In Sorrow to Bring Forth Children: Fertility amidst the Plague of HIV

Alwyn Young
London School of Economics
September 2007

Abstract

The HIV epidemic is lowering fertility in sub-Saharan Africa. This decline in fertility appears to reflect a fall in the demand for children, and not any adverse physiological consequences of the disease, as it is matched by changes in the expressed preference for children and the use of contraception, and is not significantly correlated with biological markers of sub-fecundity. A fall in fertility lowers dependency ratios and, for a given savings rate, increases future capital per person. These two effects more than offset the loss of prime working age adults and reduced human capital of orphaned children brought by the epidemic, allowing 27 of the nations of sub-Saharan Africa to cumulatively spend US$ 650 billion, or $5100 per dying adult AIDS victim, on patient care without harming the welfare of future generations. In sum, the behavioral response to the HIV epidemic creates the material resources to fight it.

In memory of Ho Hon and Nechama, terribly missed by us all. I am grateful to Ho Veng-Si, seminar participants at the Rand Corporation and the editors of this journal for many helpful comments.

I. Introduction

HIV is lowering fertility in sub-Saharan Africa. Table I below reports the average ratio of children aged 0-4 years old to women aged 15-49 enumerated in censuses in sub-Saharan countries with widespread HIV epidemics over the past three decades.[1] While the ratio of infants to fecund women rose slightly during the 1970s, it fell 3% in the 1980s and a full 16% in the 1990s.[2] As shown in the table, this decline in fertility is correlated with the development of the HIV epidemic. When the ln ratio of infants to fecund women in each country x census year is regressed on a full set of country dummies and the average countrywide HIV infection rate in the census and four preceeding years, one estimates a coefficient on HIV of -1.54.[3] The inclusion of country specific time trends reduces the magnitude of the point estimate, but does not change the conclusion. Fertility in sub-Saharan Africa is strongly negatively associated with the spread of HIV.

The association between fertility and HIV shown in Table I cannot be explained by the

---

[1] The table reports the data available to me for the 39 sub-Saharan countries identified by the U.S. Census Bureau as currently having generalized widespread epidemics (essentially, all of sub-Saharan Africa except Mauritania, Somalia, Sudan & the island nations). As sub-Saharan censuses are frequently sporadic, I use paired within country observations to identify changes decade by decade. Censuses are assigned to decadal years using the usual rounding rules. With the exception of South Africa where a systematic effort has been made to adjust for the census undercount, I use the population as enumerated (as adjustments elsewhere tend be fairly ad hoc). Where possible, I use the de facto population in preference to the de jure, but otherwise just try to maintain a consistent population definition within countries.

I should note that, in my experience, infants are systematically undercounted in third world censuses (one can see this by observing the growth of the absolute size of infant cohorts, suitably aged, from one census to the next). Consequently, the data in the table understate true fertility. However, unless the degree of the undercount is correlated with the spread of HIV, this will not bias the inferred relation between HIV and fertility. As will be shown further below, microdata surveys of fertility suggest similar effects.

[2] Throughout this paper, I use the term "percent" to refer to ln derivatives, as contrasted with changes in decimal fractions, which I indicate with the term "fraction."

[3] Without country dummies, the coefficient on HIV is actually -1.60 (.224), but it is obviously more sensible to allow for country fixed effects. The reader might worry that the epidemic has reduced the proportion of women in the prime child bearing ages of 20-39, thereby artificially lowering the fertility ratio described above. Given the slow progression of HIV in adults (discussed further below), this is not an issue. If I use the ln ratio of infants to women aged 20-39 as the dependent variable, I get coefficients of -1.57 (.176) and -1.14 (.243) for the specifications run in columns (1) and (2) of Table I, respectively.

Table I: Fertility & HIV

(A) Decadal Census Averages

	1970	1980	1980	1990	1990	2000
infants/fecund women	.765	.765	.771	.750	.736	.630
average HIV	.000	.000	.000	.020	.024	.117
# of countries	15	15	22	22	16	16

(B) Regression (ln infants/fecund women on average HIV)

	(1) With country dummies	(2) With country dummies and time trends
average HIV	-1.54 (.184)	-1.13 (.255)
N	90	90
R2	.798	.938

Notes: Infants/fecund women = the ratio of infants aged 0-4 to women aged 15-49 as recorded in each census year. Average HIV = the average infection rate for women aged 15-49 in the actual census year and four preceeding years, as estimated by the U.S. Census Bureau (see Section II below). # of countries = number of countries in each paired comparison. N = the number of countries in the sample, using all countries with more than one census observation. For the decadal averages censuses are assigned to decadal years using the usual rounding rules, but for the regression analysis (and the calculation of the average HIV infection rate) the actual census year is used instead. The regression analysis uses more data than is presented in the paired averages, as some census observations cannot be paired with an adjacent decadal observation but can still be used to calculate levels and trends. The data are presented in a spreadsheet, available from the author.

physiological effects of the disease. Medical studies indicate that African women infected with HIV, who are believed to be unaware of their infectious status, have perhaps a 1/4 reduction in their probability of becoming pregnant.[4] About 1/3 of children born to African women with

[4] Carpenter et al (1997) and Hunter et al (2003), in cohort studies in Uganda and Tanzania, respectively, find that HIV infected women have age-adjusted pregnancy odds ratios of about .74 and .71 of normal, respectively, while in a similar cohort study in Uganda, Sewankambo et al (1994) find an HIV-induced relative pregnancy risk of .8. Desgrées du Loû et al (1999) in a retrospective birth interval study of pregnant women arriving at antenatal clinics find that HIV infected women have .83 times the pregnancy risk of uninfected women, while Glynn et al (2000), in similar studies in Cameroon, Kenya and Zambia, find adjusted hazard ratios of

HIV become infected themselves, and of these about 1/3 will die of the ravages of the disease within one year of birth, and 3/5 within four.[5] Cumulatively, these physiological effects suggest a coefficient, in a regression of the ln ratio of infants to fecund women on the average infection rate, of about -.38.[6] Clearly, physiological effects alone cannot explain sub-Saharan trends.

In this paper I argue that the HIV epidemic has led to a decline in the desire, or demand, for children in sub-Saharan Africa. Using the microdata of 78 Demographic and Health Surveys (DHS) and World Fertility Surveys (WFS) in 29 sub-Saharan African countries, I show that the association between HIV and declining fertility suggested by aggregate census data are supported by household surveys. This decline in fertility is matched by a correlation between rising HIV infection rates and reductions in the expressed preference for children. The spread of HIV is associated with a rise in the use of condoms, but also with a very substantial increase in the use of non-viral-protective contraception. This suggests that the reduction in fertility is desired, and not merely the unintended consequence of efforts to limit the exchange of bodily fluids. Biologi-

---

between .74 and .84. The notable outlier is Gray et al (1998), who, in a cross-sectional analysis of a Ugandan community, find an HIV induced pregnancy odds ratio of .45. In a cohort study, Ross et al (1999) argue that reverse causation (i.e. pre-existing sub-fertility leading to marital problems, infidelity and HIV infection) accounts for half of the association between HIV and fertility, reducing their estimate of the HIV induced reduction in fertility from .47 to .25. One can easily conceive of other forms of endogeneity, e.g. marital problems leading to sub-fertility, infidelity and HIV infection. Although not recognized in this literature, it is also possible that the reduced fertility of infected women is a behavioral change stemming from knowledge of their own infectious status.

[5] See Bobat et al (1999), http:/www.avert.org/motherchild.htm (10/6/2005), Unicef (2002), and UNAIDS (2002).

[6] Assuming that prior to infection HIV positive women have average fertility, the reduction in the infant to fecund women ratio should be on the order of 1-h + h*p*(1-i + i*s), where h is the adult female HIV infection rate, p the post-infection relative fertility of HIV positive women, i the mother to child transmission rate, and s the survival rate of infants. I use values of 1/3 for i, .75 for p and .537 for s (arrived at by integrating the UNAIDS (2002) estimates of HIV related child mortality from ages 0-5, under the assumption that births are uniformly distributed across the five years preceeding the census). I regress the ln of this, calculated, ratio on the average HIV infection rate in my census data sample to arrive at the coefficient suggested above. Focusing not on covariance but on means, another way to summarize this result is to state that based upon the physiological effects of the disease, as summarized by the formula 1-h + h*p*(1-i + i*s), the movement from an average infection rate of .024 to .117 between 1990 and 2000 described in Table I can only explain 3.5% of the observed 16% decline in the fertility ratio.

cal markers that are associated with sub-fecundity, such as amenorrhea or miscarriages, are not significantly correlated with infection rates, providing further evidence that the change in fertility is behavioral and not biological. A host of independent and diverse measures indicate that the HIV epidemic has lowered the overall demand for children.

A decline in fertility implicitly creates resources that can be expended, today, in fighting the epidemic, without harming the welfare of future generations. For a given savings rate, a fall in fertility endows future cohorts with more capital per person. Furthermore, declines in fertility lower dependency ratios, so that at given per capita consumption levels savings rates actually rise, endowing future cohorts with more capital. Together, these effects suggest that is is possible to allocate resources to medical care without adversely affecting future living standards. This is only true, however, if the fertility effect dominates other mechanisms through which the HIV epidemic adversely affects current and future living standards. Two of these are participation and human capital accumulation. The loss of prime working age adults works to increase dependency ratios during the height of epidemic induced mortality, so that any attempt to maintain normal per capita consumption levels (let along pay additional medical costs) necessarily eats away at the physical capital bequethed to future generations. At the same time, the orphans left by the epidemic acquire less human capital, which lowers lifetime productivity; a loss which is perpetuated for countless generations as less educated persons have larger numbers of less educated children. In simulations further below, I find that the impact of fertility on dependency and population growth is overwhelmingly dominant, allowing the sub-Saharan countries to fund extensive medical care for the victims of HIV. I show that it is possible for 27 of the African countries afflicted by the epidemic to cumulatively expend an extraordinary 650 billion US$, or $5100 per patient in the year of each AIDS victim's death, in humanitarian patient care, while safeguarding the per capita living standards of current and future cohorts.

The paper proceeds as follows: In Section II below I discuss the DHS/WFS and HIV infection data used in the paper. Section III presents the analysis of the correlation between HIV

infection rates and fertility, preferences for children and contraceptive behavior, producing the results noted above. Before turning to simulations, Section IV examines the impact of the epidemic on the accumulation of human capital, showing that, outside of the adverse individual consequences of orphaning, the spread of HIV has not otherwise produced significant community wide effects. Section V then runs simulations to calculate the maximal possible expenditures on patient care consistent with safeguarding living standards. Section VI concludes.

II. Data and Methods

I will explore the response of fertility and other aspects of individual behavior to communal HIV infection using the microdata files of the Demographic and Health Survey (DHS) and its predecessor the World Fertility Survey (WFS). The DHS and WFS projects, supported by the U.S. Agency for International Development, have conducted irregular, but in-depth, household level surveys of fertility and health in developing countries. The survey at present consists of three questionnaires: (1) a survey of women aged 15-49, obtaining a detailed retrospective fertility history as well as information on current contraceptive use, fertility preferences, marital status, and personal characteristics (e.g. education and age); (2) a survey of adult men, collecting similar information, but lacking detailed histories of fatherhood; and (3) a household questionnaire, assembling basic demographic information on all members of the household. In the WFS and early rounds of the DHS, the men's and household surveys were frequently not administered. I have access to a total of 78 female, 57 male and 63 household surveys covering 29 of the 39 sub-Saharan countries with generalized HIV epidemics. Eight of the female surveys are from the WFS, with the remaining files coming from the DHS.[7]

In estimating the direct impact of communal HIV infection on individual behavior, it is important to control for spurious correlations generated by country specific characteristics, time trends, and indirect influences. The spread of HIV within sub-Saharan African countries is very likely to be correlated with country specific characteristics (e.g. incomes, location, and religious beliefs and practices), which may also have a major impact on individual behavior. As HIV infection rates have increased over time, they will easily be correlated with otherwise indepen-

[7] The appendix at the end of this paper lists the sample countries and available surveys. The data files can be acquired at www.measuredhs.com and opr.princeton.edu/archive/wfs/. There is some variability in the coverage of the surveys, with some surveys or questions restricted to ever-married or currently married women. I have taken care to ensure that this does not bias the estimates. Where most of the surveys restrict the relevant question to women of a particular marital status, I focus the entire analysis on that group alone. Where a small group of surveys restrict the question, I exclude these from specifications which do not include country specific controls for marital status.

dent trends in behavior. Finally, the impact of HIV might stem from its impact on other determinants of behavior (e.g. marital status, the survival of children and parents, and household wealth), rather than through any direct response to the presence of the epidemic and its associated mortality risk. In the analysis that follows I address these concerns using three sets of controls. First, as a baseline, I include a complete set of country specific constants, educational category dummies, and age and age squared terms.[8] In this case, the effect of HIV is identified through country specific trends in behavior, adjusted for the age and educational composition of the population. Second, I add country specific time trends, identifying the impact of HIV from non-linear deviations in trends within each country. Finally, I add country specific terms controlling for marital and familial status and, where available, the survival of parents and household wealth.[9]

The inclusion of country specific constants and time trends as control variables has the unfortunate consequence of reducing the sample variation used to estimate the impact of HIV. Table II below tabulates the sample countries by the number of surveys. As the reader can see, by the time country specific time trends are included, identification of the effect of HIV is reduced to 15 countries in the women's sample and 9 countries in the men's sample. I should

[8] Consequently, as none of the identification of the impact of HIV comes from cross-country variation, I do not include in my sample the DHS/WFS sub-Saharan countries (Comoros, Madagascar and Sudan) which do not have widespread epidemics.

[9] As fully detailed below, the control variables used in the analysis of fertility behavior and educational outcomes are slightly different, reflecting the availability of control variables. In particular, controlling for household wealth in the analysis of female behavior results in the loss of a large number of early country observations (although the basic results are not changed), so I do not include these in the analysis of female behavior.

One could argue that, aside from omitted variable bias (which I try to address), there is also the issue of endogeneity, as behavior affects the spread of HIV. I do not think this is a severe problem. First, by including country specific dummies, I focus on changes, eliminating the obvious cross-country correlation between levels of behavior and the levels reached by the epidemic. Second, changes in behavior affect the flow rate of new infections, not the stock of current infections. The fact that half of HIV infected adults survive 9 years of more (estimates of UNAIDS 2002) ensures a fairly long lag structure, which greatly ameliorates any endogenous correlation between changes. Finally, to the degree that endogeneity is a problem, it works against the conclusions of this paper, as the behaviors I find associated with HIV infection rates (e.g. increased use of condoms) would tend to lower future infection rates.

Table II: Distribution of Countries by Number of DHS/WFS Surveys			
	Women's Survey	Men's Survey	Household Survey
one survey	6	6	6
two surveys	8	11	10
> two surveys	15	9	11

note, however, that the paper's central results, concerning women's fertility, fertility preferences and contraceptive use, are *completely consistent* across various formulations (i.e. whether identifed across 23 countries without time trends, 15 with time trends, or 15 with time trends and additional individual level controls). Further, for one of the key variables, women's realized fertility, I am able to use the information in the DHS surveys to construct retrospective individual histories. In this case, identification of the effect of HIV comes from the entire 29 country sample. I find that the results survive the expansion of the effective sample. As regards the fertility preferences and contraceptive behavior of men, the influence of HIV mirrors its effect on women, but is more sensitive to the inclusion of time trends (when identification is reduced to a tiny handful countries).[10] Consequently, I relegate the analysis of male preferences and behavior to a supporting appendix, available from the author.

To conduct the analysis, the irregular dates of the DHS have to be matched to contemporaneous data on HIV infection rates by country. I make use of the U.S. Census Bureau's estimates of annual infection rates amongst pregnant women in sub-Saharan African countries. The Census Bureau uses the 2003 UNAIDS/WHO Estimation and Projections Package (EPP) to fit an epidemic progression curve to observations on the infection rates of pregnant women attending antenatal clinics in the subject countries. The EPP model is basically one of contagion from

[10] For example, only two countries have more than two male surveys providing information on contraceptive use.

the existing number of infected individuals to an estimated "at risk" group, with a small ad-hoc initial impulse function used to start the epidemic.[11] As the EPP model fits a fairly flexible curve to a relatively short time series, the reader's main concern should probably not be the modelling,[12] but the representativeness of the underlying data.

Estimates of HIV infection rates are naturally bedeviled by the problem of sample selection, i.e. who consents to being tested. In this regard, pregnant women form a natural sample, as they regularly submit blood for tests. However, the association between HIV, fertility and sexual behavior might produce serious biases. On the one hand, if HIV physiologically or behaviorally lowers individual fertility, the prevalence of HIV in pregnant women will understate infection rates in the population at large. On the other hand, both pregnancy and HIV infection are positively associated with risky sexual behavior, which would make antenatal clinic data overstate community infection rates. Perhaps most fundamentally, the correlation between risky sexual behavior, fertility and HIV could easily change in the course of the epidemic, producing antenatal clinic data quite out of keeping with the "true" community time series.

To evaluate the representativeness of samples drawn from pregnant women, I have developed a dataset matching antenatal clinic data to the infection rates of women of child-bearing age in the surrounding community for 50 regions in 8 African countries.[13] Using this data, I run a

---

[11] The estimated parameters are the epidemic start date, the "at risk" population, the contagion rate, and a behavioral change adjustment to the "at risk" group based upon cumulative cohort AIDS related mortality. I thank Ms. Laura Heaton of the Health Studies Branch, Population Division, U.S. Bureau of the Census, for graciously providing her estimates and discussing their construction. The EPP model is explained more fully in UNAIDS 2003.

[12] I tried fitting simple polynomial functions of time (without the formal EPP model of contagion) to some of the underlying Census Bureau data, and produced very similar time paths.

[13] The data are presented in a spreadsheet, available from the author. I draw on antenatal/community comparisons for 12 regions in Cameroon, Kenya, Malawi, Tanzania, Uganda and Zambia as reported in the medical studies of Changalucha et al (2002), Crampin et al (2003), Fylkesnes et al (1998, 2001), Glynn et al (2001), Gregson et al (2002), Kigadye et al (1993), and Killian et al (1999). I also match the DHS microdata on HIV infection rates for Ghana 2003 and Zambia 2001/02 to regional urban and rural antenatal surveillance surveys for the same years (as reported in the Census Bureau's HIV Surveillance Data Base) to produce another 37 regions. Finally, I contrast the countrywide data of the 2002 South African Nelson-Mandela Survey (HSRC 2002) with the 2002 South African maternity clinic seroprevalence results (RSA DOH

Table III: Community vs. Antenatal Clinic Infection Rates (logit analysis)		
	(1)	(2)
Sample (community = 1)	.051 (.025)	.064 (.054)
Sample interacted with regional dummies		.009 (.033)
N	86684	86684
Pseudo R2	.099	.099

Notes: All equations include a full set of regional dummies (D_i). N = number of observed women. With the interaction term, the ln odds ratio = $D_i + D_s * S + D_{sxd} * S * D_i$, where S = 1 if the sample is from the community, and D_s and D_{sxd} are the estimated coefficients reported above.

logit of the probability a women is infected on a complete set of regional dummies and her sample (community vs. antenatal). As shown in Table III, column (1), I do not find a particularly significant difference between community and antenatal infection rates. Furthermore, when I interact the community dummy with the level of regional antenatal infection, I find no significant effect at all (column 2). Whatever the biases of antenatal clinic data, they appear to cancel out over the range of observed infection rates. While antenatal clinic data are potentially problematic, they are the most widespread available and, in this paper, I follow the Census Bureau in assuming that they reflect community patterns of infection.

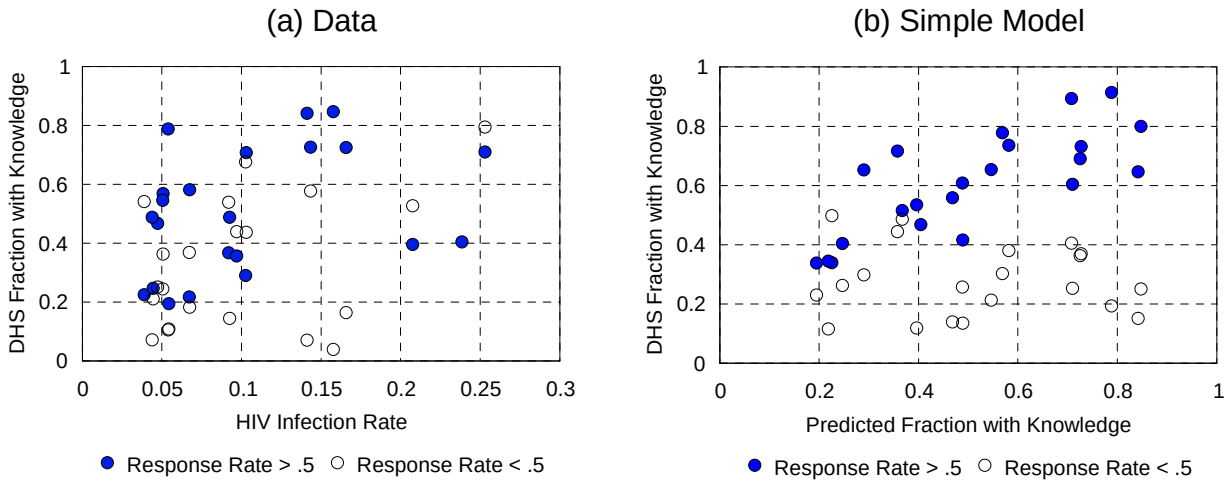
2003). The Nelson-Mandela Survey (of males and females and all age groups) provides extensive information on infection rates, but is somewhat sparse in its description of sample sizes. I estimate the number of black and non-black females aged 15-49 sampled by using data on age group infection rates and the assumption that the ratio of black to non-black females in this group equals that amongst all adults aged 15 and above. Adjustment for race is important, as blacks make up .60 of the HSRC survey, but .85 of the DOH survey, a fact which I incorporate into the likelihood equations by calculating separate infection rates for blacks & non-blacks. Without a great number of assumptions, it is not possible to calculate the sample sizes by province, so I do not compare the HSRC and DOH data at that level.

Finally, I address the issue of whether the community infection rates known to the econometrician are actually known to the sample. The development of HIV in adults is gradual, with individuals on average remaining largely asymptomatic for almost a decade. Pediatric HIV, however, develops at a ferocious pace, killing 1/3 of perinatally infected children within one year of birth. As about 1/3 of children born of HIV positive mothers in African countries are infected, women in communities with generalized epidemics will rapidly become aware of the existence and prevalence of the disease.[14] As it so happens, 43 of the DHS surveys at my disposal have asked the respondents whether they know someone who has AIDS or has died of AIDS. In Figure I(a) below I graph the average response for these surveys against the Census Bureau's estimate of the average community infection rate in the survey year. As the reader can see, even at low infection rates respondents already have personal contact with the epidemic. As a logical consistency check, I ask how many women a survey individual would need to know to maximize the probability of the data being observed under the assumption that all experiences with AIDS come from observations of the deaths of infants of female acquaintances.[15] The value that maxi-

[14] On disease progression in adults see http://www.avert.org/hivstages.htm (10/6/2005); on progression in children see Blakeslee (2001) and Bobat et al (1999); and on rates of mother to child transmission see Bobat et al (1999), http://www.avert.org/motherchild.htm (10/6/2005), and Unicef (2002). Bobat et al find that symptoms of diarrhoea, severe thrush, marasmus, and pneumonia were present in 3/4 of South African infants dying of AIDS, while more than 1/2 showed significant lymphadenopathy and neurological abnormalities. While some of these are commonly associated with infant mortality, their extraordinary simultaneous occurrence in infants would quickly alert mothers to the presence of a new source of mortality.

[15] With 1/3 mother to child transmission, the probability a respondent gets an observation of an AIDS related infant death from an acquaintance a years ago is p*(ha/3)*(1-Sa), where p is the probability a female acquaintance of the respondent gives birth, ha the female HIV infection rate a years ago, and Sa the probability an HIV infected child survives a years. Thus, the probability a respondent does not get an observation of an AIDS related death from n iid acquaintances is 1 - p*(ha/3)*(1-Sa) raised to the nth power. For my value of p I use the average number of births per woman in the preceeding 12 months in the DHS sample for that survey, while Sa is taken from the UNAIDS (2002) recommended children's mortality profile. The 1999 DHS female survey for Nigeria actually includes information for individuals as young as 10 and, in the cross section, shows a rapid rise in the fraction of women who know of an AIDS death from low levels at age 10. Consequently, I assume that observation of community fertility begins at age 10 and take the probability of not having observed an AIDS death as the product of all the annual probabilities of no observation from age 10 to the respondent's current age. The assumption of no intertemporal correlation in observations from individual acquaintances is obviously wrong, but, given the relatively short history of the epidemic (the average infection rate 10 years before

Figure I: Knowledge of Someone with AIDS



mizes this likelihood is 28, which is not unreasonable, and, as shown in Figure I(b), the predicted values of the model match the average survey responses fairly well.[16] Arguably, perinatal transmission ensures that women in communities with generalized epidemics know the infection rate as well, if not better, than the researcher.

each observation being .04), not likely to impart a great bias.

[16] A regression of the survey average response on the predicted averages yields an R2 of .44. If one restricts the sample to DHS surveys where more than .5 of women responded to the question, the fit improves dramatically and the R2 rises to .6 (see Figure I(b)). A regression of the ln fraction of the sample responding on the year of each survey yields a coefficient of -.070 (s.e.=.022), or -.080 (.025) if one includes country dummies. Women are increasingly not responding to this question.

I should note that the predicted fraction in Figure I(b) is not monotonic in the current infection rate. A country with a low current infection rate, but a long epidemic history, could have a high predicted fraction (as women have had the opportunity to observe more children for longer), while a country with a high current infection rate, but a short epidemic history, could have a low predicted fraction.

III. Fertility

In this section I show that communal HIV infection, as measured by the infection rate of pregnant women, lowers average fertility. This reduction in fertility appears to be a conscious decision, as it is consistent with changes in the desired number of children and the use of contraception. It does not appear to be a physiological consequence of the disease, as there is no significant association between HIV infection rates and amenorrhea and miscarriages. The decline in fertility seems to reflect a powerful community-wide response, as it goes well beyond measured changes in the relative fertility of infected women. As noted in the previous section, I will examine the impact of HIV on each dependent variable using, in sequence, three sets of controls: (1) a complete set of country specific constants, educational category dummies and age and age squared terms; (2) the preceeding, plus country specific time trends; (3) all of the above, plus country specific controls for marital status, urban/rural location, and the number and square of born and living children.[17] Although my focus is on the influence of HIV, I report the average country specific impact of educational attainment in each equation as a means of gauging the magnitude of the HIV coefficient and comparing its influence across tables with disimilar dependent variables. The construction of the variables and the samples appearing in each equation are described in the appendix at the end of the paper.

As shown in Table IV, the average HIV infection rate in the year of each DHS/WFS survey is negatively associated with the probability a respondent gave birth in the preceeding 12

[17] To clarify, all control variables are entered at the individual level, reflecting individual respondent characteristics, but I estimate separate coefficients for each control variable (age, marital status, etc.) in each country.

months.[18] With the baseline controls, column (1), I find that, on the margin, each .01 increase in the fraction of the population infected with HIV lowers the ln probability that a woman gives birth in a given year by -1.6 percent.[19] Adding country specific time trends and other controls, in columns (2) and (3) of the table, raises the point estimate and standard error of the HIV variable, as the effective sample variation is reduced, but otherwise does not change the picture. The magnitude of HIV's impact on fertility can be gauged by comparing it with the influence of education. At infection rates of .1, present since the year 2000 in 10 sub-Saharan countries,[20] HIV reduces fertility by a proportional amount equivalent to the estimated impact of raising the entire population's educational attainment from none to secondary, as estimated including other important determinants of fertility such as urban/rural location (column 3 in the table). As indicated by the table's estimates of current[21] fertility rates with and without the epidemic, with mean country infection rates of about .081 in my sample, HIV appears to be associated with about a .02 to .03 reduction in the fraction of women giving birth in a recent year.

---

[18] For the logit model presented in the table, the probability a respondent gave birth is given by P = exp(xB)/(1+exp(xB)). In the binary and multiple choice models I use below, the interpretation of results is complicated by the fact that there is no simple relation between coefficient estimates and the response of the dependent variable. To ease discussion of magnitudes, in reporting results I calculate the average derivative of the ln expected value (in the case of Table IV, the probability of giving birth) with respect to the variables reported in the table (using the delta method to calculate standard errors). In calculating the average ln derivative, which depends upon sample values, I weight all countries equally, i.e. take the average across the countries of the average within country sample derivative. The "derivatives" with respect to the educational dummies represent the average change in the ln probability associated with a movement from no education to each of the higher educational categories.

[19] As noted earlier, I use the term "percent" to refer to ln derivatives, as contrasted with changes in decimal fractions, which I indicate with the term "fraction."

[20] That is, Botswana, Central African Republic, Lesotho, Malawi, Mozambique, Namibia, South Africa, Swaziland, Zambia and Zimbabwe.

[21] The "current" mean represents the average, across the sample countries, of the predicted sample mean in the most recent survey available for each country. As DHS surveys occur irregularly, the actual date varies by country (see Table A1 in the appendix). For 22 of my 29 countries, it is between 2000 and 2005, but, regrettably, for a small handfull it is in the 1990s or (for two countries) the late 1980s. In looking through the tables, the reader may notice that the predicted mean with HIV varies between specifications and does not exactly match the current sample mean. This stems from the fact that the coefficients are estimated using a broader sample than that used to summarize current patterns of behavior.

Table IV: Realized Fertility

	Past Year (logit model)			Retrospective Annual (logit model)		
	(1)	(2)	(3)	(4)	(5)	(6)
Primary Educ.	-.159	-.153	-.011	-.079	-.073	.021
	(.013)	(.013)	(.012)	(.004)	(.004)	(.004)
Secondary Educ.	-.611	-.605	-.204	-.467	-.459	-.124
	(.016)	(.016)	(.016)	(.005)	(.005)	(.005)
Tertiary Educ.	-.945	-.934	-.414	-.971	-.963	-.391
	(.052)	(.052)	(.055)	(.017)	(.017)	(.014)
HIV	-1.58	-2.06	-2.07	-1.30	-1.06	-1.10
	(.127)	(.355)	(.345)	(.030)	(.046)	(.040)
"Current" Mean	.183	.183	.183	.167	.167	.167
w/out HIV	.206	.211	.210	.185	.180	.180
w/ HIV	.184	.181	.182	.170	.168	.168
Sample Mean	.191	.191	.192	.159	.159	.159
Pseudo R2	.074	.075	.132	.134	.135	.186
N	563474	563474	567031	10427761	10427761	10502123

Notes: Column (1) contains a complete set of country specific constants, educational category dummies (none, primary, secondary & tertiary), age and age squared terms. Column (2) adds country specific time trends. Column (3) adds country specific controls for marital status (never, currently or formerly married), urban/rural location, and the number and square of born and living children. The reported "coefficients" represent the average country ln derivative (or, for education, ln change) of the predicted probability with respect to the variable in question, with standard errors calculated using the delta method. All standard errors are adjusted for clustering. The "current" mean refers to the average country sample mean in the most recent survey year, while the same with and without HIV refers to the predicted average country values in that year with and without the existing levels of HIV. These notes apply to all subsequent tables in this section, with exceptions noted as necessary. For the retrospective annual projections in columns (4)-(6), the number of born and living children refers to the cumulative number prior to the year in question, while the marital status dummy takes a value of one from the year of the first union on (as, for individuals who are no longer married, I do not have information on the year in which the separation occurred). The samples expand in columns (3) and (6) as, with the addition of marital status dummies, I am able to add the 1977 WFS survey of Lesotho, which only covered ever married women.

One can use the birth history reports in the DHS to construct a retrospective record of each woman's fertility in each calendar year of her life since age 10.[22] This creates a time series for each survey so that, despite the country specific constants and time trends, each of 29 countries contributes to the identification of the impact of the HIV variable.[23] As can be seen from a comparison of columns (4)-(6) and (1)-(3) of Table IV, without a time trend the impact of the HIV variable on the retrospective histories is close to its influence on fertility in the past year, but with a time trend the effect is dramatically weaker. The issue here is not the sample, but the calculation of the time trend. If I limit the estimation of column (5) to the fifteen countries with three or more surveys which produce the estimated derivative of the ln probability with respect to HIV of -2.06 in column (2), I get an estimated derivative of -1.17 in column (5). With the retrospective histories the estimated time trends change, and this changes one's inference about the impact of HIV. The results are quite robust to an expansion of the effective sample, but are

[22] While some women report births back to age 6, I take age 10 as the earliest possible onset of menarche, ignoring the earlier records. The reader will notice that the sample and current means are lower in the right-hand panel of Table IV. The right-hand panel of the table includes fertility between the ages of 10 to 14 (when pregnancies are extremely rare), while the left-hand panel is limited to women currently aged 15 or above. In addition, the DHS estimate of "past year" fertility, used in the left-hand panel, is fertility in the survey month and the 12 months preceeding the survey. This produces fertility rates which are somewhat higher than the calendar year fertility rates I develop for the right-hand panel of the table.

[23] As noted in the preceeding section, when there is only one year of observation of the dependent variable per survey, the country specific constants limit the identification of the impact of HIV to the 23 countries with 2 or more surveys. With time trends, identification is limited to the 15 countries with 3 or more surveys. With a time series on the dependent variable calculated from each survey, this problem disappears. I should note that, throughout, I keep all available countries in the sample, even when they do not contribute to the identification of the impact of HIV, so as to make the comparison of the impact of other variables (e.g. education) consistent across the tables.

sensitive to alternative calculations of the time trend.[24] Based upon the right-hand panel of Table IV, HIV can be associated with only about a .01 reduction in the fraction of women giving birth in a recent year. However, at .1 infection rates the impact of HIV is still close to that of secondary education after controlling for other determinants of fertility, as the estimated impact of education is lower in the right-hand panel of the table as well.

[24] The sensitivity with respect to the time trend can, however, make the analysis sensitive to the sample. In my analysis of fertility in South Africa (Young 2005), in a formulation similar to that of column (5) above, I estimated the ln derivative of expected births with respect to HIV to be -1.63. In that paper I was trying to estimate the response of household behavior to lifetime differences in income and consequently restricted the sample to women aged 25 and over at the time of the survey, whose education I could take as completed. I now realize that if I had used the entire South African sample of women aged 15 and above, the estimated ln derivative with respect to HIV would have been only -.86. This comes from the fact that with a younger sample more of the retrospective observations are recent, so that a steeper time trend takes away the impact of the recent rise in HIV. In Table IV above, if I similarly restrict the sample to women aged 25 and above at the time of the survey, the ln derivative with respect to HIV in column (5) increases to -1.50 (this sensitivity does not appear in column (4), which does not include time trends). As I am now aware of this problem, I estimate all of the effects using the full sample, as this produces the lowest estimates.

Table V: Desired Fertility

	Ideal Number of Children (poisson count model)			Want No More Children (logit model)		
	(1)	(2)	(3)	(4)	(5)	(6)
Primary Educ.	-.173 (.003)	-.167 (.003)	-.121 (.003)	.372 (.014)	.350 (.014)	.317 (.015)
Secondary Educ.	-.366 (.004)	-.355 (.004)	-.231 (.004)	.479 (.016)	.448 (.017)	.570 (.020)
Tertiary Educ.	-.560 (.008)	-.542 (.008)	-.341 (.007)	.343 (.043)	.298 (.044)	.788 (.045)
HIV	-1.65 (.063)	-1.56 (.127)	-1.61 (.113)	4.86 (.160)	4.30 (.392)	5.34 (.434)
"Current" Mean	5.15	5.15	5.15	.308	.308	.308
w/out HIV	5.91	5.76	5.77	.217	.239	.234
w/HIV	5.23	5.13	5.13	.299	.313	.312
Sample Mean	5.28	5.28	5.32	.276	.276	.276
Pseudo R2	.081	.084	.098	.272	.279	.360
N	482260	482260	492038	353535	353535	353521

Notes: The analysis of the desire for more children is restricted to currently married individuals (and consequently excludes marital status dummies) as in earlier rounds of the DHS/WFS this question was asked of this group alone. The "coefficient" estimates for the desire for more children (as well as other binary choice variables presented later in this paper) represent the average derivative of the ln probability, as explained earlier above.

The reduction in fertility described above reflects changes in desired fertility. As shown in the first three columns of Table V, each .01 increase in the HIV infection rate reduces the ideal number of children reported by DHS/WFS respondents by about -1.6 percent, an amount similar to the estimated impact on realized fertility in the past year presented above.[25] It also has a powerful influence on the probability a respondent definitely indicates that they want no more children,[26] raising the ln probability by 4 to 5 percent (columns 4 to 6 of the table). Both variables indicate that, as in the case of realized fertility, an infection rate of .1 is roughly equivalent to a movement of the entire population from no education to secondary attainment, as estimated using all controls (columns 3 & 6). At average current population infection rates, HIV is found to lower the expressed ideal number of children by about .6, while raising the fraction of respondents indicating that they want no more children by about .08.

Communal HIV infection leads to purposeful fertility control. As shown in columns (1)-(3) of Table VI, the average national HIV infection rate increases the probability that a

[25] As the ideal number of children takes zero values for some respondents, the left half of Table V uses a Poisson count model rather than linear regression. In the poisson model the expected response is given by exp(xB), so that the coefficient estimate represents the derivative of the ln mean predicted response with respect to the dependent variable. Thus, the reported coefficient estimates for HIV and education (the mean of the country specific dummies) are comparable to the mean ln derivatives of the predicted probability reported in the tables with logit models.

[26] I code "no more" and "sterilized" as indicating a definite desire for no more children (dependent variable = 1), using the responses "have another" and "undecided" as the opposite (dependent variable = 0), and excluding from the sample ambiguous responses such as "declared infecund" and "never had sex." As noted in the table, the sample is restricted to currently married women, as this question was only asked of this group in earlier rounds of the DHS/WFS. If I include unmarried women and run the specification of column (6) with marital status controls, I get an average estimated derivative of the ln probability with respect to HIV of 6.02.

Table VI: Use of Contraception (married women only)						
	General (logit model)			By type (multinomial logit model)		
	(1)	(2)	(3)	(4)	(5)	(6)
				Condoms		
Primary Educ.	.746 (.019)	.727 (.019)	.622 (.020)	1.44 (.092)	1.40 (.092)	1.23 (.092)
Secondary Educ.	1.35 (.019)	1.33 (.019)	1.21 (.021)	2.46 (.092)	2.41 (.092)	1.99 (.094)
Tertiary Educ.	1.57 (.033)	1.55 (.033)	1.53 (.035)	3.08 (.123)	3.03 (.124)	2.43 (.130)
HIV	4.28 (.180)	2.34 (.456)	2.78 (.461)	10.60 (.833)	4.74 (1.51)	5.06 (1.50)
"Current" Mean w/out HIV w/HIV	.280 .189 .267	.280 .234 .281	.280 .229 .281	.031 .012 .026	.031 .021 .030	.031 .021 .030
Sample Mean	.223	.223	.223	.017	.017	.017
				Other Contraceptives		
Primary Educ.				.754 (.024)	.731 (.024)	.638 (.024)
Secondary Educ.				1.28 (.026)	1.26 (.027)	1.17 (.028)
Tertiary Educ.				1.47 (.045)	1.44 (.047)	1.48 (.051)
HIV				4.07 (.174)	2.25 (.449)	2.70 (.454)
"Current" Mean w/out HIV w/HIV				.262 .176 .249	.262 .221 .265	.262 .215 .265
Sample Mean				.204	.204	.204
Pseudo R2 N	.149 315013	.159 315013	.191 315000	.142 285718	.152 285718	.182 285705

DHS/WFS female respondent is using any form of contraception.[27] One might argue that this increased use of contraception reflects the need of the respondents to protect themselves from sexually transmitted diseases, rather than their desire to control fertility per se. Alternatively, the increased supply of condoms in countries with rising epidemics, motivated by the endogenous response of governments and aid agencies, could be behind the overall rise in contraceptive use. Some insight into these issues is provided by columns (4)-(6) of Table VI, where I model individual behavior as a choice amongst three alternatives: no contraception, the use of protective condoms, or the use of some other form of contraception.[28] HIV evidently leads to a comparatively large rise in the ln probability of using condoms, although the change in the absolute fraction of current use is quite small (about .01), given the low usage levels prevailing in sub-Saharan Africa. Perhaps more interesting, however, is the fact that the probability of using other forms of contraception rises with the HIV infection rate as well, with my estimates suggesting an increase of between .04 and .07 in the current usage associated with the epidemic.[29] Whether for reasons of demand or supply, individuals in communities with rising rates of HIV

[27] The sample excludes women who are currently pregnant, as well as those who declare themselves to be infecund. As the question was only put to currently married women in early rounds of the DHS and WFS, I restrict the analysis to that group alone. If I include currently unmarried women and run the specification of column (3), with marital controls added, I get a derivative of the ln probability of using contraceptives with respect to HIV of 3.21 (s.e.=.415). As for column (6), with the full sample and marital controls I find an average derivative of the ln probabilities of condom use and other contraceptives with respect to HIV of 7.91 (1.17) and 2.90 (.423), respectively.

[28] In the multinomial logit model of this part of the table, the probability action i, amongst a set of potential actions, is taken is given by $\exp(xB_i)/\sum\exp(xB_i)$. The B_i for a base category, in this case no contraception, are normalized to equal 0. As in other tables, the reported "coefficients" represent the average country derivative of the ln probability of each action with respect to the variable in question.

[29] However, the reader will note that the impact of HIV relative to education is definitely lower in this table, with my estimates indicating that a .1 infection rate would produce a response equal to about 1/4 of that associated with a population wide movement from no education to secondary attainment (as estimated with all controls). This stems mainly from the fact that the educational elasticities of contraceptive behavior are much larger than those estimated for other variables in earlier tables and, moreover, are quite insensitive to the inclusion of additional controls. The relationship between contraceptive use and realized and desired fertility is obviously not ln linear.

infection are increasing their use of condoms. They are also, however, increasing their use of non-protective, fertility inhibiting, contraception.[30]

The impact of HIV on fertility reported above might reflect the biological effects of the disease, rather than behavioral changes. As noted in the introduction, a number of medical studies have found reduced fertility among HIV positive women who are otherwise believed to be unaware of their seroprevalence status. If HIV biologically reduces the fecundity of infected women, then, obviously, average fertility will be negatively correlated with average levels of HIV infection. While struggling to find a compelling biological explanation, the medical literature conjectures that the fecundity inhibiting effect of HIV may be a consequence of amenorrhea, known to occur at the late stages of the disease, and miscarriages.[31] I investigate these mechanisms in Table VII. In the left hand panel, I project the probability that a DHS respondent reports she has menstruated in the previous six weeks on the average HIV level and the usual

[30] About .006 of my sample (.028 of those using a contraceptive method other than condoms) reports using viral protective total abstinence (as opposed to periodic, rhythm, abstinence) as a contraceptive method. However, this response does not appear credible, as two-thirds of these individuals also report either having given birth or having had sex in the previous 12 months. I should also note that it is not likely that the increased use of non-protective contraception reflects major misperceptions about methods to avoid transmission of the virus. In 47 of my DHS surveys the respondents were asked whether they think a person can avoid AIDS by using condoms during sex. Affirmative responses are given by .35 of non-contraceptive users, .79 of condom users, and .58 of other contraceptive users. A separate AIDS and STDs module available for 16 surveys asks respondents whether they have ever heard of using condoms to avoid AIDS. The response to this "heard of using" question is affirmative .53 of the time that the "think can avoid" question is negative (but the converse is only true .05 of the time). Affirmative responses to this question are given by .53 of non-contraceptive users, .96 of condom users, and .82 of other contraceptive users. When asked whether they think people can avoid AIDS by abstaining from sex, .33, .35 and .34 of the three groups respond in the affirmative. Otherwise, the DHS respondents are not asked whether they believe non-barrier contraception will prevent transmission of the virus.

[31] Krieger et al (1991) find abnormal sperm in men with AIDS (who do not survive long in Africa), but no abnormalities whatsoever in minimally asymptomatic men with HIV. Widi-Wirski et al (1988), in a study of AIDS defining symptoms, find a strong association between amenorrhea and HIV in women, but .7 of such women in their sample had already lost more than .1 of their body weight. D'Ubaldo et al (1998) and Miotti et al (1990) find a significant correlation between HIV and retrospective histories of miscarriage. Gray et al (1998), who find the strongest negative influence of HIV infection on individual fertility, also engage in the most honest head scratching about its cause.

sequence of controls. I include the controls for education, urban/rural location, etc., not because these influence the biological likelihood of amenorrhea, but because they proxy for the probability that a woman, unknowingly, is actually pregnant.[32] Thus, as shown in the table, educational attainment is positively associated with menstruation. With regards to the impact of HIV, with minimal controls I do find a significant negative association between the average infection rate and the probability of a recent menstrual period (col. 1). However, once I add time trends and further individual controls, the significance of the relationship vanishes, although the coefficient does remain negative. The right hand panel of Table VII investigates the relationship between HIV and the probability that a pregnancy in a given year successfully resulted in a live birth, as opposed to ending in a stillbirth, miscarriage or abortion.[33] I find no consistently significant relation between HIV and the probability of successful pregnancies. These results suggest that the presence of sub-fertility in otherwise asymptomatic HIV infected women found in African studies may be a consequence of reverse causation[34] or a behavioral response to knowledge of their infection.

[32] I exclude from the sample all women who are currently breastfeeding and all women who are sure they are pregnant.

[33] I construct the data using the DHS "calendar" information on events in the five years prior to the survey. This is available for seventeen surveys in thirteen countries (the sample is listed in the appendix). I treat each termination or live birth as an event, which could have taken the opposite outcome, and classify it according to the year in which it occurred. Multiple events can occur for a single respondent in a single year. The most significant determinant of this variable is the number of born and living children (both positively associated with successful pregnancies), which, in the last column, render the educational dummies insignificant.

[34] While Ross et al (1999) focus on the possibility of sub-fecundity leading to marital problems and infection, there is also the possibility that marital problems lead to sub-fertility and infection. In this regard, Selwyn et al's (1989) study of New York female intravenous drug users in a methadone programme is interesting, as the primary cause of infection within the sample moves from marital and sexual issues to bad draws in needle sharing. Within the universe of intravenous drug users, they find no difference in pregnancy or miscarriage rates between asymptomatic HIV positive and HIV negative women.

Table VII: Biological Effects of HIV

	Menstruated in Past Six Weeks (logit model)			Successful Pregnancy (logit model)		
	(1)	(2)	(3)	(4)	(5)	(6)
Primary Educ.	.032 (.003)	.032 (.003)	.019 (.003)	-.016 (.002)	-.016 (.002)	-.008 (.002)
Secondary Educ.	.113 (.003)	.113 (.003)	.088 (.003)	-.053 (.004)	-.052 (.004)	.001 (.003)
Tertiary Educ.	.103 (.008)	.104 (.008)	.063 (.011)	-.120 (.018)	-.120 (.019)	.016 (.007)
HIV	-.285 (.036)	-.140 (.090)	-.105 (.090)	-.268 (.049)	-.086 (.108)	-.014 (.110)
"Current" Mean w/out HIV w/ HIV	.844 .859 .842	.844 .851 .842	.844 .849 .842	.900 .942 .926	.900 .916 .910	.900 .910 .909
Sample Mean Pseudo R2 N	.833 .137 289556	.833 .138 289556	.833 .150 289552	.931 .024 164261	.931 .028 164261	.931 .119 164261

Notes: The analysis of the probability of a menstrual period excludes all women who are currently breastfeeding or indicate that they are pregnant. Successful pregnancy is the probability that a pregnancy resulted in a live birth, as opposed to terminated in a stillbirth, miscarriage or abortion.

I close this section with further evidence that the association between the HIV epidemic and declining fertility reported above reflects broad communal responses, rather than the physiological or behavioral response of infected women alone. In some recent DHS surveys the respondents have been tested for HIV. I am able to link these seroprevalence results to individual fertility records in 11 surveys, allowing an analysis of trends in the relative fertility of infected and uninfected individuals. In column (1) of Table VIII below I examine fertility in each of the calendar years 16 to 20 years before the DHS survey, seeing how the logit probability of a birth in each of those years varies with a woman's infectious status at the time of the survey,

with the usual controls for education, marital status, age, etc.[35] In columns (2)-(4) I then look at fertility in the calendar years 11 to 15 years before, 6 to 10 years before and 1 to 5 years before the DHS survey, before finishing, in column (5), with fertility in the 12 months prior to the survey. As the reader can see, 16 to 20 years before the survey, when they could not possibly have yet been infected, the fertility of HIV positive women was not significantly lower than the average. However, as one gets closer and closer to the survey year, their relative fertility falls, so that, by the fifth column, one finds that HIV positive women have about a 30 percent lower probability of giving birth in the previous 12 months, a result well in keeping with those of the medical studies cited in the introduction to this paper. These results do not indicate whether the increasingly lower relative fertility of HIV infected women is due to the biological consequences of the disease, reverse causation, or a behavioral response to knowledge of their own infectious status. They do, however, clearly indicate that the decline in fertility recorded earlier in Table IV, which is at least 4 times greater than that estimated here, cannot possibly be coming from the response of HIV infected women alone. The trends reported in earlier tables reflect a broad and powerful communal response to the epidemic.

[35] As in the case of Table IV, I use the fertility histories to generate observations of a birth or no birth in each year of a woman's life going back to age 10. Each of these annual records, falling within the time periods specified, provides an observation. Education and urban/rural location are the reported DHS values, while age, age squared, marital status and the number and square of born and living children (coming into each calendar year) are calculated from the fertility histories and reported year of first marriage. I calculate, as usual, country specific coefficients for each of the controls. I do not include time trends, as the analysis is broken down by time period.

Since younger women obviously cannot provide records 15 to 20 years before the survey, the number and current age of women providing birth records changes as one moves from left to right in the table. This, however, does not appear to be a problem. If I reestimate columns (2) through (4) using only the data for the sample of women appearing in column (1), I get estimated HIV effects of -.101 (s.e.=.025), -.146 (.022), and -.329 (.028). The specification of column (5) cannot be estimated with this sample as there are very few observations, producing degenerate control categories where all of the women take a common action.

Table VIII: Relative Fertility of HIV Positive Individuals (logit model)

	Period before the DHS Survey				
	(1)	(2)	(3)	(4)	(5)
	16-20 Years	11-15 Years	6-10 Years	1-5 Years	Past Year
HIV positive	-.058	-.098	-.114	-.327	-.303
	(.031)	(.026)	(.023)	(.025)	(.055)
Sample Mean	.158	.160	.159	.148	.173
Pseudo R2	.210	.184	.185	.150	.155
N	116445	159406	211545	241170	48234

Notes: Each column includes a full set of country specific controls as in columns (3) and (6) of Table IV earlier, but without time trends (as the analysis subdivides time periods) and with tertiary and secondary education sharing a joint dummy (as the number of tertiary educated women is frequently exceedingly small). The dependent variable is whether or not a DHS woman gave birth in a given year. As in the other logit tables, I estimate a common coefficient for the impact of HIV across countries and then, to ease interpretation of the results, report the cross-country average of the within country derivative of the ln probability of a birth with respect to HIV status. As HIV is a discrete variable, this "derivative" reflects the estimated impact on fertility of all women in the sample moving from being HIV negative to being HIV positive. The higher sample mean in column (5) reflects the fact that the DHS measure of fertility in the past year covers 12 to 13 months and women aged 15-49 (as opposed to 10-49 in the retrospective histories), as explained earlier above.

IV. Human Capital Accumulation

While reductions in current fertility increase the per capita resources available to future generations, there are other mechanisms whereby the HIV epidemic might adversely affect the welfare of future cohorts, most notably through its impact on the accumulation of human capital. Clearly, the many orphans left by the epidemic are likely to receive less education. More generally, the higher expected mortality brought by the epidemic is reducing the return to potential human capital investments faced by parents and their children, which could lower overall investment in education. Thus, just as in the case of fertility, the HIV epidemic might lead to behavioral changes in human capital investment that generate effects well beyond the immediate physiological consequences of the virus. In this section I show that no such effects are discernible in the DHS data.

The household files of the DHS contain data on both current school enrollment and the educational attainment of all household members under the age of 25. As in the previous section, I will examine the impact of the average female HIV infection rate on these variables using three sets of controls: (1) country specific constants, dummies for the educational category of the household head,[36] and controls for the age, age squared and gender of the individual in question; (2) the above, plus country specific time trends; (3) all of the preceeding, plus country specific controls for urban/rural location and the presence of a radio, bicycle and electricity (each entered separately) in the household. For individuals 14 and younger, the DHS collects information on whether their parents are alive or not. Consequently, I split the sample into two groups, children aged 6-14 and youths 15-24, using country specific dummies on the survival status of parents as an additional set of controls for the younger age group.

[36] I use the educational attainment of the household head, as opposed to that of parents, as the head is always identified. This is particularly apropos for the cases where I estimate the impact of orphaning.

Table IX: School Attendance (logit model)						
	Ages 6-14			Ages 15-24		
	(1)	(2)	(3)	(4)	(5)	(6)
Head Primary	.364 (.006)	.364 (.006)	.270 (.006)	.445 (.012)	.444 (.012)	.294 (.012)
Head Secondary	.562 (.006)	.564 (.006)	.413 (.007)	.797 (.013)	.794 (.013)	.502 (.014)
Head Tertiary	.629 (.009)	.631 (.009)	.443 (.012)	1.09 (.016)	1.09 (.016)	.747 (.019)
Mother Dead			-.084 (.011)			
Father Dead			-.045 (.008)			
Both Parents Dead			-.109 (.027)			
HIV	1.23 (.200)	1.30 (.582)	.210 (.581)	.026 (.289)	1.48 (.827)	.251 (.810)
Sample Mean Pseudo R2 N	.606 .243 586371	.606 .254 586371	.611 .289 545247	.308 .239 373430	.308 .245 373430	.308 .272 364512

Notes: The first column for each age group includes a complete set of country-specific estimates of the impact of age, age squared, gender, and the educational category of the household head. The second column adds country specific time trends. The third column includes country specific dummies for urban/rural location and the presence in the household of a radio, bicycle or electricity (all entered separately), as well as (for children 6-14) whether the mother, father or both parents are dead. The reported coefficients and standard errors are for the average across the countries of the sample average derivative (or discrete change) of the ln probability with respect to the variable in question, as explained in the previous section. All standard errors are adjusted for clustering. Unless otherwise noted, these comments apply to all tables which follow in this section.

Table IX above estimates the impact of communal HIV infection rates on the probability a DHS individual is currently in school. Surpisingly, enrollment rates for children aged 6-14 appear to be positively associated with HIV infection rates, even when country time trends are added to the analysis. The significance of this relation, however, does not survive the introduc-

tion of additional controls for location, parental survival and household wealth. With regards to youths aged 15-24, the coefficients on HIV, while positive, are never significantly different from zero. These estimates suggest that, despite its effect on the return to investments in education, the HIV epidemic, per se, is not having a general negative influence on the accumulation of human capital. The epidemic does, however, clearly create orphans who with the loss of both of their parents might be expected, based upon the estimates of the table, to receive 24 percent fewer years of education.[37]

In sub-Saharan Africa, school enrollment and educational attainment are by no means equivalent, as there is frequent failure and repetition of grades.[38] In this context, it is clearly important to estimate the impact of HIV not only on an input (student time), but also on realized educational outcomes. To this end, Table X uses an ordered logit model to estimate the influence of communal HIV infection on a DHS individual's progression through the ordinal categories of no schooling to primary, then secondary and, in the case of young adults, possibly tertiary

[37] I include country specific dummies for dual orphans (with both parents dead) since Case et al (2004) find an increased effect from the death of both parents, as do I. Although cross sectional studies of African couples find that .3 to .4 of the partners of infected males and females are HIV negative (Carpenter et al 1999, Gray et al 1998, Hira et al 1990 and Serwadda et al 1995), seroconversion over time suggests that most AIDS orphans will ultimately be dual orphans.

[38] On this, see the surveys of the Monitoring Learning Achievement Project (Chinapah et al 2000 and Strauss 1999), which find, for example, that more than .20 of 4th grade students in Madagascar, Malawi, and South Africa had repeated grade 1, while .15 or more of 4th grade students were repeating that grade in Botswana, Malawi, Senegal, South Africa, Tunisia, and Uganda. In the extreme, in Senegal, .07 of 4th grade students were in the midst of their 4th or greater attempt at completing that grade. Across my DHS sample, cumulative average enrollment rates by age group from age 6 on imply a sample average educational attainment of 5.5 years by age 15 and 8.2 years by age 24, as compared with the actual sample average attainments of 4.0 years and 5.5 years at those ages, respectively.

Table X: Ordinal Educational Attainment (ordered logit model)

	Ages 6-14			Ages 15-24		
	(1)	(2)	(3)	(4)	(5)	(6)
Head Primary	.907 (.014)	.905 (.014)	.722 (.015)	.885 (.016)	.895 (.016)	.644 (.015)
Head Secondary	1.77 (.019)	1.79 (.019)	1.28 (.020)	1.96 (.019)	1.95 (.019)	1.26 (.019)
Head Tertiary	2.59 (.036)	2.62 (.036)	1.83 (.039)	2.88 (.033)	2.88 (.033)	1.90 (.034)
Mother Dead			-.127 (.025)			
Father Dead			-.033 (.018)			
Both Parents Dead			-.205 (.054)			
HIVAVG	1.89 (.416)	2.32 (.706)	1.42 (.714)	2.89 (.287)	.120 (.954)	.711 (.922)
C(secondary)	5.86 (.022)	5.91 (.022)	6.27 (.023)	2.64 (.014)	2.65 (.014)	2.90 (.015)
C(tertiary)				6.87 (.028)	6.89 (.029)	7.32 (.030)
Sample Mean						
None	.363	.363	.361	.266	.266	.266
Primary	.615	.615	.618	.450	.450	.452
Secondary	.022	.022	.020	.274	.274	.272
Tertiary				.010	.010	.010
Pseudo R2	.262	.269	.309	.185	.187	.238
N	616596	616596	574454	397433	397433	388342

Notes: Since attainment is a stock variable, I use the average country HIV infection rate from the time the individual was age 6 to the year of the DHS survey (HIVAVG) as the independent variable. Since it is difficult to summarize ln probability derivatives for multiple categories, I depart from my practice in the logit tables and report, in this table, the actual coefficient estimates. In the case of the education and mortality dummies, the reported coefficients are the equally weighted average of the country specific coefficients (and the standard error of the same, i.e. not the average of the standard errors).

education.[39] I find that the average communal HIV infection rate present during an individual's educational life (i.e. from age 6 to her/his current age)[40] is positively associated with educational attainment. While this relation is not robustly significant for youths, it remains statistically significant for children even when all the additional conditioning variables are added to the equation. As an alternative, in Table XI I abandon ordinal measures in favour of the DHS's somewhat problematic quantitative measure of realized years of attainment.[41] Interestingly, this variable produces diametrically opposite results as, with all conditioning variables, HIV is now

[39] In the ordered logit model an individual's choice is determined by the value of a latent variable xB + e, where e is a logistically distributed individual error term, relative to decision cutoff points. I normalize constants so that the cutoff from no education to primary equals 0, and then estimate a primary to secondary cutoff C(secondary) and, for youths, a secondary to tertiary cutoff C(tertiary). The probability of no education is equal to 1/(1+exp(xB)), the probability of any intermediate educational outcome I is 1/(1+exp(xB-C(I+1))) - 1/(1+exp(xB-C(I))), and the probability of the highest educational outcome H is exp(xB-C(H))/(1+exp(xB-C(H))). I report actual coefficient estimates, as opposed to the ln probability derivatives listed in the logit tables, as there are too many categories to summarize easily. The reported coefficient estimates for the head of household education and parental mortality are the equally weighted average of the country specific dummies.

[40] Since educational attainment is a stock variable, a consequence of years of investment, I use the average annual HIV infection rate from age 6 to the individual's current age as the independent variable in this table and those that follow in this section.

[41] The DHS tries to calculate realized years of educational attainment by combining each individual's report of their highest level of education attended (e.g. primary or secondary) with the number of years completed at that level. Thus, for example, an individual who completed 3 years at secondary might be calculated, by the DHS, to have completed 9 years of education. The problem is that individual reports on years completed at their final level often refer to attendance rather than completion, as evidenced by the fact that large numbers of individuals report completing more than 8 years at the primary or secondary level. Consequently, the DHS measure is a mixture of attainment and attendance.

To add to the confusion, the DHS uses highly inconsistent coding *within countries*, e.g. giving a base of 6 years as completed primary in one survey, 7 in the next, and 6 in the following, or giving different base years when recoding the male questionnaire, female questionnaire and (when viewing the same individuals) household questionnaire, all within the same survey. The resulting variable reported by the DHS, hv108 "education in single years", is remarkably inconsistent. I solve this problem by recalculating years of educational attainment using consistent base quantities for primary and secondary education across surveys and individuals (I determine the base quantities, which vary country by country, by using the most common adjustment applied by the DHS within that country). This still leaves a measure which is a mixture of attainment and attendance (as described above), but at least removes the coding inconsistencies.

Table XI: Years of Educational Attainment (poisson model)						
	Ages 6-14			Ages 15-24		
	(1)	(2)	(3)	(4)	(5)	(6)
Head Primary	.375 (.006)	.371 (.006)	.244 (.006)	.364 (.007)	.364 (.007)	.222 (.006)
Head Secondary	.693 (.007)	.691 (.007)	.402 (.006)	.688 (.007)	.681 (.007)	.371 (.006)
Head Tertiary	.904 (.008)	.896 (.008)	.498 (.008)	.880 (.008)	.874 (.008)	.469 (.008)
Mother Dead			-.066 (.009)			
Father Dead			-.007 (.006)			
Both Parents Dead			-.047 (.019)			
HIVAVG	.554 (.111)	-2.50 (.260)	-2.70 (.246)	.585 (.092)	-.554 (.319)	-.563 (.290)
Sample Mean Pseudo R2 N	1.64 .311 616109	1.64 .314 616109	1.63 .338 574357	4.91 .181 397161	4.91 .183 397161	4.89 .229 388147

seen to have a significantly negative influence on the educational attainment of both children and youths. The results of Table XI indicate that dual orphans receive about 12 percent fewer years of education, the discrepancy with the results of Table IX reflecting the difference between attendance and attainment in the sub-Saharan African context.

The ordered logit and poisson count models provide poor descriptions of the decisions involved in human capital accumulation as they attempt to characterize the entire process with a

simple distribution.[42] To provide a more nuanced specification of human capital accumulation and, in all honesty, to further illustrate the gross inconsistencies between the results of different specifications, Table XII below breaks downs the process into phases. In the first panel (three columns) of the table I examine the probability that an individual has ever been to school. As the fraction of the sample that has never been to school levels out at age 11, I use the average HIV infection rate between the ages of 6 and 11 as the independent variable,[43] under the assumption that individuals who have not been to school by age 11 no longer have a reasonable opportunity to do so. In the next panel I examine the probability that individuals aged 12 and over have made it to secondary school, conditional on having attended primary (i.e. conditional on a positive decision in the first panel). Finally, in the third panel I run a poisson analysis of years of realized educational attainment conditional on having attended school (again, a positive decision in the first panel).

The results in the table speak for themselves. For children aged 6-14, the average HIV infection rate present during their educational life has, once all conditioning variables are entered, no significant influence on the probability they have ever attended school or made it to secondary. However, conditional on having attended school, HIV has negative effects on total years of attainment (despite having no influence on the probability of making it to secondary). For youths currently aged 15-24, once the full panoply of controls are added, the HIV infection rates present during their childhood are seen to have a significant negative influence on the probability they ever attended school, but also a very strong and significant positive influence on their probability of reaching secondary conditional on having made it to school. However, despite this

[42] The poisson count model in particular, while matching the mean number of years of schooling, fails to predict the large mass of individuals who have zero years of education, i.e. have never attended school.

[43] Or, for individuals under 11, the average infection rate from age 6 to their current age. In a similar fashion, the age terms are the actual age for individuals under 11, and 11 years of age for individuals over 11. Thus, the logit is essentially one of whether an individual, between ages 6 and 11, ever started school.

Table XII: Phases of Education									
	Ever Attended (logit model)			Reached Secondary (logit model)			Completed Years (poisson model)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Ages 6-14								
Head Primary	.347	.348	.268	.417	.428	.213	.091	.090	.049
	(.005)	(.005)	(.005)	(.048)	(.048)	(.051)	(.004)	(.004)	(.004)
Head Secondary	.521	.526	.394	1.46	1.45	.875	.304	.304	.177
	(.006)	(.006)	(.006)	(.046)	(.046)	(.052)	(.004)	(.004)	(.004)
Head Tertiary	.586	.592	.430	2.18	2.20	1.43	.489	.483	.289
	(.006)	(.006)	(.009)	(.054)	(.054)	(.064)	(.006)	(.006)	(.006)
Mother Dead			-.045			-.308			-.028
			(.009)			(.091)			(.006)
Father Dead			-.016			-.010			.006
			(.007)			(.046)			(.004)
Both Parents Dead			-.090			.164			-.004
			(.021)			(.146)			(.012)
HIVAVG	.461	.146	-.157	3.41	5.81	1.63	.286	-2.59	-2.79
	(.161)	(.267)	(.270)	(.651)	(3.56)	(3.90)	(.089)	(.193)	(.182)
Sample Mean	.637	.637	.639	.084	.084	.079	2.58	2.58	2.55
Pseudo R2	.284	.294	.332	.287	.291	.331	.226	.230	.239
N	616596	616596	574454	147813	147813	137795	391993	391993	366706
	Ages 15-24								
Head Primary	.335	.338	.256	.225	.235	.096	.071	.072	.030
	(.006)	(.006)	(.006)	(.011)	(.011)	(.011)	(.004)	(.003)	(.003)
Head Secondary	.465	.467	.333	.794	.791	.483	.277	.273	.149
	(.006)	(.006)	(.006)	(.011)	(.011)	(.010)	(.004)	(.003)	(.003)
Head Tertiary	.502	.505	.335	1.04	1.04	.693	.437	.430	.252
	(.007)	(.007)	(.009)	(.012)	(.012)	(.014)	(.005)	(.005)	(.005)
HIVAVG	1.02	-.712	-.658	2.94	3.05	2.68	.505	-.416	-.460
	(.093)	(.197)	(.192)	(.233)	(.861)	(.821)	(.061)	(.223)	(.202)
Sample Mean	.721	.721	.721	.388	.388	.384	6.68	6.68	6.66
Pseudo R2	.274	.279	.327	.206	.213	.263	.075	.078	.096
N	377815	377815	368839	291840	291840	285051	291568	291568	284856

Notes: HIVAVG = For columns (4)-(9), as in Table X. For columns (1)-(3), the average country HIV infection rate from the time the individual was age 6 to the year of the DHS survey or age 11, whichever comes first. The individual age variable in columns (1)-(3) also reaches a maximum at age 11 (i.e. the logit is of the probability an individual, by age 11, ever went to school). For the children's sample of columns (4)-(6) the analysis is restricted to individuals aged 12-14.

positive influence on the probability of reaching secondary, HIV infection rates have a negative effect on years of attainment conditional on having attended school. These mutually inconsistent results are usefully compared with those for fertility listed in the previous section where, across a variety of measures and specifications, the communal HIV infection rate has a consistently negative influence.

Without wishing to offend the reader, I interpret the grossly inconsistent influence of the communal HIV infection rate on measures of educational enrollment or achievement presented above as simply reflecting spurious correlations with otherwise exogenous non-linear trends.[44] Unlike its influence on fertility, the simple presence of the HIV epidemic in the community at large does not appear to exert any consistent dampening effect on the accumulation of human capital.[45] This is not to say that it will not adversely, through parental mortality, affect the endowments of future cohorts. The HIV epidemic will leave large numbers of orphans, which will reduce the average human capital of the cohorts which mature in its shadow. Outside of this mechanism, however, the DHS data are not suggestive of a generally negative impact of the epidemic on human capital accumulation.

[44] In principle, the results on fertility can also be dismissed as reflecting spurious correlations with non-linear time trends. However, different measures of a concept have slightly different non-linear trends. When the direction and significance of the influence of HIV on a concept is highly sensitive to the measure used (as in the case of education), it is obvious evidence in favour of spurious correlations. However, when the sign and magnitude of the influence of HIV is consistent across a number of diverse measures (as in the case of fertility discussed above), I think this is suggestive of actual causality.

[45] This might reflect a failure of human capital investment decisions to respond to mortality risk or, alternatively, a compensating shift towards "quality" on the part of surviving parents, who are now having considerably fewer children.

V. Supporting Patient Care

In this section I calculate the maximum expenditures on AIDS patient care consistent with maintaining the living standards that would have been enjoyed absent the epidemic. My simulations show that the reduced dependency ratios and capital dilution effects brought by the fall in fertility dominate the harmful effects of adult mortality on participation and the human capital accumulation of children, allowing extraordinary humanitarian expenditures on medical care without any expectation of economic benefit whatsoever. I begin by presenting my key population and macroeconomic assumptions, before reporting the results.

(a) Assumptions

I simulate stylized economies consisting of cohorts aged 0 to 97 who consume, work, reproduce, educate their young, age, and die of HIV and non-HIV causes. I estimate the key fertility, mortality, education, participation, and relative labour productivity equations of these economies from the DHS files, set their macroeconomic characteristics to match long term historical data, and then simulate their evolution in the presence and absence of the HIV epidemic. I eschew educational averages and subdivide each cohort into 17 explicit educational categories (0-16 years of education), with special emphasis on the probability of no education at all. Less educated women tend to have more and less educated children, so that the variance of educational outcomes is important, with a mean preserving spread in educational attainment worsening average outcomes for subsequent generations. By keeping track of the most detailed educational distribution possible I maximize the long term economic impact of orphaning. I have sufficient data to calculate lifecycle patterns of behavior and run simulations for 27 of my 29 DHS countries.[46]

[46] In the table that follows, I use all available DHS country data to estimate lifecycle patterns of behavior, but can only put together a full set of country parameter estimates for 27 countries. The two missing countries are Burundi and Liberia.

Table XIII: Lifecycle Equations Estimated from DHS Data

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Fertility	Mortality		Education		Participation		Wages
	logit	logit	logit	logit	poisson	logit	logit	regression
	Ages 12-49	Ages 0-11	Ages 12+	Attended Ages 6-11	Attainment Ages 12-22	Males Ages 22+	Females Ages 22+	M & F Ages 22+
Age	.598	-.736	.008	2.51	.418	.245	.158	.087
	(.001)	(.003)	(.001)	(.032)	(.005)	(.008)	(.005)	(.009)
Age^2	-.011	.038	.000	-.118	-.010	-.003	-.002	-.001
	(.000)	(.000)	(.000)	(.002)	(.000)	(.000)	(.000)	(.000)
Education	-.006					-.050	.027	.095
	(.001)					(.007)	(.004)	(.008)
$Education^2$	-.005					.000	.001	.003
	(.000)					(.000)	(.000)	(.000)
Mother's Education				.304	.038			
				(.005)	(.001)			
Mother's $Education^2$				-.009	.000			
				(.000)	(.000)			
Sex		-.109	-.296	-.160	.036		.159	
		(.004)	(.008)	(.010)	(.002)		(.166)	
HIV	-1.46							
	(.034)							
Sample Mean	.175	.026	.004	.572	4.84	.681	.681	10.5
Pseudo R2	.096	.138	.014	.255	.127	.112	.112	.993
N	8920049	10449633	28651772	302951	141354	475913	475913	9358

Notes: Education equals years of attainment, while sex equals 1 for females. For the adult mortality equation, age is age-12. For the fertility and mortality equations, where I use histories of fertility and survival and death, N (the number of observations) denotes the number of person x years of observation. For the mortality equation, the model is not strictly logit (although the probability of death from non-HIV causes in each year is), as the likelihood is modified to take into account the (independent) probability of death from AIDS, as detailed in the supporting appendix. For the wage regression the R2 is not pseudo. Each equation includes a complete set of country dummies. The coefficients listed above are the actual estimates, as the derivatives of the ln expected value (presented in earlier tables) are less informative for the quadratic forms used above.

Table XIII above presents the basic lifecycle equations estimated off of DHS data.[47] I begin, as shown in column (1), by estimating fertility from ages 12 to 49 as a function of age, years of educational attainment and the average HIV infection rate.[48] Across the 27 sub-Saharan countries in my simulation, the average woman with 0 years of education gives birth to 6.9 children, with lifetime fertility falling to 6.0, 3.9, and 2.5 for women with 6, 12 and 16 years of education, respectively. The coefficient on HIV, -1.46,[49] translates into an average country derivative of the ln probability of giving birth with respect to the infection rate of -1.21, which is in the lower range of the results found with census data in the Introduction and with DHS/WFS microdata in Section III. In the simulations that follow, I will consider cases where the HIV coefficient is arbitrarily lowered to -.73 (i.e. a ln derivative of -.60), to show the robustness of the conclusions to a weaker influence, and 0, to show their definite dependence upon some fertility effect.

To continue, the DHS women's reports of the survival status of their children and siblings allow me to estimate the probability of death from causes other than HIV in each year of a per-

[47] The construction of variables and samples is explained in the appendix. Each equation includes a complete set of country dummies. Unlike the analysis earlier in the paper (where my intent was to test the robustness of the HIV variable), I do not include country interactions with the age, sex or education terms. As the samples are often quite thin, such interactions occasionally produce bizarre patterns which would make the simulations, for some countries, somewhat nonsensical. I should also note that, to keep the presentation simple, I generally do not interact sex with the age and education variables (i.e. estimate separate gender profiles) as I have found, in almost every case, that doing so produces very similar predictions to those arrived at with equations with simple gender dummies, but no interactions (see the supporting appendix available from the author). The exception is the participation equation, where male and female educational profiles have different signs, as reported below.

[48] Aside from being a reasonable average age for menarche, age 12 is also computationally convenient as I assume that all children infected at birth with HIV die by age 12 (see below). The DHS does not provide birth reports beyond age 49, by which point, in any case, fertility is negligible.

[49] For the quadratic forms presented in the table, the derivatives of the ln probability with respect to the independent variables are less informative, as they treat the linear and quadratic components as varying independently and dont indicate at which point the influence of the variable in question changes sign. Consequently, I depart from my practice elsewhere in the paper and report the actual coefficient estimates.

son's life as a function of their sex and age (columns (2) and (3) of Table XIII).[50] I divide the mortality estimates into two groups, ages 0-11 and 12 and above, so as to allow simple quadratics in age to capture the extremely rapid decline of mortality early in life and its more gradual rise later on. Across the 27 simulation countries, in the absence of the epidemic the average probability of dying in the first year of life is .084, while life expectancy at birth is 59.6 years for men and 64.2 years for women.

I model children's educational outcomes, in columns (4) and (5) of Table XIII, as a function of their mother's education.[51] As noted in the previous section, the proportion of my DHS sample that has ever attended school levels out at about age 11. Consequently, I divide the educational process into two segments. First, I use a logit model to estimate, amongst children currently aged 6 to 11, the probability an individual has ever attended school as a function of their age, sex and mother's educational attainment. Next, I use a poisson count model to calculate the probability distribution of educational attainment conditional upon having attended

[50] As the epidemic is ongoing, the estimating procedure has to adjust for mortality due to HIV, which I take as being independent of mortality from non-HIV causes. For children under 12, the likelihood equation includes a Bayesian calculation of the probability the mother, at the time she gave birth, was HIV infected (given that she is now alive), and the impact this would have on her child's mortality. For persons 12 and above, the likelihood adjusts for HIV infection and associated mortality amongst adults. The full procedure is laid out in the supporting appendix, available from the author.

I should note that, with regards to mortality at early ages, the DHS sibling survival reports suggest much lower mortality than is indicated by mother's reports of the survival of their children. There is likely to be a downward bias here, as individuals are less likely to remember siblings who died when very young. For this reason, in estimating mortality between the ages of 0-11, I disregard the sibling data and use the childrens' survival reports alone.

[51] Earlier in the paper, where I wished to assess the impact of orphaning, I used the education of the household head as the determinant of children's outcomes. However, in terms of calculating an evolving population distribution, it is not easy to model who becomes a household head and which children reside in which household. Consequently, I model each child's outcome as a function of their mother's education. I then incorporate adjustments for the impact of orphaning, as explained further on.

school for individuals aged 12 to 22, assuming that education is completed by age 22.[52] In the simulations which follow, I fold the rather thin upper tail of this distribution into the 16 years of education category (i.e. individuals whose predicted education is greater than 16 years are treated as having 16 years of education).[53] With regards to the impact of orphaning, I append the adjustments implied by my estimates earlier in Table XII. Thus, I assume that if an individual's mother, father or both parents die before they complete their education (at age 22), the ln probability of their ever having attended school falls by -.045, -.016 and -.151, respectively, while, in a similar fashion, the ln expected number of years, conditional upon having attended school, changes by -.028, .006, and -.026. Together, these adjustments predict that AIDS dual orphans will acquire 18 percent less years of educational attainment.[54]

Finally, I use the DHS data to estimate patterns of participation and labour income. Rather than calculate a complicated model involving simultaneous choices amongst schooling and market participation, I simplify the analysis by assuming that schooling takes place prior to age 22 and work choices arise thereafter. Since male and female participation patterns are quite different, I include a sex dummy and allow each gender to have separate age and educational

[52] For the simulations, I'm actually only interested in an individual's ultimate educational attainment at age 22. For estimation purposes, however, using younger age groups (where it is easier to identify mothers), provides a much broader sample. The educational attainment of 22 year olds is then calculated using the equation of column (4), with age set at a maximum of 11, and column (5), with age set at 22. In the supporting appendix, available from the author, I show that my fertility, mortality and education equations and the DHS population distribution all form an internally consistent whole, in that they can produce (as the predicted offspring of the older cohorts in the DHS sample) the age and educational distribution of individuals 0 to 21 in the DHS sample. Thus, these equations are good estimates of recent patterns of behavior.

[53] This procedure, using a poisson count model and then collapsing the thin top tail, is considerably easier (in terms of simulation) than using the coefficients of an ordered probit or logit model to limit the formal probability distribution to the 17 educational categories.

[54] This effect is stronger than the 12% reduction I found in Table XI, where I simply ran a poisson of years of attainment (i.e. did not separate the educational process into two equations).

profiles.[55] As shown in columns (6) and (7) of Table XIII, sub-Saharan male participation tends to decline with education, while female participation tends to rise with years of attainment.[56] From 0 to 16 years of educational attainment, average years of participation between the ages of 22 and 65 falls from 37.2 to 32.6 years amongst men, while amongst women it rises from 26.8 to 33.5 years. I conclude with a regression linking market incomes to age and educational attainment (column (8) of Table XIII). I exclude sex from the regression, as I find no significant differences between genders. Lifetime incomes rise by 155 percent from 0 to 12 years of education, and an additional 70 percent if an individual reaches a full 16 years of attainment.[57]

Turning to the HIV epidemic, while the Census Bureau data provide annual estimates of infection rates among women aged 15-49, these data do not indicate when the women were infected, which is necessary to calculate mortality patterns. Taking into account mortality due to AIDS and the entrance of new, uninfected, women into the 15-49 population average, I calculate the annual transmission rate to previously uninfected women aged 15-49 implied by the Census

[55] However, I estimate the male and female participation equation jointly, i.e. there is common country dummy for both genders and, thus, columns (6) and (7) refer to one estimating equation. This is because for some countries the DHS has not collected male participation data and I must, consequently, use the average male and female differences estimated off the entire sample to predict male participation rates.

[56] Prior to age 22, educational attainment is strongly negatively associated with participation in both sexes, as better educated persons are more likely to be in school.

[57] The regression is run using data for individuals who work for others (i.e. not for themselves or family). Labour income data are only available for a small subset of the countries. However, as I will be assuming a common labour factor income share for all of the countries (which determines mean incomes), these estimates are only used to calculate a common relative labour efficiency profile. I should also note that these estimates do not adjust for selectivity bias (Killingsworth and Heckman 1986). I have run more complicated (joint maximum likelihood) systems, augmenting the wage equation with probit selection equations for men and women, where the probability of working is determined by the variables in the wage equation plus marital status, the presence of electricity or a radio, television, refrigerator, bicycle, motorcycle or car (each entered separately) in the household, the number of the individual's children living in the household, and (for women) births in the past year and in the past 5 years, allowing the male and female selection error terms to have separate covariances with the wage error term. The age and educational profiles estimated with the selection adjustment are virtually identical to those reported above, while gender remains completely insignificant (see the supporting appendix available from the author). Consequently, to keep things simple, I ignore selection issues and report the direct estimates, using data for individuals working for others, listed above.

Bureau data.[58] I then use these annual entrance rates to develop cohort HIV histories, under the assumption that women are only at risk between the ages of 15-49. While knowledge of male infection rates is fairly limited, studies of African couples consistently find that about .3 to .4 of the male partners of HIV infected females and .3 to .4 of the female partners of infected males are themselves HIV negative (Carpenter et al 1999, Gray et al 1998, Hira et al 1990 and Serwadda et al 1995). This suggests that, adjusting for the age difference between male and female partners, male and female infection rates are roughly the same, an assumption that is borne out by DHS seroprevalence data.[59] Within my DHS sample, women report that on average their partners are 9 years older than they are, so I assume that male cohort infection rates parallel those of female cohorts 9 years younger in age.[60] DHS prevalence data indicate no consistent pattern

[58] The full procedure is laid out in the supporting appendix, available from the author.

[59] Altogether, I have access to 14 DHS surveys with data on the HIV status of the respondents (see the appendix at the end of this paper). The average across these of the average within country male infection rate is .048, while the average of the female infection rates is .072. This suggests that male infection rates are generally lower. These figures, however, include large numbers of uninfected males under the age of 24. When one compares the female population aged 15-49 to the male population aged 24-58 (using the 9 year age difference between male-female partners discussed further on), the cross country average male and female infection rates are .070 and .072, respectively. For 12 of these surveys, I am able to link the HIV data files to the household files (including the location and clustering information), allowing me to test the equivalence of the male/female infection rates with standard errors adjusted for clustering. Comparing females aged 15-49 with males 24-58 in a logit analysis of the individual probability of infection run on sex and a country dummy, in only one country is the female sex dummy significantly positive at the 1% level and in three countries it is actually estimated to be negative (i.e. female rates are lower). Pooling the entire sample, i.e. running a logit of individual infection on country dummies and sex, the coefficient on female sex is .066 (s.e. = .030), which is not significant at the 1% level.

[60] Thus, for example, male cohorts at age 29 have the infection rate of contemporaneous female cohorts aged 20 born nine years after them. In practice, this means that overall male infection rates are lower (as indicated by the unadjusted DHS data discussed in the preceeding footnote), as older cohorts are smaller in size.

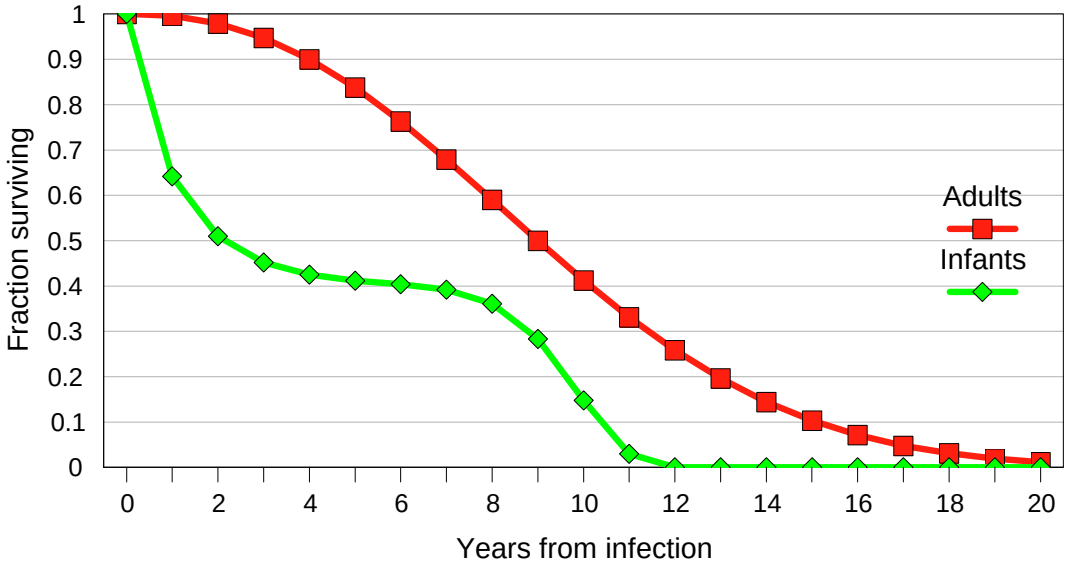
across educational groups, so I assume equal infection rates across educational categories.[61] With respect to the future evolution of the epidemic, I assume that with medical breakthroughs or behavioral change the rate at which previously uninfected females aged 15-49 are infected declines sinusoidally to zero from 2005 to 2050.[62]

Regarding mortality, adult individuals infected with HIV typically experience a short period of flu-like symptoms, after which they remain, superficially, asymptomatic, until their immune system collapses and succumbs to opportunistic infections and cancers. I use as my baseline the WHO-UNAIDS (UNAIDS Reference Group 2002) recommendations on adult survival times absent retroviral therapy, which suggest a median survival after infection of 9 years and virtually no probability of survival beyond 24 years (see Figure II below). While mother-to-child transmission can be reduced through drug therapy, cesarean delivery and avoidance of breastfeeding, these are cumulatively prohibitively costly and currently about 1/3 of children born of HIV positive mothers in African countries are infected, a proportion I assume will remain constant for the foreseeable future. For infected children, I follow the WHO-UNAIDS (2002) recommended children's mortality profile, which reproduces the appallingly rapid progression of pediatric AIDS, with 1/3 of infected children dying from the virus within one year of

[61] For the 12 DHS HIV surveys that I can link to household files, with information on clustering and the educational attainment of individuals, and limiting the sample to women aged 15-49 and men 24-58 (see above), I find that a pooled logit of individual HIV status on country dummies and educational attainment yields education coefficients of primary .334 (s.e.=.047), secondary .516 (s.e.=.053), and tertiary .163 (s.e.=.120), with no education as the base category. This suggests that infection rates are substantially lower in the least and most educated groups. Looking at the countries individually, however, I find that this pattern is only reproduced in 4 countries, with others showing lower infection rates in those with primary or secondary education than in those with tertiary education. Since there is very little data on infection rates by education, it is quite difficult to model the evolution of the epidemic by educational cateory. Consequently, I make the assumption of uniform infection described above. If indeed the upper and lower tails of the educational distribution have the lowest infection rates, then my assumption simply averages this out.

[62] Thus, if I(t) is the fraction of the uninfected female cohort aged 15-49 (and male cohort aged 24-58) in year t which becomes infected, then I(t) = sin[(pi/2)(2050-t)/45]*I(2005), where I(2005) is calculated off of the census data, as described above.

Figure II: Cumulative Survival Rates



S: UNAIDS [2002].

birth and none living to see their 12th birthday (Figure II).[63]

Together, the lifecycle behavioral equations and cohort HIV histories calculated above allow me to simulate the evolution of the population of each country in my sample in the presence and absence of the HIV epidemic. I initialize each country's population using the age x educational distribution and weights recorded in that country's earliest DHS survey.[64] Making use of each cohort's cumulative mortality estimate (including the effect of HIV) I begin by projecting backwards to determine each cohort's original size at birth. I then run simulations forwards, in the presence or absence of the HIV epidemic, charting the evolution of the population as it ages, reproduces, educates its young, participates economically and dies. In these simulations I keep track of the number of children orphaned before the age of 22 and the impact

[63] For sources on disease progression in adults and children and rates of mother to child transmission see the footnote at the end of Section II above.

[64] There is a marked preference amongst DHS respondents for ages which are multiples of 5 and those selecting such ages tend to be considerably less educated. Further, the age and educational distribution of elderly persons is extremely jagged, as the samples are quite thin. In addition, the DHS surveys, like most sub-Saharan surveys and censuses, tend to undercount the extremely young, so the population count under age 5 is quite low. The current educational distribution of those under 22 is also not relevant for the simulations, as I need to know their ultimate educational attainment to predict their behavior. Finally, in projecting the evolution of the economy absent the HIV epidemic I occasionally have to predict the number and educational attainment of very old cohorts, who are no longer alive at the time of the earliest DHS survey.

I handle these problems by producing a smoothed population over the age of 22 and then using it to predict the number and future educational attainment of persons under the age of 22, as well as all future cohorts. First, I run equations on the probability of ever having attended school and the number of years of attainment conditional on having attended school on the birth year of the DHS male and female populations (separately) of persons aged 22 to 60 (whose education is completed). I then calculate the original birth cohort size of individuals 5 to 60 in the base (earliest) DHS survey (taking into account mortality from all sources), and run a regression of the ln cohort size on birth year. Together, these two sets of equations allow me to predict the original cohort size and ultimate educational attainment of all cohorts born 22 years of more before the earliest DHS survey (as shown in the supporting appendix, these predictions closely match the average DHS population characteristics). With this smoothed population in hand, I can then project forward the size and educational characteristics of subsequent cohorts using the equations in Table XIII.

this has on their educational attainment and, by extension, lifecycle behavior.[65] I assume that the presence of HIV in the community lowers fertility and produces orphans, but otherwise does not influence communal economic behavior.[66] As the epidemiological pattern of adult HIV is one of prolonged apparent health, followed by extremely rapid deterioration and death,[67] I assume that HIV infected individuals function normally, in all respects, until the year they develop AIDS and die. I then calculate, as explained further below, the maximal expenditure on medical care in the year of each AIDS victim's death consistent with maintaining the no-HIV path of consumption. In these calculations, I explicitly avoid any assumptions that allow patient care to "pay for itself" by improving economic outcomes.[68] This emphasizes my point, that sufficient resources exist to

[65] For example, in the absence of the HIV epidemic, assuming that spousal mortality rates are independent of each other, the probability a child born to a women of age a is left a dual orphan by age 22 is (1-S[m,a+9])(1-S[f,a]), where S[i,j] is the 22 year survival rate of sex i from age j to age j + 22. In contrast, with the epidemic the probability of being left a dual orphan is H[a]{1-Sh[a]+S[h,a](1-S[m,a+9])(1-S[f,a])} + (1-H[a]){1-Snh[a]+Snh[a](1-S[m,a+9])(1-S[f,a])}, where H[i] is the probability that the mother of age i was HIV infected at the time of the child's birth, Sh[i] is the probability an HIV infected women of age i and her equally infected partner of age i + 9 do not die of HIV causes in the next 22 years (virtually zero), and Snh[i] is the probability an uninfected women of age i and her partner do not die of HIV causes (only) in the next 22 years. Similar equations are used to calculate the probability of being a maternal or paternal orphan and of reaching age 22 with both parents alive.

[66] The lack of any coherent influence of HIV infection rates on schooling (other than through orphaning) has been explored above. With regards to participation, in my DHS sample the impact of communal HIV infection is also inconsistent, as econometrically it has an enormous positive influence on female participation and an equally enormous negative influence on male participation (results in the supporting appendix). As there is very little data on participation in the DHS, and these effects are largely mutually cancelling, I treat these correlations as spurious and assume that participation is unaffected by communal HIV infection. As for consumption, the whole point of my simulations is going to be to see whether non-HIV consumption levels can be sustained while expenditures are made to fight the epidemic.

In my analysis of South Africa (Young 2005), I argued that increases in wages, brought about by shortages of labour, will eventually lead to behavioral changes (such as a reduction in fertility and rise in economic participation), but also noted that such effects are likely to be small relative to the direct impact communal HIV infection has on fertility and orphaning. Consequently, in this paper, in extending the analysis to sub-Saharan Africa as a whole, I put these issues aside and focus on the direct influence on fertility and orphaning alone.

[67] Morgan et al (2000) report a median survival time from the development of AIDS to death in Uganda of 9.3 months. On the long period of asymptomatic apparent health, see http://www.avert.org/hivstages.htm (10/6/2005).

[68] E.g., by keeping adults healthy and productive for a few more years, thereby sustaining economic participation levels and improving the educational outcome of eventual orphans.

fight the epidemic while safeguarding the welfare of future generations.

To estimate the ability of the sub-Saharan economies to support AIDS patient care, I embed the population dynamics described above in a simple macroeconomic framework. Say that income in each economy is a function of capital and effective labour:

$$(1) \quad Y = F(K, EL), \quad \text{where} \quad EL = \sum_i W_i L_i .$$

Effective labour is given by the sum of labour supplies differentiated by age and education, with fixed relative weights given by my estimates, in column (8) of Table XIII above, of relative labour productivities. Assuming a constant savings path, absent the epidemic the development of capital per effective worker follows the equation:

$$(2) \quad n_{t+1} k_{t+1} = sF(k_t, 1) + (1 - \delta) k_t$$

where k_t is capital per effective worker in year t and n_{t+1} the ratio of effective workers in year t+1 to those in year t. Consumption per capita, or living standards, is calculated on an "equivalent adult" basis, under the assumption that children aged 0-14 have consumption needs equal to .5 that of an adult:

$$(3) \quad C_t = \frac{(1-s)F(k_t, 1)EL_t}{Pop_t - .5 * Ch_t}$$

where Pop is the total population and Ch the population aged 0-14. The equivalence adjustment substantially weakens my results, as, given the lower assumed consumption of children, reductions in fertility convey smaller consumption benefits.[69]

In my simulations, I begin by calculating the paths of k_t and C_t in the absence of the HIV epidemic, and then ask what is the maximal expenditure on medical care each economy could

[69] For this reason, I follow Deaton's (1997) recommended value of .5, although, as he notes, there are methodological problems with equivalence scales and, once one adjusts for economies of scale in household size, equivalence results can actually be interpreted as giving much higher relative costs of children. I have experimented with applying the equivalence adjustment to youths aged 0-21 and/or using an equivalence adjustment of .7 (such as that used by Cutler et al (1990)), both of which improve my results (the former because the no-HIV/HIV relative ratios of youths move less than the relative ratios of children). I opt for the least favourable assumption, as presented above.

make per AIDS victim in the year of that person's death consistent with giving everyone else the no-HIV path of consumption. In this case, the evolution of capital per effective worker is governed by the equation:

$$(4) \quad n^h_{t+1} k^h_{t+1} = F(k^h_{t+1}, 1) - C_t\left(\frac{Pop^h_t - .5*Ch^h_t)}{EL^h_t}\right) - M\left(\frac{AIDS^h_t}{EL^h_t}\right) + (1-\delta)k^h_t ,$$

where superscript *h* denotes the value of variables along the HIV epidemic path, AIDS the number of persons dying of AIDS, and *M* the constant per patient humanitarian expenditure on terminal medical care. Given assumptions about the functional form of F() and simulated paths for the growth of effective labour, effective labour per capita, and the number of AIDS patients, one can solve (3) for the maximum value of M, subject to the constraints that the initial (pre-epidemic) and asymptotic capital per effective labour ratios equal those found along the no-HIV path.

I consider two forms for the income generating function F(), a closed economy constrained by the concavity of the domestic production function and an open economy with perfect capital mobility and international transfers. Assuming domestic output is a Cobb-Douglas function of capital and effective labour, the income functions for the two economies are:

$$(5) \quad \text{Closed: } F() = GDP = Ak^{\alpha} \qquad \text{Open: } F() = GNP = rk + w + ct$$

where A is the constant level of domestic total factor productivity, α the share of capital in domestic GDP, k capital per effective worker, r and w the gross rental and wage per unit of effective labour (held fixed by capital mobility), ct net current transfers per effective worker, and where, in the case of the open economy, the domestic ownership of capital per effective worker does not necessarily equal the quantity of capital used in domestic production.

For the sub-Saharan economies, an open economy framework is both more natural and more informative. Many of these economies are large net borrowers, i.e. have gross savings rates well below their rates of capital formation, so that historical gross domestic investment considerably overstates the domestic ownership of capital. This is quite significant for the analysis,

as diminished population growth only yields resources to be spent on patient care while safeguarding future living standards insofar as there are positive amounts of capital being subdivided amongst future cohorts. At the same time, many sub-Saharan countries are recipients of large current transfers in the form of both international aid and cross border remittances, so that even GNP is a poor measure of current income. Finally, an open economy framework is more flexible, allowing the consideration of a hypothetical scenario where the sub-Saharan economies begin the simulation with zero net assets, i.e. own no capital at all. This scenario allows the reader to clearly see the role reduced fertility plays, not in lowering the rate at which assets are subdivided, but in changing dependency ratios, i.e. the rate at which assets can be accumulated while maintaining living standards. For these reasons, I focus on an open economy framework, reporting results for closed economy simulations in footnotes.

I estimate key macroeconomic parameters using data for 1970-2002 from the United Nations National Accounts Main Aggregates Database.[70] The national accounts estimates of the sub-Saharan economies are, to say the least, extraordinarily volatile.[71] Consequently, rather than

[70] Downloaded from http://unstats.un.org/unsd/snaama/dnllist.aspm on 9/5/2005, with available data on current transfers extracted, by request, from the UN database 9/23/2005 and supplemented, where missing observations exist, with data on current transfers downloaded from the IMF's international financial statistics database 9/12/2007. The reader might wonder why I use local currency values and do not make use of the ever-popular Penn World Tables (Heston, Summers and Aten 2002) purchasing power parity (PPP) national accounts. PPP based measures actually provide the wrong calculations, as they misstate the transformation possibilities between domestic output and international output and between domestic consumption and domestic capital. Thus, evaluating African output (particularly consumption) at PPP dramatically overstates the ability of the African economies to purchase international goods (such as retrovirals). Similarly, calculating domestic savings rates at PPP (with lower relative prices of capital goods) grossly overstates the economic ability to transform any temporary boon to effective participation into real capital accumulation. Consequently, the Penn World Tables data provide phenomenally stronger results. I report these in footnotes below for readers who prefer PPP measures, but I do not think they are the appropriate calculation.

[71] Thus, for the 27 countries in my simulation in the period 1970-2002, the average (maximum) absolute value of the ln annual changes in real GDP, real gross capital formation, the relative gross capital formation to GDP deflator, and the current price ratio of gross capital formation to GDP equal .049 (.92!), .157 (1.62!), .124 (1.65!), and .154 (1.56!), respectively. For the US data in the UN data base, the comparable figures are .033 (.07), .079 (.23), .013 (.04), and .050 (.13). I have been discussing the problems of African data with UN statistics officials, but it is not possible to resolve these issues within the bounds of this paper.

use any particular year to initialize the simulations, I make use of averages. As is well known, in a steady state with a constant savings rate and without technical change the domestic capital-output ratio is given by $s_I/(n+\delta-1)$, where s_I is the rate of capital formation in GDP, while the ln level of total factor productivity, for the Cobb-Douglas production function used above, equals $(1-\alpha)\ln[y] + \alpha \ln[(n+\delta-1)/s_I]$, where y is output per effective worker. Using the average 1970-1990 current price ratio of gross capital formation to GDP of the UN data as s_I, the 1970 to 1990 average growth of effective labour implied by my population simulations as n, and the 1990-2002 constant price average ln of GDP per capita in the UN data and the 1990-2002 average ln population per effective workers in my population simulations to calculate ln y, I estimate a value of A for each economy. Thus, my estimate for the productivity of each economy is average productivity in the 1990s, assuming that the capital-output ratio is at a steady state value determined by savings and labour force growth in the preceeding two decades.[72] I assume that all of the economies share common values of the parameters $\alpha = .4$ and $\delta = .06$, and convert local 1990 constant price values to constant 2005 US$ using the average real exchange rate for

[72] If the reader examines equations (2)-(5), she can easily see that, holding all of the demographically driven variables constant, increasing estimated steady state labour productivity by a factor τ (and ln A by $(1-\alpha)\tau$) uniformly scales up the equilibrium values of k, C and M by τ. In my calculations I have found that estimated labour and total factor productivity for the African economies appears to be substantially lower in the 1990s. Thus, using labour productivity numbers from the 1990s produces lower estimates of maximal expenditures on patient care than might arise if one used, say, the average of 1970-2002. I think, however, that this is the appropriate benchmark, as virtually all of the changes relative to the no-HIV path in the demographic variables begin in the 1990s (see Figure III later). In practice, my simulated output series ends up understating output prior to 1990, but this is of little relevance. As the reader can see from equations (3) and (4), if there are no demographic effects from HIV the path of k in the HIV scenarios matches the no-HIV path, even if A is moving rapidly. It is the productivity of the economy when it begins to carry the burden of patient care and experience changes in participation rates and rates of labour growth that determines maximal medical expenditures and, for this reason, I benchmark with the lower numbers of the 1990s.

1990-2002 in base 1990 prices, inflated to 2005 using the US GDP deflator.[73] I use the steady state values to set the initial 1970 domestic capital per effective worker [$k = y^*s_I/(n+\delta-1)$], as well as the fixed wage [$w = (1-\alpha)y$] and fixed rental [$r = \alpha\ (n+\delta-1)/s_I$] for the open economy. I use the 1970-1990 average savings rate out of GNP plus transfers (s_Y) and average ratio of current transfers to GNP (t_Y),[74] as well as the steady state assumption to calculate an initial ownership of capital per effective worker [$k = s_Y(1+t_Y)w/(n+\delta-1-s_Y(1+t_Y)r) = s_Y(1+t_Y)w/[(n+\delta-1)(1-\alpha s_Y(1+t_Y)/s_I)]$] for the open economy simulations.[75]

(b) Results

In the pages that follow I present simulation results for three scenarios concerning the fertility impact of HIV, evaluated within a variety of real and hypothetical initial asset positions. With regards to population dynamics, I allow the average derivative of ln probability of a birth with respect to HIV to equal its estimated value (f = -1.2), half of its estimated value (f = -.6) and

---

[73] Thus, I use the UN data to calculate the 1990-2002 mean of $\ln(P_i ER_{US,i}/P_{US})$, where P_i is the GDP deflator for country i (in base 1990 prices) and $ER_{US,i}$ is the nominal exchange rate from the currency of country i to US dollars, and then add the ln 1990-2005 US GDP deflator (from Council of Economic Advisers 2007). I should note that the UN data provide conversions to 1990 US dollars using the 1990 exchange rate, but, for my sample, on average the 1990 real exchange rate is greater than the 1990-2002 average. Using the averaging procedure described above further lowers my estimates of maximal retroviral expenditures. Were I to use the 1990 exchange rate benchmark provided by the UN, and the averate ln output per capita of 1970-2002 (as discussed in an earlier footnote), the aggregate maximal total expenditures on patient care reported below are increased by about 40%.

[74] The United Nations data on current transfers, even when supplemented with the IMF's records, are far from complete (i.e. contain many missing observations). I begin by calculating the average ratio of transfers to gross national income for any available years in 1970-1990 (the years 1970-2002 for the three countries with less than 11 observations in the 1970-1990 period). I then use this average ratio to calculate the year by year and average s_Y, savings rate out of gross national disposable income, using the conventional national accounts definition, i.e. $s_Y = (GNP - C - G + CT)/(GNP + CT) = (GNP(1+t_Y) - C - G)/(GNP(1+t_Y))$.

The issue of current transfers is important in calculating individual country savings rates, as without this adjustment many of the smaller African countries have negative gross savings rates. However, as much of these transfers are interregional, this mainly affects the estimated distribution of resources, rather than the overall level. I have run simulations where I use gross national savings excluding current transfers, i.e. (GNP - C - G)/GNP, and gotten fairly similar aggregate results, as reported in a footnote below.

[75] The reader can derive this last relation using equations (2) and (5) above. For $t_Y = 0$ and $s_Y = s_I$ it reduces to the expression for domestic capital per effective worker.

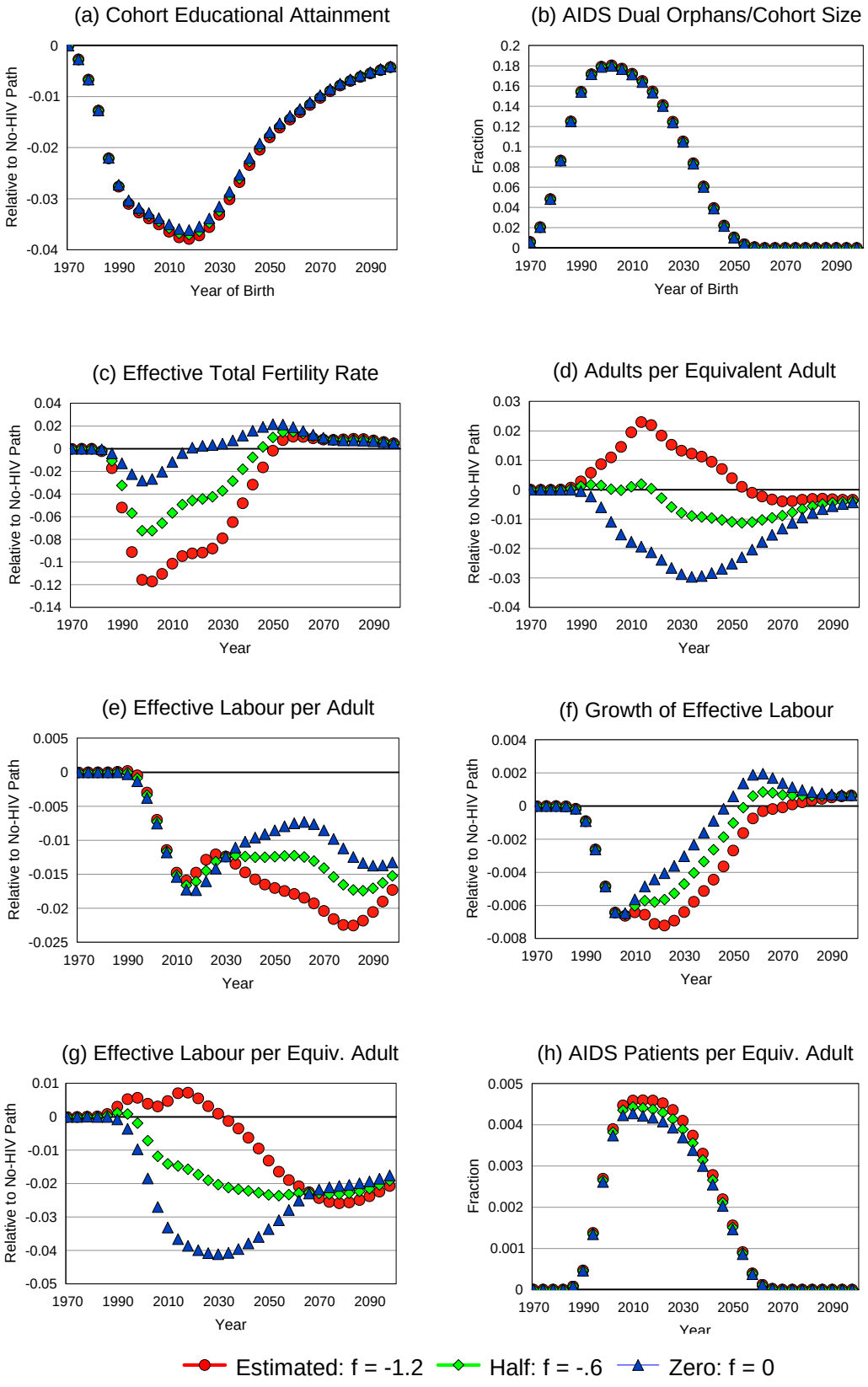
zero (f = 0). On the macroeconomic side, aside from my actual estimate of the initial asset position of the sub-Saharan economies, I also examine the special case where they have no initial assets whatsoever, as this highlights the role changing participation rates play in supporting AIDS patient care, as well as hypothetical scenarios in which they own their entire domestic capital stocks and receive no current transfers. The latter scenario has something of the flavor of "debt forgiveness", and also provides a link from the open economy simulations to the closed economy framework.

Figure III below presents cross country averages of the characteristics of the HIV population simulations, focusing on the variables that determine the ability of the sub-Saharan economies to support AIDS patient care while maintaining the living standards of the non-HIV path.[76] As shown in panel (a), the three scenarios show similar declines in cohort human capital, with, at its worst, a loss approaching 4% in average years of educational attainment. This loss of human capital is driven by a rapid rise in the number of AIDS dual orphans, panel (b), which reaches .18 of the average country cohort born early in the 21st century, attaining peaks as high as .54 and .50 for some cohorts born in Lesotho and Zimbabwe, respectively.[77] The impact of orphaning is extended for decades, as undereducated children have larger numbers of uneducated offspring, producing reductions in educational attainment that persist long after orphaning rates have

[76] In interpreting rates of change, the reader might find it useful to know that the observations (i.e. symbols) in the figure are presented at 4 year intervals. Each of the variables presented represents the unweighted average across the 27 simulation economies. Thus, for example, panel (a) depicts the ln average country cohort years of education under the different HIV paths minus the ln average country cohort years of education under the no-HIV path.

[77] The dual orphaned fraction is considerably higher than peak maternal infection rates (.30 in Lesotho and .27 in Zimbabwe), because many individuals who are born to HIV-free mothers are still orphaned before they reach 22, as their mothers contract HIV and die of AIDS. I should note that these are estimates of surviving dual orphans, i.e. excluding the 1/3 of children born to HIV positive women who die of AIDS related causes by their 12th birthday and never have the opportunity to participate economically.

Figure III: Simulation Characteristics



declined (compare panels (a) and (b)).[78]

Offsetting human capital losses is the reduction in fertility associated with the epidemic. As shown in panel (c) of Figure III, given the estimated impact of HIV on fertility (f=-1.2), the simulations indicate that the effective total fertility rate along the HIV path falls by as much as 12% relative to the no-HIV scenario.[79] The HIV epidemic strikes down working age adults, adversely affecting not only the accumulation of human capital by children, but also the overall ability of the afflicted society to support the consumption needs of its members. As indicated in panel (d), in the absence of any fertility effect the average number of adults per equivalent adult in sub-Saharan Africa would fall to about 3% below that experienced along the no-HIV path.[80] However, as shown in the same panel, the estimated communitywide reduction in fertility brought about by HIV can more than offset the mortality of adults, producing, paradoxically, a rise in adults per equivalent adult. Fertility effects cannot, nevertheless, obviate the concentration of mortality amongst younger, better educated,[81] prime working age adults. In all three scenarios, effective labour per adult falls about 1.5% with the initial onrush of epidemic induced

[78] Note also that with a .18 peak orphaning rate, the maximum loss of ln human capital (given a -18% impact of dual orphaning) should be about -.03 = ln(.18*exp(-.18) + .82). The peak loss of ln human capital, however, is -.04, reflecting the cumulative effects noted above.

[79] Effective fertility is the number of births, adjusted for childhood AIDS mortality, per woman aged 15 to 49. In this figure I remove childhood AIDS mortality, equal to 1/3 of the children born to HIV positive woman, as most of these die within a couple of years of birth and, consequently, do not impose a long term consumption burden on society (however, the consumption needs of these children, while alive, are incorporated into the simulations). For this reason, effective fertility falls even under the assumption of no communitywide fertility response to the epidemic (f=0). As should be abundantly clear by this point, I am not discussing welfare, but simply the ability of the society to sustain the no-HIV consumption levels while helping the infirm.

[80] Adults are counted as individuals aged 22 and above (who may participate economically) and equivalent adults equals the number of individuals aged 15 and above plus .5 times the number of children aged 0-14, reflecting the assumption that the consumption needs of children are half that of an adult.

[81] Although I assume that infection rates within a cohort are equal across educational groups, younger cohorts are on average better educated.

adult mortality (panel (e)).[82] Recovery from this level is hampered by the diminished educational attainment of subsequent cohorts (panel (a) again) and, in the case of the scenarios with reduced fertility, the smaller relative numbers of new youth cohorts.[83]

The elements outlined above combine to produce two key measures that determine the ability of the sample economies to support AIDS patient care, the growth of the effective labour force and the effective labour force per capita. As shown in panel (f) of Figure III, whether under the estimated, half or zero fertility effect scenarios, the growth of effective labour in the presence of the epidemic is reduced relative to the no-HIV path. With stronger fertility effects, however, this reduction is perpetuated and strengthened in later years, long after the initial mass of adult deaths brought by the epidemic, as reduced fertility lowers the size of subsequent cohorts. While slower growth of the effective labour force tends to increase capital labour ratios and produce higher labour productivities, the level of effective labour per equivalent adult, i.e. the overall effective participation rate, determines the dependency burden falling on each worker. As shown in panel (g), with the estimated HIV fertility effect the reduction in the size of youth cohorts dominates the loss of educated adult workers, at least early in the epidemic, increasing effective participation relative to the no-HIV path. With weaker fertility effects, however, adult mortality is more dominant and effective participation falls from early on, reaching a trough 4% below the no-HIV path . In all scenarios, the loss of human capital brought about by orphaning

[82] The concentration of mortality in better educated cohorts is confirmed by the fact that crude participation (number of working adults per adult) falls by no more than .4% during this period.

[83] Although new cohorts are less well educated than they might have been, they are nevertheless better educated than their parents, as, with better educated children producing better educated children, average years of education rise gradually (attaining an average steady state 15% higher than the educational attainment of the no-HIV year 2000 birth cohort). Crude adult participation in all three scenarios also actually rises above the no-HIV path in the middle of the 21st century, as new working age cohorts enter adulthood, but the relative movement with a zero fertility effect (a maximal increase in relative average crude participation of 1.5%) is much greater than that found with the estimated fertility effect (a maximal increase of .8%), which, along with the greater educational attainment of younger cohorts, produces the higher relative effective labour per adult along that path shown in panel (e).

remains long after all demographic effects have disappeared, producing a long term decline in effective labour per equivalent adult. Figure III concludes by estimating the additional dependency burden that would be created by an AIDS patient care programme, as average AIDS patients per equivalent adult approach .005, reaching sustained highs in excess of .015 in Lesotho and .011 in Namibia, South Africa and Zimbabwe.

Table XIV below reports the maximal expenditures on AIDS patient care consistent with maintaining the no-HIV path of living standards under various scenarios. To begin with an extreme hypothetical possibility, if the sub-Saharan economies had historically followed a path of zero savings and accumulated no assets (panel #1), they would still, in the face of the epidemic, be able to expend $660 per AIDS victim in the year of their death in patient care, while maintaining the no-HIV consumption levels. These results highlight the fact that the early rises in participation brought about by the decline in fertility, alone, confer considerable consumption benefits.[84] Allowing the economies their estimated ownership of capital (panel #2) introduces the capital dilution effects associated with reduced rates of population growth, increasing the implicit resources available for patient care in economies endowed with positive initial net assets and reducing them in the one country (Mozambique) with net initial indebtedness.[85] The overall effect is overwhelming positive, with the sub-Saharan African countries, in aggregate, being able to afford $5123 per patient, assuming the estimated effect of -1.2 of HIV on fertility. To put these figures in perspective, generic triple dose retroviral therapy is now available for less than $292 per year, while Brazil's large scale retroviral programme, complete with auxiliary CD4,

[84] The only economies that cannot sustain AIDS patient care under such circumstances are South Africa and Gabon, which have two of the lowest rentals and highest consumption levels in the sample. As the temporary boon in participation brought about by reduced fertility has to be translated into the targeted future no-HIV consumption level (offsetting human capital losses) through the rental, this makes them unable, in the absence of any initial capital assets, to support patient care.

[85] Where reduced population increases the debt servicing burden of future cohorts. Despite substantial current transfers, the high reported levels of consumption in Mozambique imply a negative net wealth.

Table XIV: Maximal Expenditures on AIDS Patient Care
Consistent with Maintaining No-HIV Consumption Levels (2005 US$)

				(1) Zero Assets	(2) Estimated Assets			(3) Full Ownership of Domestic Capital		
	s_Y	s_I	t_Y	f = -1.2	f = -1.2	f = -.6	f = 0	f = -1.2	f = -.6	f = 0
Benin	.04	.16	.10	1636	2175	573	-1043	3938	1983	12
Burkina Faso	.12	.23	.14	864	1699	520	-705	2462	1154	-208
Cameroon	.03	.20	-.00	1728	2151	-770	-3798	7091	3150	-936
Cen. Afr. Rep.	.09	.13	.09	1195	2450	1147	-233	2749	1430	30
Chad	.03	.14	.09	1183	1418	429	-579	2466	1247	1
Congo	.22	.30	-.01	1125	5574	1115	-3577	8540	3484	-1838
Cote D'Ivoire	.13	.19	-.03	2473	5253	1804	-1798	7858	3855	-324
Ethiopia	.13	.12	.04	450	1246	672	88	1046	534	13
Gabon	.43	.39	-.03	-7551	55595	13622	-31017	48637	8969	-33189
Ghana	.09	.09	.02	1546	3438	1775	98	3363	1740	102
Guinea	.16	.16	.03	1406	4710	2390	41	4314	2108	-126
Kenya	.18	.24	.01	707	2364	593	-1267	3207	1269	-766
Lesotho	.10	.28	.11	199	1420	-604	-3024	4297	1838	-1124
Malawi	.12	.21	.06	425	1049	234	-649	1655	736	-261
Mali	.17	.18	.09	881	2613	1239	-147	2495	1179	-150
Mozambique	-.20	.10	.09	863	86	-322	-743	1749	909	38
Namibia	.15	.23	.20	2785	12566	4425	-4601	15871	7616	-1601
Niger	.11	.19	.03	796	1605	477	-657	2383	1092	-206
Nigeria	.20	.16	-.01	1305	5496	3120	683	3973	1941	-143
Rwanda	.06	.16	.05	894	1351	369	-659	2407	1206	-51
Senegal	.08	.11	.04	2449	4651	2094	-469	5503	2763	15
South Africa	.21	.25	.00	-1298	16960	6485	-5247	22825	11510	-1169
Tanzania	.16	.24	.06	377	1285	293	-755	1862	768	-391
Togo	.21	.28	.07	985	3300	1211	-933	4069	1834	-462
Uganda	.04	.09	.04	1458	2005	876	-342	2874	1508	33
Zambia	.04	.20	-.03	834	1132	-296	-1898	3430	1551	-561
Zimbabwe	.17	.19	-.00	1329	4522	2009	-927	5144	2523	-540
Country Avg.	.12	.19	.05	780	5486	1684	-2376	6526	2589	-1622
Weighted Avg.				660	5123	1991	-1439	6437	3112	-541

Notes: $s_Y = (GNP(1+t_Y)-C-G)/(GNP(1+t_Y))$, s_I = GCF/GDP, and t_Y = CT/GNP, all averages for 1970-1990 or, exceptionally, 1970-2002 (t_Y for Guinea, Mozambique & Uganda - see footnote earlier above), with CT equal to current transfers and GCF gross capital formation. f = the average derivative of the ln probability of a birth with respect to the HIV infection rate among women aged 15-49. Country avg. = simple average across the 27 countries; Weighted Avg. = average expenditure per patient across all countries, i.e. a population weighted average.

viral load and drug resistance testing, is currently delivered at a cost of less than $1200 per patient year.[86] With an average of $5123 in available resources per patient, most of the countries in Table IV could afford to provide basic retroviral drugs, and even a full patient programme, to each AIDS victim for years.[87]

Table XIV also reports the sensitivity of the results to changes in the magnitude of the HIV fertility effect and the possibility of a better initial asset position. As shown in panel #2, with half the estimated impact on fertility (f=-.6) about half of the economies are still able to support the basic cost of retroviral drugs for years, with an average expenditure across the entire sample of $1991 per patient. With no fertility effect at all, however, the loss of prime working age adults and the human capital of future cohorts is crushing, requiring large transfers (i.e. negative expenditures) per patient to sustain the no-HIV consumption levels. As seen in panel #3, if one allows each sub-Saharan country the ownership of its own capital stock, the ability to support patient care (while sustaining an increased no-HIV consumption level for all generations)[88] is increased somewhat and, more importantly, spread more evenly across the economies, with virtually every economy able to support substantial patient care even with half the estimated fertility effect. In general, I find that different manipulations of African data to estimate different

[86] McNeil (2004) reports figures of US$292 and US$244 for Indian generic producers and notes that in October 2003 the Clinton foundation negotiated a price of $140 per patient year for large orders paid in cash. On Brazil's programme, see http://www.avert.org/aidstarget.htm (10/13/2005) and Marins et al 2003.

[87] UNAIDS (2002) estimates that retroviral therapy in sub-Saharan Africa will extend AIDS patients lives by an additional three years. Median survival times in Brazil rose from 18 months for AIDS patients diagnosed in 1995 to 58 months for AIDS patients diagnosed in 1996, when the Brazilian retroviral programme was instituted (Marins et al 2003), so the UNAIDS assumption seems to be about right. While there are substantial side effects, patients on retroviral drugs are not invalided, and (as most are adults of prime working age) would be able to earn and produce their average consumption needs.

[88] In panel #3 each economy is given the initial ownership of capital associated with s_I and the HIV simulations are constrained to give the (generally higher) no-HIV consumption path that would be associated with that level of savings and capital ownwership. Thus, future generations are given the full consumption benefits of "debt forgiveness" that would be experienced along the no-HIV path. These simulations also assume that current transfers are zero throughout.

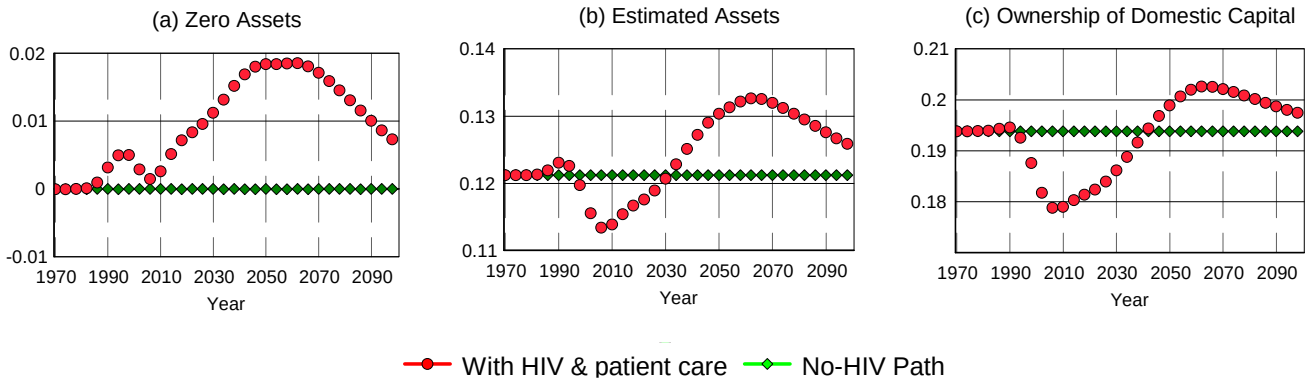
initial asset positions change the distribution of resources, but do not have a substantial influence on the average ability to support patient care.[89] In this regard, it is really the magnitude of the fertility effect that is of primary importance.

While the preceeding calculations are done in an open economy framework, the reader should not conclude that the simulated patient care expenditures require extensive international borrowing. Figure IV below depicts the average 27 country savings rate associated with the expenditures outlined in Table XIV for the full fertility effect (f=-1.2) under various initial asset assumptions. With zero initial assets, panel (a), savings rates are above the no-HIV path throughout, as the economy has to bequeath physical capital to maintain the living standards of future cohorts whose human capital is diminished by the cumulative effects of orphaning. Both this accumulation of assets and the support of patient care come from the increase in effective participation brought about by the reduced size of dependent youth cohorts. Jumping to panel (c), we see that with full of ownership of domestic capital the simulated economies save less initially, allocating more resources to patient care and relying upon reduced rates of population growth to raise the capital labour ratio above the no-HIV path. The simulation with the estimated level of assets, panel (b), lies between these two extremes. In all cases, the domestic

[89] Thus, if I ignore current transfers and simply calculate national savings as (GNP-C-G)/GNP, average possible expenditures per patient for f = -1.2 are $4921, which is very close to the number reported in panel #2 of Table XIV. Again, the main thing that changes is the distribution across countries.

Reporting additional sensitivity tests, if equivalent adults are calculated evaluating youths aged 0-21 as .7 of an adult, maximal expenditures per patient for f = -1.2 in panel #2 rise to $6661. If I shorten and lengthen the time horizon of the epidemic, assuming that infection rates decline sinusoidally to zero in 20 or 95 years, maximal expenditures per patient for f = -1.2 in panel #2 rise and fall to $5378 and $4377, respectively. Turning to the closed economy framework (i.e. giving each country ownership of its capital stock, and allowing no international borrowing/lending or current transfers), the average expenditures per patient for f = -1.2 are $6004. These figures are somewhat lower than those reported in panel #3 (but close to those in panel #2), reflecting the difficulty of substituting intertemporally in a closed economy setting. If I use the Penn World Table data, in a closed economy framework with full ownership of domestic capital and f = -1.2, I get average per expenditures of 2005 US $19120 per patient. As noted earlier, however, I do not think PPP data provide the correct calculation.

Figure IV: Average Savings Rates



ownership of capital per effective worker rarely falls substantially below that of the no-HIV path[90] as, to offset the human capital losses of the epidemic, the sub-Saharan African countries need to accumulate assets.[91]

[90] Thus, for the estimated and full ownership of assets scenarios, capital per effective labour is only more than 1% below the no-HIV path for 79 and 56 country x year observations, respectively, and the average reduction, in those cases, is only 3.2% and 2.2%.

[91] At this point, the reader might naturally wonder what is actually happening to sub-Saharan savings rates. I use three measures of savings for my 27 country simulation sample: S1 = (GNP-C-G)/GNP, which ignores current transfers; S2 = (GNP(1+trmean)-C-G)/(GNP*(1+trmean)) which uses the average level of current transfers to GNP in 1970-2002 to calculate year by year savings; and S3 = (GNP(1+tr)-C-G)/(GNP(1+tr)) which uses the actual ratio of transfers to GNP, limiting the sample to the 723 country x year observations with current transfers data (as opposed to 871 for the other measures). The coefficients in the regression of these measures on the average HIV infection rate with country dummies and with country dummies and time trends, are .056 (s.e. .066) and .413 (.113) for S1, .042 (.063) and .393 (.108) for S2, and .008 (.069) and .061 (.122) for S3. However, as emphasized above, it is the rise in the ratio of adults per equivalent adult (reduced dependency burdens) that is generating resources for the countries afflicted by the epidemic. When I regress the savings measures on the simulated ratio of adults per equivalent adult with country dummies and with country dummies and time trends, I get coefficients of .886 (.332) and 3.18 (.794) for S1, .781 (.317) and 3.05 (.759) for S2, and .518 (.342) and 1.20 (.860) for S3. Thus, although sub-Saharan national accounts data should be approached with extreme caution, it would seem that savings rates may be rising during the epidemic, particularly with reductions in the dependency ratio.

VI. Conclusion

Figure V below presents my estimates of the total annual expenditure on the support and treatment of adult AIDS patients in 27 sub-Saharan African countries consistent with granting each and every cohort the material standard of living it would otherwise have enjoyed absent the HIV epidemic,[92] while Figure VI depicts the number of individuals who will die each year absent any such treatment. As shown, if HIV has no effect on fertility (f=0), then the sub-Saharan countries are not able to support any patient care and in fact need additional transfers, cumulatively totalling $194 billion, merely to maintain living standards. With even half of the baseline impact of HIV on ln fertility (f=-.6), however, extensive expenditures on patient care become possible, while with the baseline influence (f=-1.2), supported both by DHS and Census data, potential annual expenditures in the year in which each adult AIDS victim would otherwise have died rise to $13.5 billion a year and cumulatively total $650 billion. These expenditures are after, i.e. exclusive of, allowances for the accumulation of extra physical capital to offset the human capital losses of AIDS orphans. Furthermore, they represent pure humanitarian expenditure, as I assume no economic benefit, as in participation or children's outcomes, from adult patient care.

In terms of per capita material consumption, fertility declines impart a boon, as lower dependency ratios increase effective participation and reduced labour force growth lessens the dilution of capital. Although couched in terms of a patient care programme, the expenditures laid out in Figure V are a measure of the material boon granted by the reduced demand for children brought about by the HIV epidemic. In some respects the application of this windfall, whether to current consumption or capital bequests to future generations, is in the micro control of the individuals living through the epidemic. The epidemic itself, however, brings forth a demand for public goods and services, such as information on viral transmission and treatment,

[92] These represent the total annual expenditures on patient treatment using my baseline estimates of national assets, i.e the scenarios presented in panel #2 of Table XIV. As noted earlier, f represents the average country derivative of the ln probability of giving birth with respect to the infection rate among women aged 15-49.

Figure V: Annual Expenditures on Patient Care

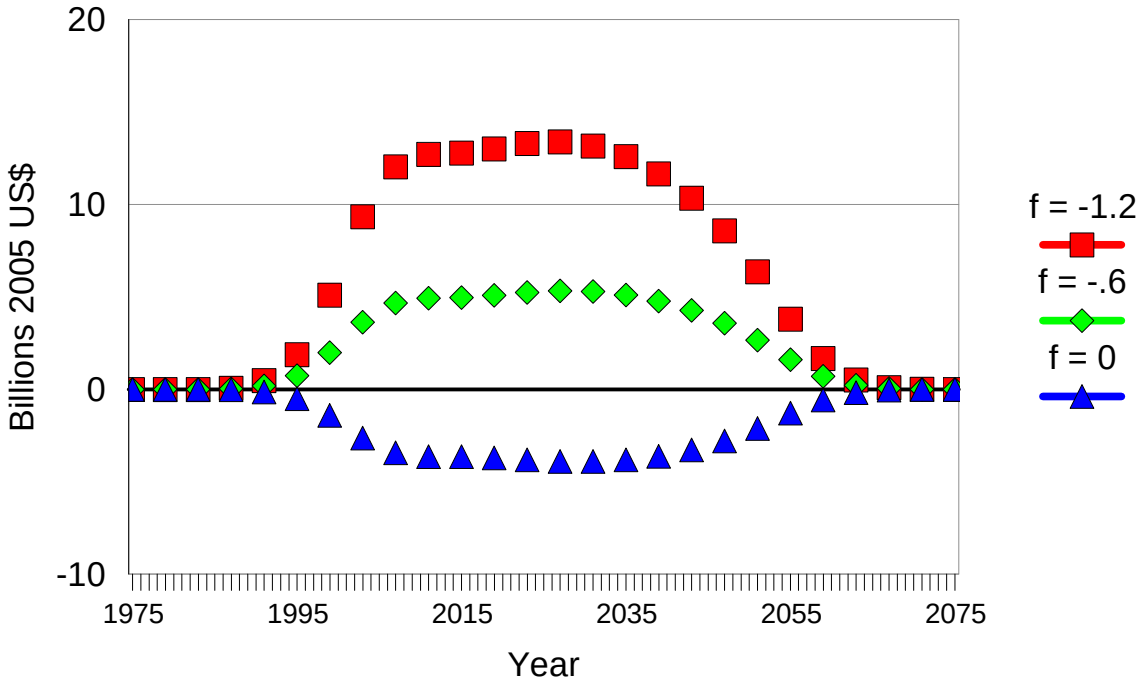
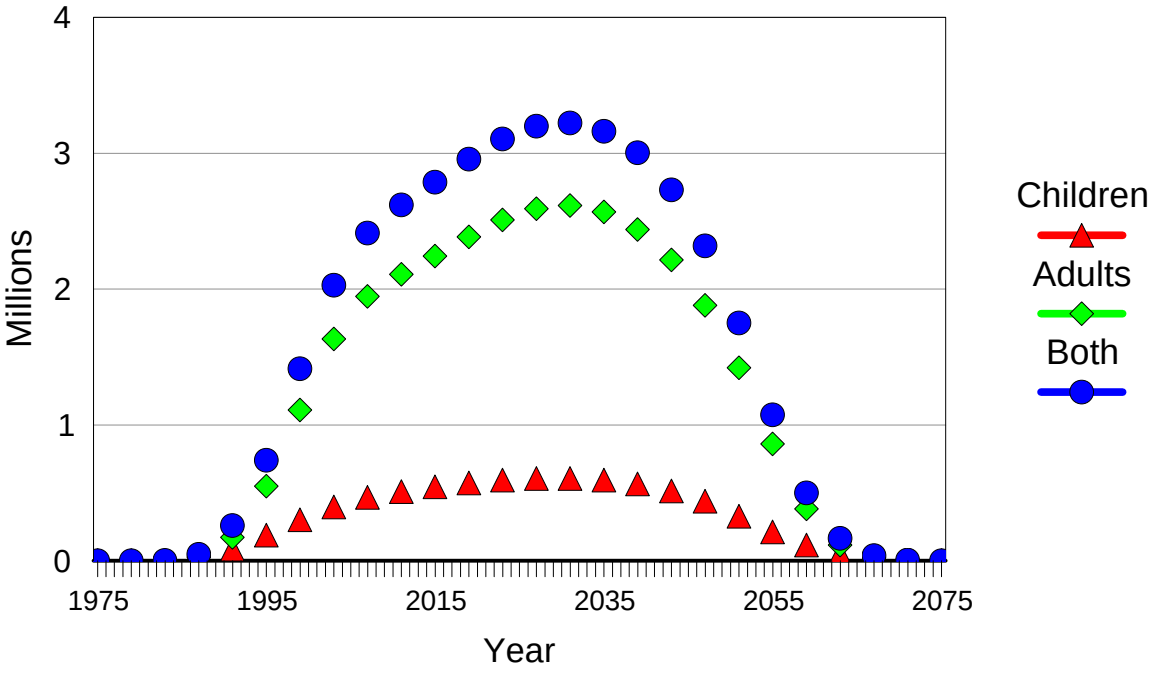


Figure VI: Annual AIDS Deaths



accreditation of treatment systems, low cost mass procurement of drugs, and insurance against family and communal abandonment as the disease progresses, that are not easily provided by the private sector.[93] In this regard, the laxity of many African governments might charitably be excused as driven by fears that costly widespread treatment of a seemingly incurable disease will only worsen the outcomes of the orphaned and already disadvantaged members of future generations. In this paper, I argue that such fears are unfounded. The behavioral response to the HIV epidemic creates the material resources to fight it.

[93] As noted in the preceeding section, African savings data suggest that savings rates may be rising with the decline in dependency ratios brought by the epidemic. However, even if this is not the case and all of the consumption boon is being privately consumed, it is not clear that this represents the most preferred outcome. In the midst of a plague, individuals might prefer well funded public programmes to increased purchases of private goods.

VII. Bibliography

Blakeslee, Dennis. "Mother-to-Child Transmission of HIV." HIV/AIDS Resource Center, The Journal of the American Medical Association. http://www.ama-assn.org/special/hiv/newsline/briefing/mother.htm (11/3/2002).

Bobat, Raziya, Hoosen Coovadia, Dhayendre Moodley, et al. "Mortality in a cohort of children born to HIV-1 infected women from Durban, South Africa." South African Medical Journal 89, No. 6 (1999): 646-648.

Carpenter Lucy M., Anatoli Kamali, Anthony Ruberantwari, et al. "Rates of HIV-1 transmission within marriage in rural Uganda in relation to the HIV sero-status of the partners." AIDS 13 (1999): 1083-1089.

Carpenter, Lucy M., Jessica S. Nakiyingi, Anthony Ruberantwari, et al. "Estimates of the impact of HIV infection on fertility in a rural Ugandan population cohort." Health Transition Review 7, Supplement 2 (1997): 113-126.

Case, Anne, Christina Paxson, and Joseph Ableidinger. "Orphans in Africa: Parental Death, Poverty and School Enrollment." Manuscript. Princeton University, March 2004.

Changalucha, John, Heiner Grosskurth, Wambura Mwita, et al. "Comparison of HIV prevalences in community-based and antenatal clinic surveys in rural Mwanza, Tanzania." AIDS 16 (2002): 661-665.

Chinapah, Vinayagum, E.L. Mostafa H'ddigui, Anil Kanjee, et al. With Africa for Africa: Towards...Quality Education for All. Pretoria, South Africa: Human Sciences Research Council, 2000.

Council of Economic Advisers. Economic Report of the President - February 2007. Washington, D.C.: United States Government Printing Office, 2007.

Crampin, Amelia C., Judith R. Glynn, Bagrey M.M. Ngwira, et al. "Trends and measurement of HIV prevalence in northern Malawi." AIDS 17 (2003): 1817-1825.

Cutler, David M., James M. Poterba, Louise M. Sheiner, and Lawrence H. Summers. "An Aging Society: Opportunity or Challenge?" Brookings Papers on Economic Activity 1990, No. 1 (1990): 1-73.

Deaton, Angus. The Analysis of Household Surveys: A Microeconometric Approach to Development Policy. Baltimore: Johns Hopkins, 1997.

Desgrées du Loû, Annabel, Philippe Msellati, Angèle Yao, et al. "Impaired fertility in HIV-1-infected pregnant women: a clinic-based survey in Abidjan, Côte d'Ivoire, 1997." AIDS 13 (1999): 517-521.

De Vincenzi, Isabelle, Corinee Jadand, Elisabeth Couturier, et al. "Pregnancy and contraception in a French cohort of HIV-infected women." AIDS 11 (1997): 333-338.

D'Ubaldo, Carla, Patrizio Pezzotti, Giovanni Rezza, et al. "Association between HIV-1 infection and miscarriage: a retrospectie study." AIDS 12 (1998): 1087-1093.

Fylkesnes, Knut, Rosemary M. Musonda, Moses Sichone, et al. "Declining HIV prevalence and risk behaviors in Zambia: evidence from surveillance and population-based surveys." AIDS 15 (2001): 906-916.

Fylkesnes, Knut, Zacchaeus Ndhlovu, Kelvin Kasumba, et al. "Studying dynamics of the HIV epidemic: population-based data compared with sentinal surveillance in Zambia." AIDS 12 (1998): 1227-1234.

Glynn, Judith R., Anne Buvé, Michel Caraël, et al. "Decreased fertility among HIV-1-infected women attending antenatal clinics in three African cities." Journal of Acquired Immune Deficiency Syndrome 25 (2000): 345-352.

Glynn, Judith R., Anne Buvé, Michel Caraël, et al. "Factors influencing the difference in HIV prevalence between antenatal clinic and general population in sub-Saharan Africa." AIDS 15 (2001): 1717-1725.

Gray, Ronald H., Maria J. Wawer, David Serwadda, et al. "Population-based study of fertility of women with HIV-1 infection in Uganda." Lancet 351 (1998): 98-103.

Gregson, Simon, Nicola Terceira, Memory Kakowa, et al. "Study of bias in antenatal clinic HIV-1 surveillance data in a high contraceptive prevalence population in sub-Saharan Africa." AIDS 16 (2002): 643-652.

Heston, Alan, Robert Summers and Bettina Aten. Penn World Table Version 6.1. Center for International Comparisons at the University of Pennsylvania (CICUP), October 2002. http://pwt.econ.upenn.edu.

Hira, S.K., B.M. Nkowame, J. Kamanga, et al. "Epidemiology of Human Immunodeficiency Virus in Families in Lusaka, Zambia." Journal of Acquired Immune Deficiency Syndromes 3 (1990): 83-86.

Human Sciences Research Council. Nelson Mandela/HSRC Study of HIV/AIDS Household Survey 2002. Cape Town, South Africa: Human Sciences Research Council Publishers, 2002.

Hunter, Susan-Catherine, Raphael Isingo, J. Ties Boerma, et al. "The association between HIV and Fertility in a cohort study in rural Tanzania." Journal of Biosocial Science 35 (2003): 189-199.

Kigadye, Rose-Mary, Arnoud Klokke, Angus Nicoll, et al. "Sentinal surveillance for HIV-1 among pregnant women in a developing country: 3 years' experience and comparison with a population serosurvey." AIDS 7 (1993): 849-855.

Kilian, Albert H.D., Simon Gregson, Bannet Ndyanabangi, et al. "Reductions in risk behavior provide the most consistent explanation for declining HIV-1 prevalence in Uganda." AIDS 13 (1999): 391-398.

Killingsworth, Mark R. and Heckman, James J. "Female Labor Supply: A Survey." In Handbook of Labor Economics, Vol. I, ed. Orley Ashenfelter and Richard Layard, 103-204. Amsterdam: North-Holland, 1986.

Krieger, John N., Robert W. Coombs, Ann C. Collier, et al. "Fertility Parameters in Men Infected with Human Immnodeficiency Virus." The Journal of Infectious Diseases 164 (1991): 464-469.

Marins, Jose Ricardo P., Leda F. Jamal, Sanny Y. Chen, et al. "Dramatic improvement in survival among adult Brazilian AIDS patients." AIDS 17 (2003): 1675-1682.

McNeil, Donald G. Jr. "Plan to Battle AIDS Worldwide is Falling Short." New York Times. 28 March 2004.

Miotti, Paolo G., Gina Dallabetta, Eric Ndovi, et al. "HIV-1 and pregnant women: associated factors, prevalence, estimate of incidence and role in fetal wastage in central Africa." AIDS 4 (1990): 733-736.

Morgan, Dilys, S.S. Malamba, B. Mayanja, et al. "Survival by AIDS defining condition in rural Uganda." Sexually Transmitted Infections 76, No. 3 (June 2000): 193-197.

Republic of South Africa, Department of Health. "Summary Report: National HIV and Syphilis Antental Sero-Prevalence Survey in South Africa, 2002." Pretoria, South Africa: Department of Health, 2003.

Ross, Amanda, Dilys Morgan, Rosemary Lubega, et al. "Reduced fertility associated with HIV: the contribution of pre-existing subfertility." AIDS 13 (1999): 2133-2141.

Selwyn, Peter A., Ellie E. Schoenbaum, Katherine Davenny, et al. "Prospective Study of Human Immunodeficiency Virus Infection and Pregnancy Outcomes in Intravenous Drug Users." JAMA 261 (1989): 1289-1294.

Serwadda, David, Ronald H. Gray, Maria J. Wawer, et al. "The social dynamics of HIV transmission as reflected through discordant couples in rural Uganda." AIDS 9 (1995): 745-750.

Sewankambo, Nelson K., Maria J. Wawer, Ronald H. Gray, et al. "Demographic impact of HIV infection in rural Rakai District, Uganda: results of a population-based cohort study." AIDS 8 (1994): 1707-1713.

Stephenson, Judith M., Anja Griffioen, and the Study Group for the Medical Research Council Collaborative Study of Women with HIV. "The effect of HIV diagnosis on reproductive experience." AIDS 10 (1996): 1683-1687.

Strauss, J.P. Results of the Monitoring Learning Achievement Project. Bloemfontein, South Africa: Research Institute for Education and Planning, 1999.

UNAIDS Reference Group on Estimates, Modelling and Projections. "Improved methods and assumptions for estimation of the HIV/AIDS epidemic and its impact." AIDS 16 (2002): W1-W14.

UNAIDS Reference Group on Estimates, Models and Projections. "Estimating and Projecting National HIV/AIDS Epidemics." Manuscript. UNAIDS: January 2003.

Unicef. "Mother-to-Child Transmission of HIV." New York: United Nations Children's Fund, 2002.

Widy-Wirski, Roslaw, Seth Berkley, Robert Downing, et al. "Evaluation of the WHO Clinical Case Definition for AIDS in Uganda." JAMA 260 (1988): 3286-3289.

Young, Alwyn. "The Gift of the Dying: The Tragedy of AIDS and the Welfare of Future African Generations." Quarterly Journal of Economics CXX (2005): 423-466.

VIII. Appendix: Samples and Variables

This appendix provides details on the samples and variables used in the paper. Table A1 below lists the DHS/WFS surveys available to me in 29 of the 39 sub-Saharan countries with generalized HIV epidemics,[94] noting the year of the survey and the type of questionnaires available.[95] The variables used in the analysis are listed in Table A2.

In regards to Table XIII's analysis of children's education as a function of mother's educational attainment, for individuals 14 and under (generally) the DHS files include the mother's line number (hv112) if she is in the household. Linking mother's line numbers to the household data, I clean these records by eliminating observations where the "mother" (as described in the household records) is actually male, where the age difference between the "mother" and child is less than 12 or more than 60 years, or where the relationship between "mother" and child is inconsistent (e.g. both are identified as the son/daughter of the household head). I then augment these data with mothers identified on the basis of the household relationship records (hv101). Specifically, for respondents who report that they are the household head, I identify their mother as the female individual who reports that she is the mother of the household head, while for individuals who report that they are the children of the household head, I identify their mother as either the household head (if female) or the female individual who reports that she is the spouse of the household head. I eliminate cases where ambiguity is created by the existence of multiple individuals who claim to be the household head or the female spouse or parent of the household head or where the age difference between "mother" and child is less than 12 or more than 60 years. Use of the household relationship data allows me to identify the mothers of almost one hundred thousand additional individuals over the age of 14.

I should note that in the case of the multinomial analysis of contraception and the logit analysis of menstruation the large number of country specific control categories generates sparse problem groups where all observations have the same outcome. This happens for a small handful of countries where I have only one or at most two surveys (i.e. countries which do not contribute to the identification of the impact of HIV when time trends are included), and I handle this by simply dropping those countries from the sample.

Finally, I should note that as some surveys take place across two years, strictly speaking my discussion in the text that with country dummies (or country dummies and time trends) a country only contributes to the estimates of the influence of HIV when it has two (three) or more surveys, is not correct. However, as HIV infection rates change little across adjacent years, in practice these countries do not contribute much to the identification of the influence of HIV (as I confirmed by rerunning some of the analysis with each survey coded as taking place in a single, average, year).

---

[94] As noted in the paper, data are also available for Sudan and the offshore islands of Comoros and Madagascar, which are not classified by the Census Bureau as having generalized epidemics, but as infection rates in these areas are negligible and none of the identification comes from cross-country comparisons, I have not included them in my sample.

[95] For the DHS, I have only made use of the DHS recoded (i.e. harmonized) files, as the non-recoded files generally contain information on only a few variables. However, to expand the sample, I have recoded the WFS surveys (which provide all the information necessary for the analysis of fertility preferences and contraceptive behavior in section III) and the household files of the Malawi 1996 and Tanzania 1994 Knowledge, Attitude and Practices surveys (which provide all of the information necessary for the analysis of education in section IV).

Table A1: DHS and Associated Surveys Used in the Paper

	Year	Type	Year	Type	Year	Type	Year	Type	Year	Type	Year	Type
Benin	2001	WHM	1996	WHM	1981/82	W*						
Burkina Faso	2003	WHMCh	1998/99	WHM	1992/93	WHM						
Burundi	1987	W										
Cameroon	2004	WHMh	1998	WHM	1991	WHM	1978	W*				
Central Afr. Rep.	1994/95	WHM										
Chad	2004	WHM	1996/97	WHM								
Congo	2005	WHM										
Cote D'Ivoire	2005	WHMh	1998/99	WHM	1994	WHM	1980/81	W*				
Ethiopia	2005	WHMCh	2000	WHM								
Gabon	2000	WHM										
Ghana	2003	WHMCh	1998	WHM	1993	WHM	1988	W	1979/80	W*		
Guinea	2005	WHMCh	1999	WHM								
Kenya	2003	WHMCh	1998	WHMC	1993	WHM	1989	W	1977/78	W*		
Lesotho	2004	WHMh	1977	W*								
Liberia	1986	W										
Malawi	2004	WHMCh	2000	WHMC	1996	H#	1992	WHM				
Mali	2001	WHMCh	1995/96	WHM	1987	W						
Mozambique	2003	WHMC	1997	WHM								
Namibia	2000	WHM	1992	WH								
Niger	1998	WHM	1992	WHM								
Nigeria	2003	WHM	1999	WHM	1990	WH						
Rwanda	2005	WHMCh	2000	WHMC	1992	WHM	1983	W*				
Senegal	2005	WHMCh	1997	WHM	1992/93	WHM	1986	W	1978	W*		
South Africa	1998	WH										
Tanzania	2004	WHMC	2003	WHMh	1999	WHM	1996	WHM	1994	H#	1992	WHM
Togo	1998	WHM	1988	W								
Uganda	2000/01	WHMC	1995	WHM	1988	W						
Zambia	2001/02	WHMh	1996	WHM	1992	WH						
Zimbabwe	1999	WHMC	1994	WHMC	1988	W						

Notes: * = WFS survey, # = KAP survey, W = women (used in sections II, III and V), H = household (used in sections IV and V), M = men (used in an appendix, available from the author), C = surveys with 5 year retrospective calendar information on pregnancy outcomes (used in section III), h = HIV data (used in sections III and V). The HIV data for Mali and Zambia cannot be linked to files with individual characteristic data, and hence are not used in much of the analysis presented in tables and footnotes in sections III and V.

Table A2: DHS Variables Used in the Paper
Independent Variables (or variables used in the construction of independent variables): Tables IV & VIII - past year's fertility (v209), retrospective fertility (b3); Table V - ideal number of children (v613), desire for more children (v602); Table VI - contraceptive use (v312); Table VII - menstruation (v216), successful pregnancies (vcal_1); Table IX - school attendance (hv110); Table X - attainment (hv106); Tables XI & XII - attainment (hv106 & hv108); Table XIII - b3, b4, b5, b6, mm1, mm2, mm3, mm6, mm7, mm15, v008, hv106, hv108, v714, mv714, v736, mv736.
Women's Conditioning Variables: age (v012), educational attainment (v106), year of survey (v007), urban/rural (v102), born children (v201), living children (v201-v206-v207), marital status (v502).
Household Conditioning Variables: age (hv105), head's attainment (hv106), year of survey (hv007), electricity (hv206), radio (hv207), bicycle (hv210), urban/rural (hv025), mother dead (hv111) and father dead (hv113).
Sample Selection Issues: Table V - ideal number of children - drop non-numeric responses such as "God's decision" or "as many as possible"; Table V - want no more children - "declared infecund" and "never had sex" dropped, "no more" and "sterilized" coded as want no more children, "have another" and "undecided" as converse, keep only individuals who are currently married (v502=1); Table VI - drop all individuals who say they are currently pregnant (v213=1) or declared infecund (v602=5), keep only individuals who are currently married (v502=1), drop Burundi, Central African Republic, Liberia, Niger, South Africa and Togo for multinomial analysis (sparse data in categories, see above); Table VII - menstruation - drop all individuals who say they are currently pregnant (v213=1) or breastfeeding (v404=1), drop Burundi, Central African Republic and Niger (sparse data), drop Senegal 1986 (recode file is a coding error, data are NA); Tables IX-XII - drop individuals who are listed as the household head (as conditioning on education of household head); Table XIII - wage estimates - keep if v714/mv714=2 (work for someone else) and v736 (incomes) records weekly, monthly or yearly pay (multiplied by 50, 12 & 1, respectively).
Notes: "v" prefix denotes the women's file, "hv" prefix the household file, and "mv" prefix the men's file. For all variables, I drop observations with missing or "don't know" responses.