the ten bottom-ranked countries in terms of life expectancies is 10, while in the ten top-ranked countries the average rate is 24. In what follows we undertake further examination of these correlations.

4. EMPIRICS.

4.1 CORRELATES OF MORTALITY

The theoretical model focuses our attention on mortality rates as a key driver of the trajectory of the AIDS epidemic. Data on life expectancies or mortality rates for HIV-infected individuals do not exist in Sub-Saharan Africa. We do, however, observe male and female mortality rates in 1980 and 1990, and it is probably safe to assume that high mortality in the population prior to the onset of the epidemic is likely to be a reasonable proxy for relative mortality among the infected once it begins. Table 2 in the Appendix shows the

results of unconditional pairwise OLS regression of 1990 mortality rates on a variety of other national characteristics.

In general, these correlations hold no surprises. We see the strong, structural correlation between mortality rates and life expectancies that we would expect. Health care provision seems generally to have been efficacious, as expenditures, clean water, doctors, and immunizations are all associated with lower mortality rates. Hospital beds and population density are uncorrelated with mortality, but per capita GDP displays a very strong negative correlation.

The bottom of Table 2 shows how 1990 mortality rates correlate with the maximum late-90's prevalence of non-HIV TB and malaria, and the maximum annual number of cholera deaths per country. The first sign of an unusual relationship between mortality and HIV comes from the fact that non-HIV TB is the only covariate in the data which displays a counterintuitive relationship to mortality; TB is worst in the countries that had the lowest mortality rates in 1990. Given that the resurgence of non-HIV TB in Africa in the 1990s was predominantly caused by the spread of AIDS, the implication of this correlation is that the spillover effects of HIV on other diseases were most muted in those countries with the highest mortality.

4.2 Prevalence of other diseases

We now examine in more detail the relationship between public health variables and the incidence of non-HIV diseases. There is potential endogeneity of placement and timing of health expenditures; if disease outbreaks cause the health infrastructure to ramp up operations then we have positive reverse causality. The same would be true for the direct effect of disease on mortality, again causing upward bias as a result of reverse causality. In order to address this problem we lag our covariates; 1980 mortality rates are used in order to ensure that cross-sectional differences in HIV were not themselves driving mortality rates. The earliest high-quality health expenditure data available is 1990; we use the 1990 public health expenditures as a percent of total government expenditure, as well as the 1995 expenditures as a percent of GDP.[7]

Table 3 shows the pairwise correlations that lagged mortality & health expenditures display with non-HIV TB, malaria, and cholera. In no case are these relationships

[7] Data from World Health Report 1999, Statistical Annex, WHO

significant, although we see that cholera outbreaks are somewhat more severe in countries where women had high mortality in 1980. One explanation for these weak relationships is that the long-term effects of health spending are minimal, whether because of a lack of duration in the effects or because the spending was ineffective in the first place. Another explanation is that the time path of these diseases did not fluctuate greatly, meaning that even the 1980 spending is endogenous to prevalence. This effect, if present, would increase spending in high-disease environments, pushing the sign of the relationship downwards and making spending appear less efficacious than it in fact is.

4.3 HIV/AIDS AND LAGGED PUBLIC HEALTH

When we move to examining correlations with HIV/AIDS, the picture changes dramatically. Table 4 gives the pairwise correlations using the same set of lagged public health outcomes and a variety of measures of the AIDS epidemic. We calculate mean and maximum prevalence and mean and maximum changes in prevalence separately for rural and urban areas for each country; mean prevalence is $E(A)$ from (2), and the mean change in prevalence is $E(\dot{A})$. 1980 Mortality displays an enormous negative correlation with subsequent AIDS prevalence, in both rural and urban areas. The lowest t-statistic for the relationship with prevalence levels is 4.4, and for prevalence changes is 3. While the relationships with health expenditures are less sharp, and not quite significant in explaining mean urban prevalence, in all other respects the relationship is significant, particularly in rural areas.

All of these correlations move in the opposite direction of what we should expect from an exogenous, efficacious health intervention, and all conform with the theory derived at the beginning of the paper. It is possible that the 1990s health care spending figures, being from years in which certain countries already had full-blown AIDS epidemics, are endogenous. While neither 1990 nor indeed 1995 had seen any large-scale response to the AIDS epidemic, the effect if present would explain the fact that those countries with the highest health care expenditures have the worst and fastest-spreading AIDS epidemics. The relationship with 1980 mortality, however, is both stronger and much harder to explain in terms of reverse causality. Because this is truly before the beginning of the epidemic, it is implausible that mortality was being driven directly by AIDS.

Changes in the prevalence rate equal the incidence rate minus the mortality rate. Average changes in prevalence, then should display a negative structural relationship with mortality which is not related to any possible effect of prevalence rates on incidence (the term on the right in (3)). Unfortunately, high quality panel data on incidence is lacking for countries in Sub-Saharan Africa, and so we are unable at this time to test this hypothesis directly. When we examine the relationship between the maximum increase in prevalence seen in a country and the contemporaneous prevalence, we see a very strong positive relationship (t-stats of 3.8 for urban women and 3.7 for rural women). While this relationship is consistent with high levels leading to high transmission, it does not allow us to decompose the effect of mortality on prevalence into the mechanical component and the component that works through transmission to the uninfected.

4.4 Conditional correlations

It is likely, of course, that other forms of heterogeneity which are correlated both with mortality rates and the trajectory of the epidemic exist. Wealthier countries with better health care may have bigger cities and thus more opportunity for the spread of the disease, and road networks, which would be denser in richer countries, may be spreading the disease more widely. Then there are a wide variety of income effects which may increase the capacity of individuals to undertake actions that spread the disease. With data from only 32 countries and no natural experiment, it is impossible to tease these effects apart in order to ascribe causality to any one channel. Further, the small-N and intercorrelations between covariates make multivariate regression sensitive to specification. Table 5 presents results for urban and rural areas when multivariate OLS in run using population density, road network density, per-capital PPP GDP, and female mortality as explanatory variables. There is some evidence that population density pushes up urban prevalence while a good national system of roads pushes up prevalence in rural areas. In all regressions, however, it is 1980 female mortality rates which are the strongest conditional correlate of female seroprevalence after some two decades (except for in urban areas where the relationship between income and the speed of spread of the disease is slightly stronger).

4.5 Difference Estimation

Because we have multiple observations for each country on several variables, panel identification of the relationships between changes in health determinants and changes in HIV are of interest. We have to be careful in setting up and interpreting these relationships, however, because once the epidemic had begun it exerted powerful causal forces on the time path of the health outcomes we have been using. In an attempt to construct a RHS variable which is truly exogenous to the subsequent trajectory of the disease, we calculate the change in life expectancy between 1972 and 1982. Correlations between HIV and this quantity are interesting because countries in the sample were experiencing almost linear increases in life expectancy over the decades prior to the onset of AIDS, whereupon they plummet in most countries by 20 years. Hence this linearized rate of growth over the decade prior to the onset of AIDS may provide a reasonable counterfactual for what would have obtained in the absence of the disease. We also use as a regressor the change in ppp GDP p/c from 1980 to 1990.

Table 6 reports the results of these changes on changes regressions. The top panel uses pairwise correlations, running a separate regression to estimate the correlation of each factor independently with changes in prevalence, while the bottom panel includes the three regressors simultaneously. The pairwise correlations are consistent with the predictions of our theory; changes in life expectancy are associated with more rapid subsequent increases in seroprevalence, although these effects are insignificant using changes in the UNAIDS discrete categories. The strongest effects in both panels are seen in the relationship between lagged increases in female life expectancy and subsequent increases in the rate at which HIV spreads in rural areas. Given the large lags between the changes on the RHS and those in the outcome, and the very small number of observations, it is in some ways surprising that such strong relationships should be seen.

5. Case Studies.

For the reasons described above, we cannot interpret contemporaneous changes in public health as 'shocks' off of which to identify the effects on HIV, because they are strongly endogenous to the path of the disease. We can, however, step away from the data in order to examine cases where radical changes in the public health environment occurred as a result of political shocks which we think were not caused by the AIDS epidemic.

5.1 ZIMBABWE & KENYA

These two countries provide case studies in the path of AIDS in economies that were relatively rich at the time the epidemic struck but which rapidly contracted thereafter. In Zimbabwe, Mugabe's erratic grip on the country reduced access to imported medical supplies and food, and its people saw a huge drop in standards of living and have left the country in large numbers. A recent study on the subject (UNAIDS, 2005) concludes that "The decline in national HIV prevalence between 2000 and 2004 resulted from a combination of declining HIV incidence and rising adult mortality occurring form the mid- and early- 1990s, respectively." Kenya's institutional decline was less precipitous, but the ppp dollar value of the decline in per-capita GDP was roughly the same.

Figure 6 plots the trajectory over time of smoothed urban HIV prevalence; in the early years the shape of the curve for Zimbabwe resembles that of Botswana, and similarly Kenya resembles South Africa. During the '90s, Kenya's ppp GDP p/c fell by $136, and Zimbabwe's by $144, while Botswana's increased by $1,795. In this way the subsequent trajectories of the disease provide a comparison between countries which are rapidly improving the standards of their people and those in which they are slipping backwards. When it comes to HIV, again, we see a worse-is-better response. The rapid increases in prevalence come to an end in contracting economies; both countries show a slight fall in incidence around the mid-1990s, whereupon rates appear to plateau (at 15% in Kenya and 30% in Zimbabwe). Prevalence in the growing economies, on the other hand, continues to climb for roughly twice as long, eventually reaching levels which are 15% higher. While these pairwise comparisons may be invalid, and causality is far from clear from this picture, the pattern is consistent with the 'difference in differences' intuition from our theoretical model.

5.2 UGANDA

Uganda provides a hopeful case study, although for a reason different from the one usually claimed. The countries of the Great Lakes region had similar female life expectancies in 1982, ranging from 48 in Rwanda to 52 in the DRC. The resolution of Uganda's long civil war by Museveni's ascension to power set the country on a path of stability and growth through the 1990s, during which it was one of the 10 fastest-growing economies in the world. Thus Uganda's AIDS epidemic began in a poor country and ended in a richer one.

Rwanda, Burundi, and the DRC, on the other hand, collapsed into chaos during this time. The DRC has by far the largest fall in ppp GDP p/c during the '90's, $877, and Rwanda fell by $96 and Burundi by $282 while Uganda increased by $393. This means that, according to the relationships seen elsewhere in the data, for a given set of initial conditions we would expect prevalence to be higher and fall more slowly in Uganda than in its neighbors. On the other hand, since the Great Lakes region was the first hit by HIV, and had very low levels of public health at that time, we would expect the epidemic to have peaked first there and fallen fastest there even in the absence of any effective AIDS policy. Thus Uganda should not be considered a success story if it has *decreasing* prevalence, but rather if it has prevalence which is not decreasing more slowly than these neighbors.

Figure 7 displays the smoothed prevalence trajectories for these countries, and we see that Uganda is a success by both measures. It is the case that the country sees sharp falls in prevalence rates, but then so do all of Great Lakes countries. What is more important is that the rate and total percentage of decrease is higher than in any of these countries despite the large relative increases in Uganda's standard of living. The conclusion is that despite changes in variables that should otherwise have increased incidence, Uganda has successfully implemented other policies which have slowed the rate of new infections. Whether it was able to influence sexual behavior among the infected (θ) or to decrease the rate of transmission (π), Uganda would seem to have been unique in its ability to combine rapid growth with decreases in prevalence. Indeed, only one other country, Burkina Faso, had increasing ppp GDP p/c during the '90's and also saw prevalence fall on average during that decade (the average fall in prevalence in Burkina Faso was a tenth of the average fall in Uganda). Since the A and B in Uganda's ABC speak to decreasing θ and the C to decreasing π, so we cannot identify the channel through which it worked. What we can say is that Uganda is a success not because it brought down AIDS rates (many failing states achieved this) but because it did so while growing. Our data provides no sign of the recent upturn which has seen scattered reports in the press.

V. CONCLUSION.

This paper connects two rather obvious propositions: that good public health policy reduces mortality, and that ceteris paribus high mortality hastens the course of an epidemic. By linking the two, the possibility is raised that Sub-Saharan Africa's high burden of endemic

disease has created an unintended consequence of a history of good public health. The data shows that nations which had previously done the best job of caring for their citizens are hit hardest by HIV/AIDS, and countries which had the worst prior public health see the epidemic roll through their populations more quickly, and with much lower peak prevalence rates. If African governments had more effective tools to combat AIDS directly, we might see that health care's direct negative effect on AIDS would be predominant; instead we see a positive relationship which is consistent with high background mortality suppressing the spread of AIDS. While peak female seroprevalence rates are positively correlated with a variety of measures of public well-being in the 1980s and 1990s, the most robust relationship appears to be with rural female life expectancy during those decades. While this relationship is an immediate result of a simple epidemiological model, it is nonetheless strongly counterintuitive in a policy sense.

Whether there is indeed any policy that flows from this observation depends on the process of transmission of the disease. We can see the relevant distinction in the equation that gives the change in the time path of the infection with respect to the mortality rate (equation (3)). The first term in this expression gives the mechanical effect of increased mortality on reducing the number of infected people who will be present in the population at the time of any seroprevalence test. If this term, and only this term, is non-zero, then in essence there is no policy conclusion to be drawn from this paper. The same point could be made of any disease, and it is repugnant to hope to reduce prevalence through the mortality of the infected. Even if only this mechanism is active, however, the results shown here suggest that low historical mortality rates may be a driving force in cross-country differences in the severity of the HIV epidemic in Africa. This is an important thing to know.

The second term in this equation of motion shows the way in which the incidence rate of the disease is linked to the prevalence. Two factors drive this rate of transmission: the share of the infected who remain sexually active, and the probability of contracting AIDS per sex act. If both of these quantities are non-zero, then a high prevalence rate, by increasing the share of the sexually active who are infected, leads directly to a higher prevalence rate. In this case, we see an exceedingly thorny policy problem emerge, in which actions taken to extend the lifespan of those suffering with HIV/AIDS lead in a direct causal sense to new infections. Such a policy, therefore, effectively trades off years of life among the currently infected against years of life among the currently uninfected.

The specific policy which is most pertinent to this macabre bargaining problem is the distribution of anti-retroviral drugs, or ARVs. In practice these drugs are usually distributed in concert with programs aimed at decreasing sexual activity among the infected (or decreasing θ) or increasing condom use and decreasing STDs (thereby decreasing π). Because such secondary policies may causally decrease the incidence rate, in some probabilistic sense they balance the negative externality imposed through the increase in the share of the sexually active who are infected. The logic of this paper argues that such programs, which might otherwise appear to be auxiliary to the distribution of ARVs, are in fact a moral imperative.

Several factors make this policy debate one fraught with agency problems. Firstly, the mechanism for this relationship is somewhat obscure, and the desire of any society to care for its members causes a natural response to do whatever possible to alleviate suffering here and now. Inadvertently, this condemns some number of the healthy to future infection, and the fact that these individuals are unknown at a moment in time makes their future suffering no less real. This diffusion and unpredictability of costs creates a second problem even in cases where domestic political processes decide these tradeoffs. There is a strategic analogy to debates over trade barriers: because one party to the debate is a focused, extant entity while the other is diffuse, political processes naturally favor groups that can organize over those that cannot. Finally, due to the high hard-currency expenses, foreign donors are playing a disproportionate role in providing ARVs rather than domestic governments. Hence it may be that the onus of deciding this tradeoff will fall on principals such as PEPFAR which have little accountability to the developing world.

Seen through this bleak lens, Uganda provides a positive example in a manner somewhat different from the usual interpretation. Uganda is a Great Lakes country that was poor at the time the epidemic took off, and so it should come as no surprise that the country has seen a fall in prevalence rates (Rwanda, Burundi, and the DRC all saw large decreases in prevalence at the end of the '90's, and none of these countries can be considered to have been preoccupied by AIDS prevention during that decade). What is impressive is the fact that Uganda is also a country that has seen an enormous relative increase in well-being during the course of the epidemic, and this improvement seems not to have shown up through higher eventual incidence. Uniquely among rapidly-growing economies, Uganda has seen falls in prevalence which appear almost linear; from 1991 to 2002 prevalence fell by 1.6

percentage points every year. The fact that this took place during a period of rapidly improving public health provision suggests that, if the determinants of infection can be tackled head on, it is indeed possible to achieve improvements in health outcomes without provoking a corresponding increase in prevalence.

The idea that prolonging the life of those suffering with AIDS could lead to more infections is at once relatively straightforward and extremely distasteful. The fact that it has not been made more forcefully before is presumably due to two factors: first, to even discuss the idea appears to degrade the infected by representing them as vectors. Secondly, hastening mortality is clearly a completely unacceptable policy conclusion, and a great deal of attention has already been focused on trying to decrease incidence. However, failing to discuss the relationship between prevalence and incidence does not remove the connection, and particularly at a moment when outsiders seem to be mobilizing to provide ARVs, the discussion is a necessary one. It would be flatly unethical, as well as counterproductive in the long run, to do nothing but distribute these drugs. Along with ARVs, which are clearly needed in order to alleviate the unimaginable suffering of the millions of infected Africans, must come other policies aimed at decreasing sexual activity and transmission of the virus. The experience of Uganda suggests that it is possible simultaneously to improve public health and lower AIDS rates, but both theory and evidence suggest it is difficult to do. Needed now are more well-designed policy studies which allow us to compare the efficacy of different interventions and hence to resolve the basic conundrum outlined here.

REFERENCES

Brauer, F. (1990) "Models for the spread of universally fatal diseases", *Journal of Mathematical Biology,* Vol 28, pp. 451-462.

Castillo-Chavez, C., K. Cooke, W. Huang, and S.A Levin. (1989) "On the role of long incubation periods in the dynamics of acquired immunodeficiency syndrome AIDS)", *Journal of Mathematical Biology,* Vol 27, pp. 373-398.

Evans, D, and T. Miguel. (2005) "Orphans and schooling in Africa: A longitudinal analysis", Working paper.

Garnet, G.P and R.M. Anderson. (1996) "Sexually transmitted diseases and sexual behavior: insights from mathematical models", *Infectious Diseases,* Oct;174 Suppl 2:S150-61,

Gray, RH, Wawer MJ, Brookmeyer R, Sewankambo NK, Serwadda D, Wabwire-Mangen F, Lutalo T, Li X, vanCott T, Quinn TC (2001) "Probability of HIV-1 transmission per coital act in monogamous, heterosexual, HIV-1-discordant couples in Rakai, Uganda", *Lancet,* Apr;357(9263):1149-53

Hethcote, H. (2000). "The mathematics of infectious diseases", *SIAM Review,* Vol. 42, No. 4, pp. 599-653.

Hueveline, P.., (2003) "HIV and population dynamics: A general model and maximum-likelihood standards for East Africa", *Demography,* Vol 40, No. 2, pp. 217-245.

May, R., and R. Anderson, (1987). "Transmission dynamics of HIV infection", *Nature,* Vol. 32 pp. 137-142.

Oster, E., (2005) "Sexually transmitted infections, sexual behavior change and the HIV/AIDS epidemic", *Quarterly Journal of Economics,* 120(2), pp. 467-515.

Palloni, A., (1995) "The Demography of HIV/AIDS", CDE Working Paper No. 95-21, Center for Demography and Ecology, University of Wisconsin – Madison.

Quinn, TC, MJ Wawer, N Sewankambo, D Serwadda, C Li, F. Wabwire-Mangen, M Meehan, T Lutalo, and R. Gray., for the Rakai Project Study Group, (2006) "Viral Load and Heterosexual Transmission of Human Immunodeficiency Virus Type 1", *New England Journal of Medicine,* Vol 343, No 13, pp. 921-929.

Sweat, M., S. Gregorich, G. Sangiwa, C. Furlonge, D. Balmer, C Kamenga, O. Grinstead, and T. Coates, (2000) "Cost-effectiveness of voluntary HIV-1 counselling and testing in reducing sexual transmission of HIV-1 in Kenya and Tanzania", *Lancet*, Vol 356 pp. 113-21.

Thornton, R., (2006) "The demand for and impact of learning HIV status: Evidence from a field experiment", Mimeo, Harvard University.

UNAIDS., (2005) "Evidence for HIV decline in Zimbabwe: A comprehensive review of the epidemiological data", UNAIDS/05.26E.

Young, A., (2005) "The gift of the dying: the tragedy of AIDS and the welfare of future African generations", *The Quarterly Journal of Economics,* Vol. CXX Issue 2, pp. 423-466.

Appendix.
Table 1: Summary Statistics.

Data Ranked by 1982 Female Life Expectancy.

Country	Female Life Expectancy, 1982	Female Mortality rate/10,000 1990	Health Exp as % of GDP, 1995	Peak Urban Prevalence	Peak Rural Prevalence
Comoros	51	.	0.9	0.1	0.0
Senegal	48	409	2.5	0.9	1.0
Niger	41	453	1.6	2.0	3.4
Benin	52	397	1.7	3.3	3.6
Ghana	56	279	1.4	3.7	3.7
Gabon	51	387	0.5	4.0	4.7
Nigeria	48	516	0.3	4.3	5.2
Guinea-Bissau	41	507	1.1	4.7	.
Guinea	41	517	1.2	5.0	2.4
Mali	44	458	1.2	5.6	5.2
Dem. Republic of Congo	52	375	0.2	6.6	5.4
Sierra Leone	37	527	4.3	6.9	.
Congo	54	.	3.2	7.1	10.0
Burkina Faso	46	349	2.3	7.9	7.9
Cameroon	53	415	1	8.5	7.9
Angola	44	401	4.1	8.6	8.0
Côte d'Ivoire	52	413	1.4	12.4	9.2
United Rep. of Tanzania	53	370	2.5	13.6	12.3
Central African Republic	50	424	1.9	15.0	15.0
Kenya	58	339	1.6	15.3	15.5
Togo	52	346	1.2	16.5	4.6
Mozambique	46	361	4.6	17.4	15.2
Ethiopia	44	362	1.6	19.2	15.5
Burundi	49	400	0.8	22.8	9.1
Malawi	46	362	2.3	25.2	18.7
Uganda	49	395	1.8	26.6	12.3
Namibia	55	232	3.7	27.4	21.9
South Africa	61	.	3.6	27.7	26.3
Rwanda	48	453	1.9	29.1	7.8
Zambia	52	.	2.6	29.6	17.6
Zimbabwe	58	321	2.2	32.0	33.5
Swaziland	55	366	2.8	36.5	38.5
Lesotho	56	334	4.1	38.2	27.0
Botswana	62	278	1.6	42.9	36.6

Table 2: Pairwise Correlates of 1990 Mortality.

Outcome:	Female 1990 Mortality	Male 1990 Mortality	nobs
Male Life Expectancy, 1982	-10.951	-12.442	
	-(5.37)	-(6.83)	28
Female Life Expectancy, 1982	-10.551	-11.628	
	-(6.09)	-(7.35)	28
Health exp as % of GDP, 1990	-28.620	-33.901	
	-(1.62)	-(2.01)	23
% with good drinking water, 1990	-1.668	-2.143	
	-(1.83)	-(2.25)	23
% immunized for DPT, 1990	-1.327	-1.510	
	-(2.16)	-(2.58)	26
% immunized for measles, 1990	-1.640	-1.815	
	-(2.27)	-(2.62)	26
hospital beds per 1000, 1990	-0.282	-4.005	
	-(0.01)	-(0.19)	20
doctors per 1000, 1990	-441.810	-874.069	
	-(0.90)	-(1.80)	16
% with improved sanitation, 1990	-0.924	-1.471	
	-(0.98)	-(1.49)	23
Population density, 1980	0.251	0.285	
	(0.79)	(0.87)	27
p/c GDP PPP, 1990	-0.022	-0.025	
	-(2.31)	-(2.68)	28
Non-HIV TB, max prevalence	-115.406	-117.425	
	-(1.59)	-(1.57)	28
Malaria, max prevalence	0.548	0.317	
	(0.79)	(0.44)	28
Cholera deaths, max prevalence	2621.293	1674.099	
	(2.62)	(1.52)	27

(t-statistics in parentheses)

Table 3: Relationship between Lagged Health Outcomes and Other Diseases.

Outcome:	Non-HIV TB prevalence, max	Malaria prevalence, max	Annual Cholera deaths, max	nobs:
Female Mortality, 1980	-0.001	-0.032	0.000	30
	-(1.10)	-(0.56)	(1.84)	
Male Mortality, 1980	-0.001	-0.039	0.000	30
	-(1.38)	-(0.65)	(0.74)	
Health care as % of gov expenditures, 1990	0.048	2.939	-0.001	27
	(1.06)	(0.63)	-(0.38)	
Health Expenditures as % of GDP, 1995	0.017	1.632	-0.001	33
	(1.06)	(0.63)	-(0.38)	

(t-statistics in parentheses)

Table 4: Relationship between Lagged Health Outcomes and HIV/AIDS.

Urban Prevalence	HIV prevalence, average	HIV prevalence, maximum	Annual change in prevalence, average	Annual change in prevalence, maximum	nobs:
Female mortality, 1980	-0.086	-0.113	-0.007	-0.016	30
	-(4.46)	-(4.66)	-(3.02)	-(3.88)	
Male mortality, 1980	-0.093	-0.121	-0.008	-0.018	30
	-(4.81)	-(4.97)	-(3.44)	-(4.31)	
Health care as % of gov expenditures, 1990	3.892	6.384	0.632	0.895	28
	(1.88)	(2.42)	(2.98)	(2.09)	
Health Expenditures as % of GDP, 1995	1.906	3.585	0.417	0.566	34
	(1.42)	(2.10)	(3.10)	(2.11)	
Rural Prevalence					
Female mortality, 1980	-0.082	-0.121	-0.010	-0.020	28
	-(5.19)	-(5.72)	-(4.97)	-(4.89)	
Male mortality, 1980	-0.077	-0.113	-0.009	-0.017	28
	-(5.29)	-(5.74)	-(4.70)	-(4.26)	
Health care as % of gov expenditures, 1990	3.900	5.877	0.558	0.881	27
	(2.44)	(2.56)	(2.56)	(2.27)	
Health Expenditures as % of GDP, 1995	2.771	4.213	0.312	0.880	32
	(2.67)	(2.85)	(2.09)	(3.63)	

(t-statistics in parentheses)

Table 5: Multivariate Regression.

	Urban Prevalence		Rural Prevalence	
	Maximum prevalence	Maximum change in prevalence	Maximum prevalence	Maximum change in prevalence
Population density, 1980	0.079	-0.006	-0.014	0.002
	(1.72)	-(0.88)	-(0.41)	(0.28)
Road network, 1990	0.111	0.036	0.232	0.033
	(0.58)	(1.33)	(1.60)	(1.17)
p/c GDP 1990	0.002	0.000	0.001	0.000
	(1.36)	(2.01)	(1.21)	(0.80)
Female mortality, 1980	-0.085	-0.008	-0.091	-0.016
	-(2.85)	-(1.91)	-(3.55)	-(3.20)
nobs:	27	27	25	25

(t-statistics in parentheses)

Table 6: Estimation using Changes on Changes.

Pairwise Correlations, Changes on Changes:

Outcome:	Ordered Probit on UNAIDS discrete prevalence categories	Average Change in Urban Prevalence	Average Change in Rural Prevalence
Change in GDP p/c ppp, 1980-90	0.0001	0.0004	0.0004
	(0.34)	(1.75)	(1.71)
Change in male life expectancy, 1972-82	0.131	0.195	0.216
	(0.78)	(1.84)	(1.94)
Change in female life expectancy, 1972-82	0.089	0.207	0.220
	(0.64)	(2.30)	(2.38)
nobs:	25	34	32
	(z-statistics)	(t-statistics in parentheses)	

Multivariate OLS, Changes on Changes:

Outcome:	Ordered Probit on UNAIDS discrete prevalence categories	Average Change in Urban Prevalence	Average Change in Rural Prevalence
Change in GDP p/c ppp, 1980-90	0.0002	0.0003	0.0002
	(0.74)	(1.35)	(1.15)
Change in male life expectancy, 1972-82	-0.490	-0.197	-0.194
	-(1.24)	-(0.86)	-(0.87)
Change in female life expectancy, 1972-82	-0.143	0.316	0.409
	-(0.50)	(1.55)	(2.18)
nobs:	20	25	23
	(z-statistics)	(t-statistics in parentheses)	

Figure 1

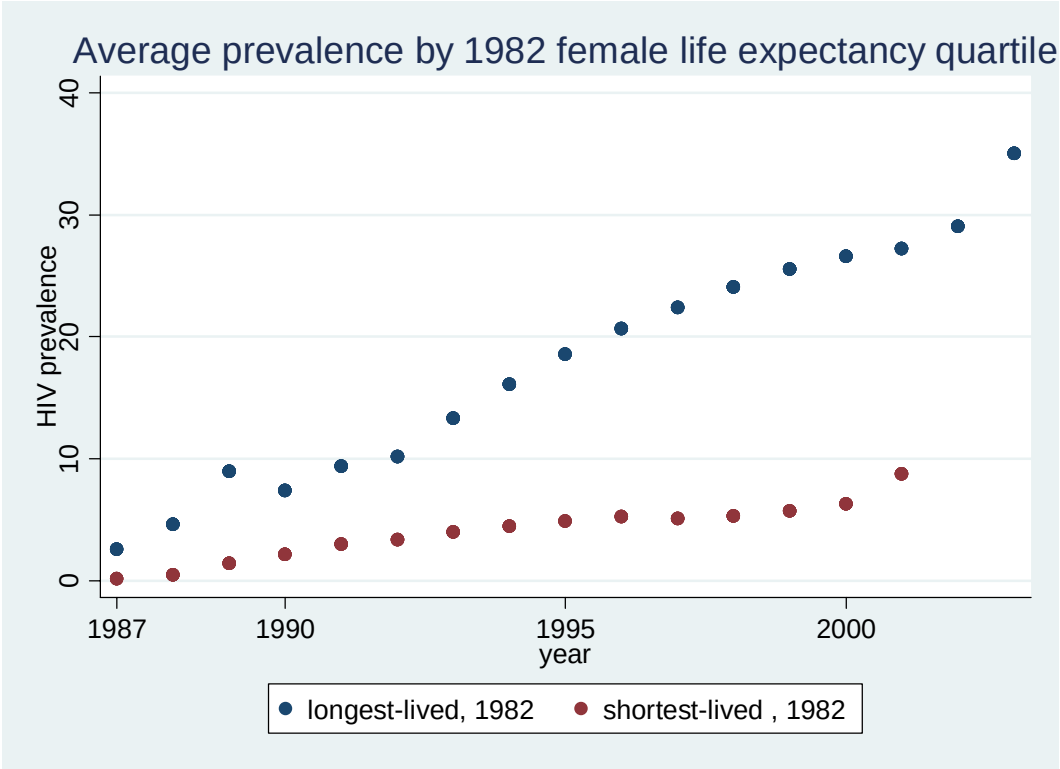

Figure 2.

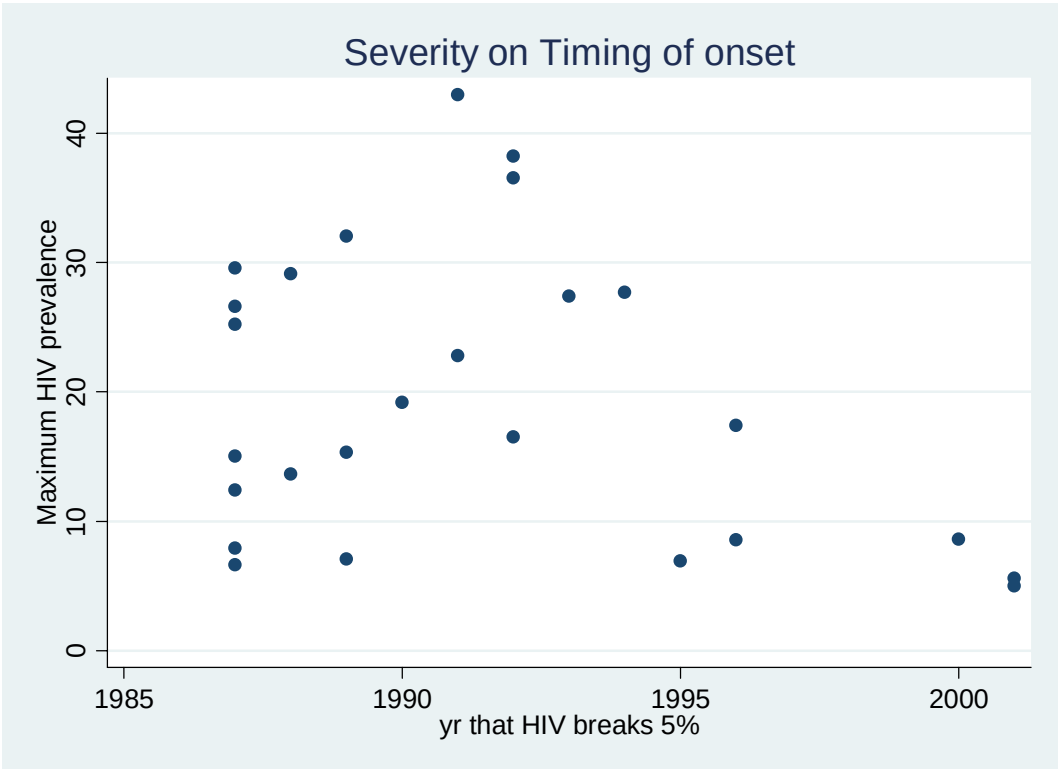

Figure 3.

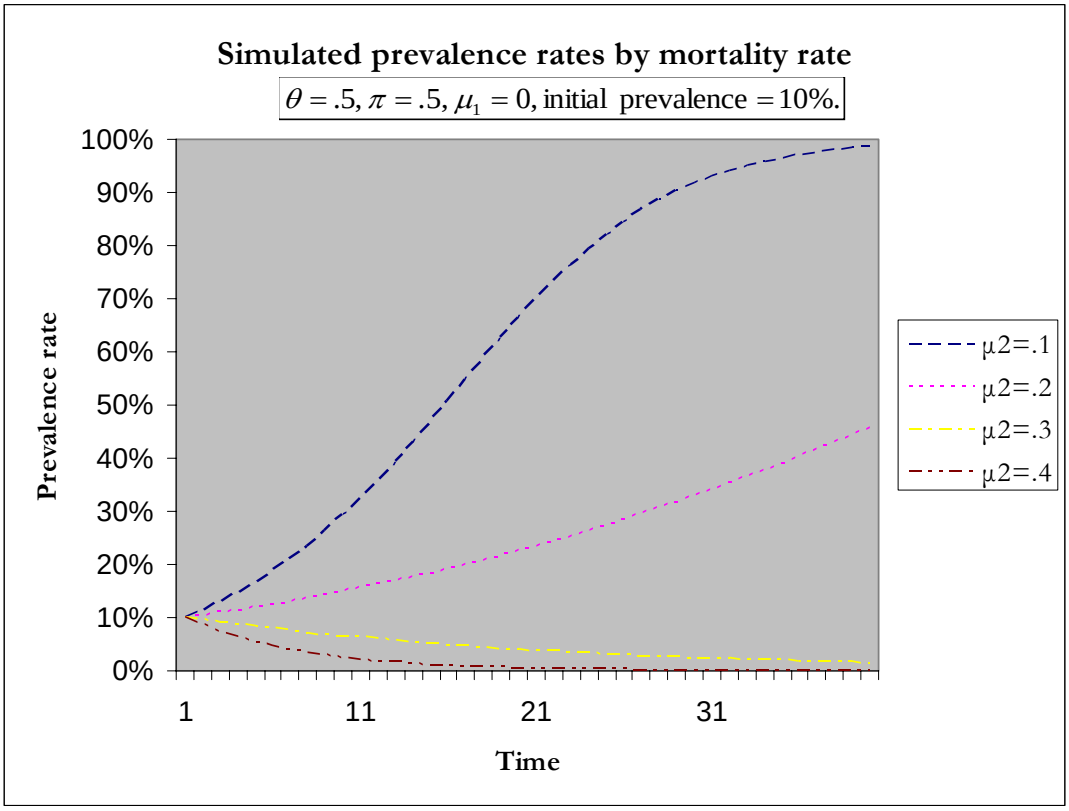

Figure 4. Phase plane diagram of predator-prey model.

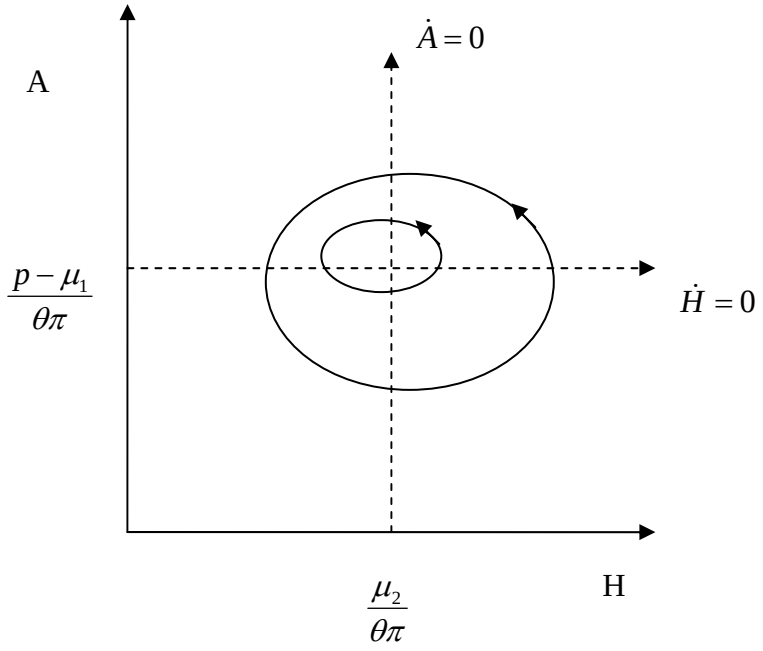

Figure 5. Median & smoothed seroprevalence rates from UNAIDS, by country/year.

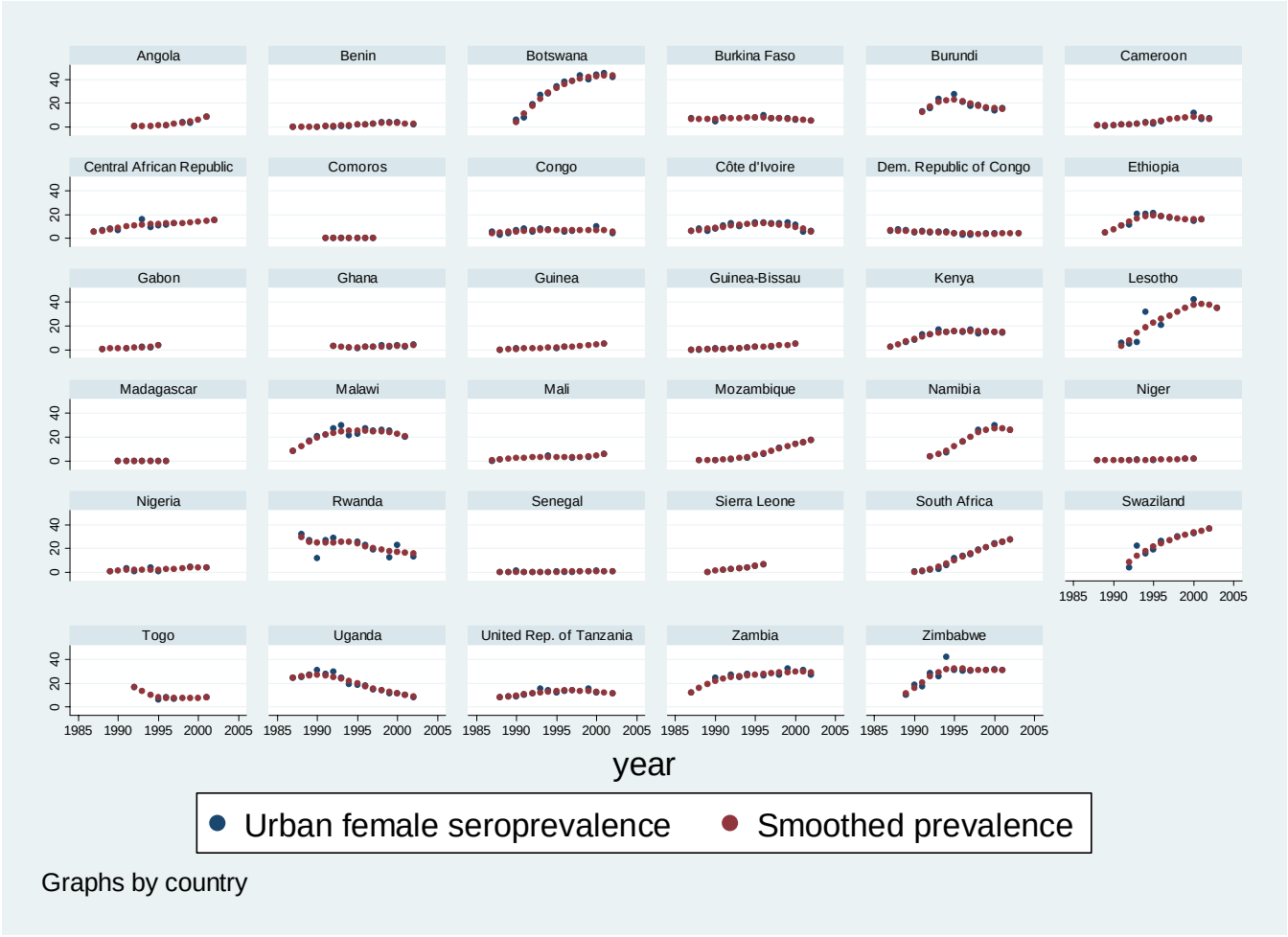

Figure 6.

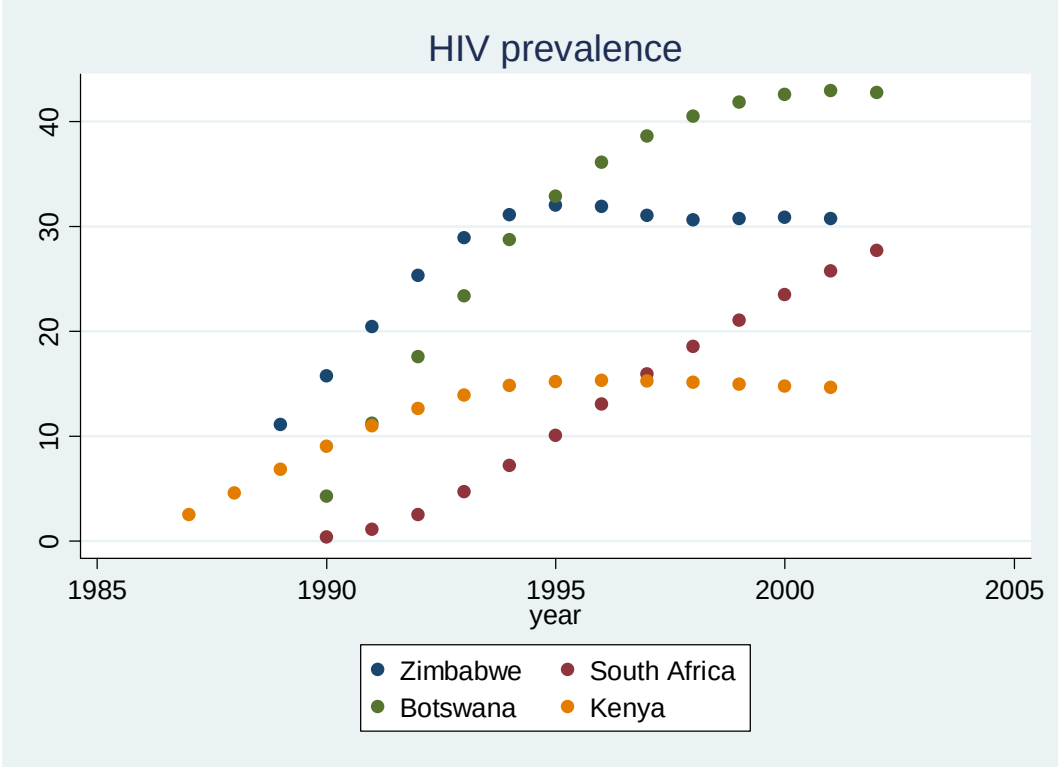
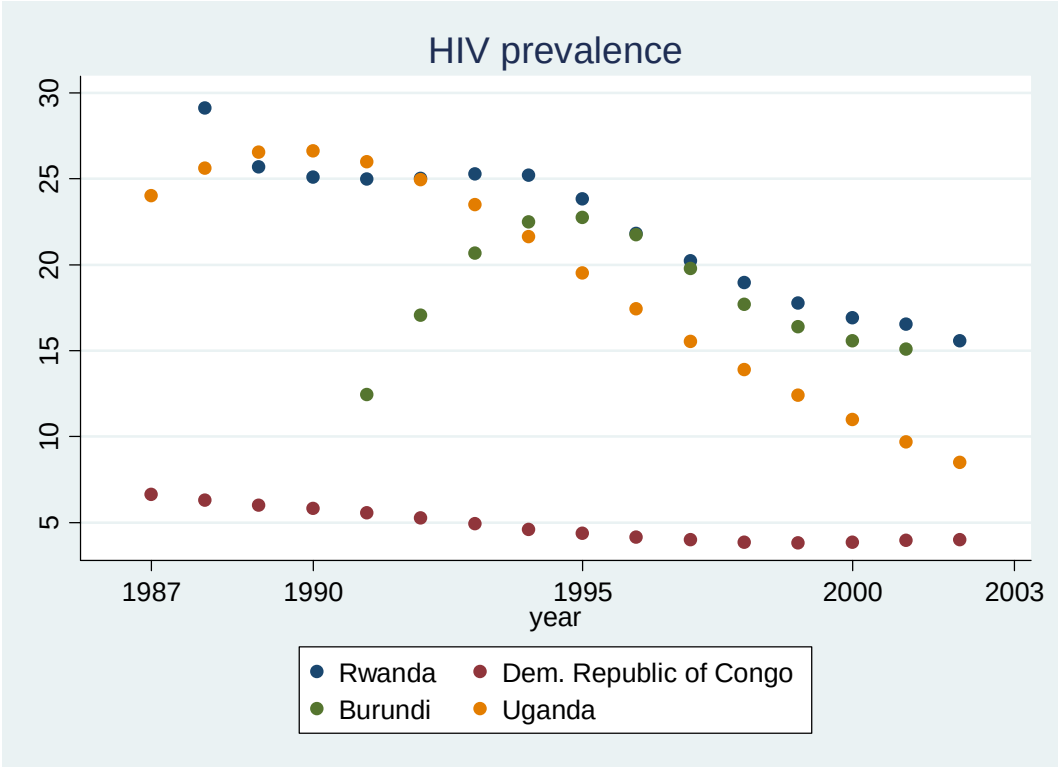
Figure 7.