NBER WORKING PAPER SERIES

AIDS, "REVERSAL" OF THE DEMOGRAPHIC TRANSITION AND ECONOMIC DEVELOPMENT: EVIDENCE FROM AFRICA

Sebnem Kalemli-Ozcan

Working Paper 12181
http://www.nber.org/papers/w12181

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 2006

I thank Daron Acemoglu, David Bloom, David Canning, Aimee Chin, Jane Fortson, Andrew Francis, Chinhui Juhn, Gerhard Gloom, David Meltzer, Emily Oster, Yona Rubinstein, Bent Sorensen, Michele Tertilt, Rebecca Thornton, Dietz Vollrath, Romain Wacziarg, David Weil, seminar participants at Brown University, Harvard University School of Public Health, and participants of the conference on "Health, Demographics and Economic Development" at Stanford University, of the AIDS Workshop at Amsterdam Institute for International Development for valuable comments and suggestions. I thank Ms. Laura Heaton of the Health Studies Branch, Population Division, the U.S. Bureau of Census for providing her HIV estimates. Belgi Turan provided excellent research assistance. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

AIDS, "Reversal" of the Demographic Transition and Economic Development: Evidence from Africa

Sebnem Kalemli-Ozcan

NBER Working Paper No. 12181
May 2006, Revised January 2008
JEL No. O11,I12,J11,J13

ABSTRACT

Theoretical models of demographic transition imply that fertility declines as a response to a decline in mortality. These models take their cue from the historical pattern of the demographic transition, which suggests that fertility declines follow mortality declines, followed by a rise in human capital accumulation and economic growth. The HIV/AIDS epidemic is a shock to mortality that threatens to reverse this path. Using country, region, and household level data on fertility rates from a panel of African countries during 1985-2000, this paper mostly shows a positive effect of the epidemic on fertility. There are three sets of results: 1) Between estimates based on country data suggest a strong positive effect of HIV/AIDS on fertility both in OLS and in IV. These estimates predict that a country with a high level of HIV/AIDS prevalence, such as Zambia, have 1 more child per woman on average compared to a country with a low level of HIV/AIDS prevalence, such as Senegal. 2) Within country estimates show mixed results due to their sensitivity to different time trends. 3) Within estimates based on individual data from South Africa suggest a positive effect of HIV/AIDS in the 1990s, whereas pooling the survey data for 27 African countries suggests a zero effect.

Sebnem Kalemli-Ozcan
University of Houston
Department of Economics
Houston, TX 77204
and NBER
sebnem.kalemli-ozcan@mail.uh.edu

1 Introduction

"Economics is judged ultimately by how well it helps us understand the world, and how well we can help improve it." Gary Becker.

Nobody doubts that AIDS is the plaque of the 21th century that has disastrous effects especially in the developing countries. The impact of the epidemic on economic development is, on the other hand, a fiercely debated issue.[1] The epidemic has altered the patterns of both morbidity and mortality. In the 35 highly affected countries of Africa, life expectancy at birth dropped by 7 to 10 years in the last decade, bringing it down to 35 years left to live for a newborn in Botswana in 2007.[2] Thus, we should start by examining the large theoretical literature that links life expectancy to economic development.

Neoclassical growth models identify two effects. The first order effect of increased life expectancy is to increase population. Absent behavioral responses in fertility, reductions in mortality increase population, thus reducing capital-labor and land-labor ratios and depressing per capita income. This effect is offset to some degree if increased life expectancy, and more generally, better health, raises TFP and the rate of human capital accumulation. Models in the tradition of Becker and Barro (1988) that endogenize fertility (e.g. Cervellati and Sunde (2007), Tamura (2006), Soares (2005), Kalemli-Ozcan (2003), Boldrin and Jones (2002), Lucas (2002), Galor and Weil (1999), and Ehrlich and Lui (1991)) show that fertility may respond to reinforce this latter effect towards higher investment and growth. Hence,

[1] While most of the researchers find negative effects of the epidemic on economic growth, some find no effect and some even find positive effects. Bloom and Mahal (1997) run cross-country regressions of growth of GDP per capita on HIV/AIDS prevalence and find no effect, whereas Bonnel (2000) finds a negative effect within a similar framework. Over (1992), who also uses cross-country data, finds a reduction of 0.5 percent per year in per capita growth rates as a result of the epidemic. Papageorgiou and Stoytcheva (2007) find a negative significant effect of AIDS on income per worker but the effect is small. Werker, Ahuja, and Wendell (2006) instrument HIV/AIDS prevalence by national circumcision rates and show that there is no effect of the epidemic on growth of the African countries. Corrigan, Gloom, and Mendez (2005) show calibration results that imply large negative effects of the epidemic on growth. The results of Lorentzen, McMillan, and Wacziarg (2007) imply significant long-run costs of AIDS on various outcome variables.

[2] A similar picture emerges if we look at life expectancy at age 20 instead of life expectancy at birth, where the latter might be affected from infant mortality.

declines in mortality could lead to a quantity-quality trade-off where parents have fewer children but invest more in each child. These models suggest that fertility and mortality are positively related and behavioral response in fertility can undo and even reverse the initial rise in population size.[3] The HIV/AIDS epidemic has generated a negative shock to life expectancy that threatens to reverse the path to growth laid out in these models. A key question, then, is the following: will fertility responses further reinforce, mitigate or even reverse the disease-induced population declines brought about by the HIV/AIDS crisis?[4]

Drawing a parallel between AIDS and the "Black Death," Young (2005) suggests that population declines will lead to higher capital-labor ratios and eventually to higher per capita income in the affected countries. While the epidemic will have a detrimental impact on human capital accumulation, he postulates that widespread community infection will lower fertility, both directly through a reduction in the willingness to engage in unprotected sexual activity, and indirectly, by increasing the scarcity of labor and the value of women's time. Using household data on fertility from South Africa and relying on between-cohort variation in country-level HIV infection and fertility, he estimates a large negative effect of HIV prevalence on fertility. He concludes that even under the most pessimistic assumption for human capital destruction the fertility effect dominates and hence future per capita income of South Africa improves. Using similar household data on fertility from other African countries and HIV prevalence rate by country and time, Young (2005b) reaches a

[3]While not directly related to HIV/AIDS, a recent paper by Acemoglu and Johnson (2006) find no effect of life expectancy on level and growth of per capita income. They instrument changes in life expectancy with dates of global interventions in disease prevention. Their results suggest that an increase in life expectancy leads to an increase in population and fertility responses are insufficient to compensate. It may be the case, however, that many of the countries in their sample have not yet completed the demographic transition.

[4]While the focus of this study is the fertility channel, an equally important question is the effect of HIV/AIDS on human capital investment. A large number of papers cover this topic and generally find substantial negative effects. Meltzer (1992) argues that AIDS raises mortality of young adults, which is going to have the biggest effect on the rate of return on educational investment. He claims for a 30 percent HIV positive population like Botswana, there would be a 6 percent reduction in the rate of return to education relative to no HIV. Bell, Shantayanan, and Gersbach (2003), using household survey data from South Africa argue that the long-term economic costs of AIDS could be devastating because of the cumulative weakening from generation to generation of human capital. Fortson (2007) shows children currently growing up in Africa, including non-orphans, will complete 0.3 fewer years of schooling compared to the case of zero HIV prevalence.

similar conclusion.

This paper investigates the effect of the epidemic on fertility, using country, region and individual level data for fertility from a panel of African countries during 1985–2004. Is fertility in Africa really declining? As a first cut, data from Demographic Health Surveys (DHS) do not seem to suggest so. Figure 1 plots data on total fertility rate (TFR) from DHS for Kenya and shows that after more than a decade of rapid decline, the total fertility rate actually increased starting in the late 1990s. Westoff and Cross (2006) find the increase in fertility in Kenya is most pronounced for the least educated group of women. They also find a significant increase in the percentage of women who report wanting more children for each age group. Similar data from ten other countries show either an uptick for fertility, such as in Nigeria and Mozambique, or a stall in fertility transition, such as Uganda and Cote D'Ivoire, as shown in figure 2.[5] Hence the survey data from DHS might indicate the start of a "reversal" in the fertility transition.

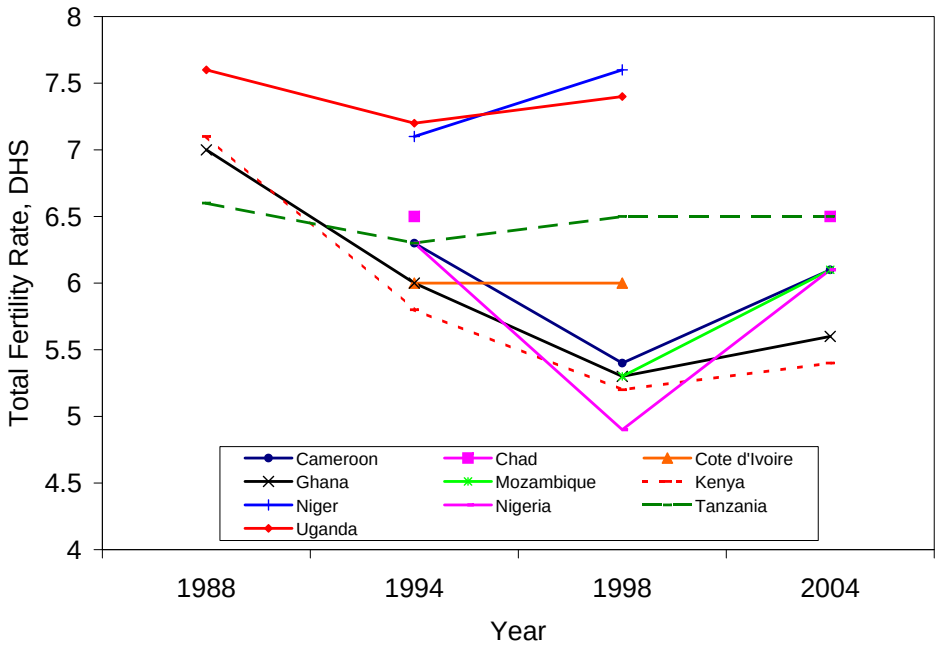
I use four different indicators for HIV/AIDS, two of which are available both at the country and at the regional level. There are three sets of results: 1) Between estimates based on country data suggest a strong positive effect of HIV/AIDS on fertility both in OLS and in IV. These estimates predict that a country with a high level of HIV/AIDS prevalence, such as Zambia, have 1 more child per woman on average compared to a country with a low level of HIV/AIDS prevalence, such as Senegal. 2) Within country estimates show mixed results due to their sensitivity to different time trends. 3) Within estimates based on individual data from South Africa suggest a positive effect of HIV/AIDS in the 1990s, whereas pooling the survey data for 27 African countries suggests a zero effect.

The positive results are consistent with various theoretical models that concentrate on two channels in general. On the one hand, rising adult mortality shortens the time horizon of parents. Hence parents may invest less in their own human capital and in their children's human capital and have more children, as in Soares (2005). On the other hand, parents, faced with a high mortality environment for young adults in the future, may develop a

[5]Each countries survey year is on or around the dates shown on the x-axis.

Figure 1: Fertility in Kenya, Demographic Health Surveys: 1989, 1993, 1998, 2003

Figure 2: Fertility in Africa, Various Demographic Health Surveys



precautionary demand for children due to uncertain survival for adolescents and hence may choose to have more children and provide them with less education, as in Kalemli-Ozcan (2003).

The rest of the paper is structured as follows. Section 2 outlines the conceptual framework and also discusses the multidisciplinary literature on HIV/AIDS. 3 examines the data and discusses the various issues surrounding the country level HIV/AIDS estimates. Section 4 presents the econometric framework, identification strategy and the empirical analysis. Section 5 compares the results of this paper to those of Young (2005, 2005b). Section 6 concludes.

2 Conceptual Framework

In this section I present a simple reformulation of the theoretical models that link fertility to an increase in mortality; specifically simplified variants of Soares (2005) and Kalemli-Ozcan (2003).

2.1 Deterministic Survival

The models of Meltzer (1992) and Soares (2005) rely on the fact that the longer the adults live the more human capital investment they will undertake in themselves, which in turn will lead to a quality-quantity trade-off. To demonstrate their mechanism simply consider an economy inhabited by adult individuals who live for a deterministic amount of time and allocate their time to invest in their *own* education, work, consume, and have children.[6] A fraction β of children born die before reaching adulthood. Adults live for T periods and derive utility from their own consumption, c and from the human capital of their children, h, which is a linear function of their own human capital, H, given as $h = bH + d$.[7] Children have a time cost, b. Parents invest in their own education, e. Hence, adult human capital

[6]This section borrows heavily from Soares (2005).

[7]Note that only the partial equilibrium is being presented here. Economy wide production will be a function of adult human capital, H.

production is given as, $H = eh_0 + D$, where h_0 is the basic parental human capital inherited from own parents.[8] Parents also care about the number of children, n and how long they live combined in an altruism function, ρ, that multiplies the utility from children's human capital. Hence the utility function and the budget constraint are given as (ignoring the time subscript),

$$\begin{aligned} U &= T\frac{c^\sigma}{\sigma} + \rho(n, T, \beta)\frac{h^\alpha}{\alpha} \\ TH &= Tc + n + (bn + e)H \end{aligned} \tag{1}$$

To present the static implications of longevity losses in partial equilibrium we use the first order conditions for maximization to show,

$$\frac{\partial\rho/\partial n}{\rho/n} = \alpha \tag{2}$$

Hence, the individual equates the elasticity of the altruism function with respect to the number of children to the constant elasticity of the utility from human capital of children. Combining above equation with the altruism function and using the implicit function theorem gives,

$$\frac{dn}{dT} = -\frac{\rho n - \frac{\partial\rho}{\partial T}\frac{\partial\rho}{\partial n}n}{\rho n[\frac{\partial^2\rho}{\partial n^2} - \frac{\partial\rho/\partial n}{\rho}(\frac{\partial\rho}{\partial n} - \frac{\rho}{n})]} < 0 \tag{3}$$

The sign follows from the assumptions that a decrease in adult longevity, T, and an increase in child mortality, β, increase the marginal utility of fertility; and the elasticity of the altruism function is decreasing in the number of children.[9] The way I interpret this model in the context of HIV/AIDS is that the epidemic will cause a decrease in T. Given the representative agent framework, T should be declining as a result of the community HIV, which will lead a rise in fertility.

[8] d and D represent innate human capital in the absence of any investments.

[9] See Soares (2005) for the justification of these assumptions.

2.2 Uncertain Survival of Adolescents

An alternative modeling strategy will rely on the uncertain survival of adolescents generated by the high mortality risk as argued by Sah (1991), Kalemli-Ozcan (2003), and Tamura (2006). The framework presented in section 2.1 abstracted from this type of uncertain survival in order to focus on the impact of *adult* longevity on the economic incentives faced by the individuals. However, in the context of HIV/AIDS, the uncertain survival of adolescents might have important consequences. Rising adult mortality will shorten the time horizon of parents leading to a quality-quantity trade-off as argued above. It is also plausible that, parents faced with a high mortality environment for young adults, may develop a precautionary demand for children due to uncertain survival and hence may choose to have more children and provide them with less education.

Consider a similar structure as before, where parents have a total time of unity instead of, T. The difference is that here the number of survivors, N, is a random variable. Hence N will not be equal to the expected number of survivors, $n(1-\beta)$ as before. Parents get utility from their own consumption and from the total amount of human capital of their survivors, where with the education investment in each child being, s, the human capital production function is given as $h = s^{\alpha}h_0$. Hence each newborn has a fixed cost, v, and an education cost but only some survive. The expected utility and the budget constraints are given as,

$$\begin{aligned} U &= U(c) + E(U(N,h)). \\ h(1-(v+e)n) &= c. \end{aligned} \tag{4}$$

Let q be the survival probability of each child, which is fixed over time. N, the number of survivors, is a random variable drawn from a binomial distribution. Thus, the probability that N out of n children will survive is,

$$f(N;n,q) = \binom{n}{N} q^N (1-q)^{n-N} \qquad N = 0, 1, \ldots, n. \tag{5}$$

Hence maximization becomes,

$$E(U) = \sum_{N=0}^{n} \Big\{ \Big[U(c) \Big] + \Big[U(N,h) \Big] \Big\} f(N; n, q). \tag{6}$$

This formulation implies that the number of children born and the number of surviving children are represented as nonnegative integers, which is a discrete representation. To have continuous representation we linearize the utility function around the mean and the variance of the binomial distribution and get the first order conditions, which will give us the following comparative static via the implicit function theorem,[10]

$$\begin{aligned} \frac{-(v+e)}{1-(v+e)n} + \frac{1}{n} + \frac{1(1-q)}{2qn^2} &= 0. \\ \frac{dn}{dq} &< 0, \ \forall q. \end{aligned} \tag{7}$$

In the context of HIV/AIDS, the epidemic will cause a decrease in q, as a result of the community HIV, which will lead a rise in fertility. Parents might respond to the higher mortality environment by having more children to guarantee a certain number of survivors, where most of the AIDS related deaths come later in life and hence it will be biologically impossible to replace the dead children (apart from the ones die in infancy due to AIDS). I am assuming that parents presume their children will be infected via sexual activity, which probably will not start before early teen years.

2.3 Possible Fertility Responses in the Special Case of HIV/AIDS

I have so far considered HIV/AIDS as a shock to adult longevity, T or to the survival probability of the child, q. However, there are characteristics of HIV/AIDS which suggests that this formulation is overly simplified. First, field evidence suggests that there is a direct

[10]The form of the utility function does not matter as long as the third derivative is positive. The closed-form solutions shown here are obtained using a log utility; See Kalemli-Ozcan (2003).

biological impact of the disease which lowers the fecundity of infected women, an effect which should be considered separately from the behavioral responses.[11] Fecundity is reduced by HIV infection due to higher rates of miscarriage and stillbirth and high rates of co-infection with other sexually transmitted infections, which may cause secondary infertility.[12]

Second, since it is a sexually transmitted disease, the impact on fertility can come through changes in sexual behavior. The impact of the disease on sexual behavior in Africa has proven to be much debated topic. Mwaluko et al. (2003), Bloom et al. (2000), Stoneburner and Low Beer (2004), Lagarde et al. (1996), Lindan et al. (1991), Ngwshemi et al. (1996), Williams et al. (2003), Caldwell et al. (1999) all find no change or very small change in sexual behavior. Luke and Munshi (2004) find that within Kenya, in a high AIDS prevalence environment married men are no different than single men in the number of non-marital partners. One would expect the number of non-marital partners to fall more for the married men if unprotected sexual activity is an issue or if wives could influence husband's extra-marital sexual activity.[13] Oster (2005), using DHS data on sexual behavior from a subset of African countries finds that sexual behavior changed relatively little since the onset of the epidemic. Other researchers finds some evidence of risky behavior reductions in Zambia and Zimbabwe such as reductions in multiple partners; see Cheluget et al. (2006), Fylkesnes et al. (2001), and the epidemiological modeling study of Hallet et al. (2006).

Oster (2005) suggests that the relatively little response in sexual behavior may be in part

[11]Many African studies, both clinic and cohort based, indicate lower fertility (around 40 percent) and childbearing odds among HIV positive woman. Gray et al. (1998), in a cross-sectional analysis of a Ugandan community, find an HIV induced reduction in fertility of 55 percent. Carpenter et al. (1997) and Hunter et al. (2003), in cohort studies in Uganda and Tanzania, respectively, find a 30–40 percent reduction in probability of becoming pregnant. See also Zaba and Gregson (1998) and Noel-Miller (2003).

[12]It is hard to separate out the biological effect from the behavioral response without data on individual HIV status. In Juhn, Kalemli-Ozcan, and Turan (2007), we take a first step in separating these two effects by utilizing recent rounds of the Demographic Health Surveys (DHSs) which link an individual woman's fertility outcomes to her *own* HIV-status, based on testing. The data allows us to distinguish the effect of own positive HIV status on fertility from the behavioral response to higher mortality risk, as measured by the local community HIV prevalence. We find that the disease significantly lowers an infected woman's fertility. In contrast to Young (2005), however, we find that local community HIV prevalence has no significant effect on non-infected women's fertility when we pool all communities while it has a significant *positive* effect in the sub-sample of communities with non-zero prevalence.

[13]Sociologists have long argued that in Africa married women don't have a lot of power over their husband's extra-marital sexual activity.

explained by low levels of knowledge about the disease. Data from DHS surveys, presented in table 1, show that the percentage of the female population that requests an HIV test, gets tested, and receives results is very small, the mean being 5.7 percent across 10 African countries with an average HIV prevalence of around 15 percent. There is little systematic evidence that countries with higher prevalence have better knowledge or perceptions of risk, as shown in figure 3.[14] The percentage of 15-49 years old women who know that HIV can be transmitted from mother to child is plotted, with a mean of 38 percent.[15] Recent evidence in Thornton (2006), however, suggests that knowledge alone may not account for the limited response of sexual behavior in many African countries. The paper is based on a randomized experiment in Malawi in which individuals were given monetary incentives to get tested and learn about their HIV status. Using randomized incentives as instrument for knowledge, she finds that those with positive HIV status were more likely to purchase condoms, a limited amount though, and there was no change in behavior among those with negative HIV status. A recent paper by Oster (2007) also argues along these lines suggesting that shorter life expectancy and lower income could account for the large differences in behavioral response between individuals in Africa and the gay population in the U.S. Young (2005) strongly argues, on the other hand, that the reduction in willingness to engage in unprotected sex is a major component of the overall fertility response to HIV/AIDS.[16]

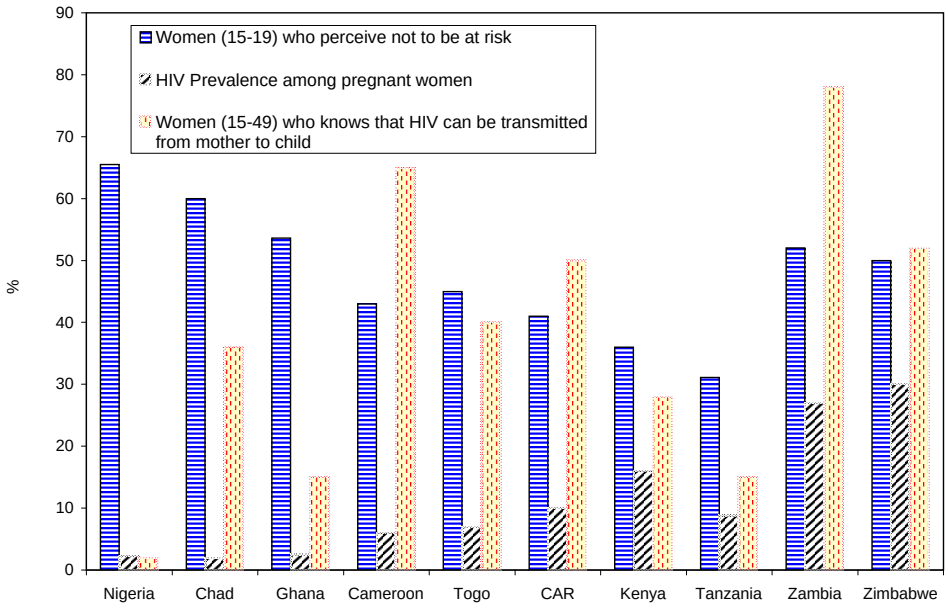
Third, regardless of changes in sexual behavior, it may be the case that infected women who know their own status and have knowledge about mother-to-child transmission would

[14]Sentinel surveillance programs (a form of surveillance relates to a particular group) monitoring HIV/AIDS epidemic in Africa are not designed to detect and notify at-risk individuals. They are conducted using anonymous and unlinked blood samples from hospital blood donors, pregnant women attending antenatal clinics (ANC), or sexually transmitted disease (STD) clinic attenders. Thus, those with HIV who are tested will not receive a notification of their status.

[15]Mother-to-child transmission is 30 percent at birth and 3 percent with every month of breastfeeding. One must also note that the questions on knowledge and perceptions are typically asked to those who already heard about AIDS. Given the high fraction of the ones who heard about AIDS this does not constitute a big issue.

[16]The Economist (2006) reports that, in the recent trial of Jacob Zuma, South Africa's former deputy president, Jacob Zuma said that he took a shower to minimize the risk of infection after having sex with an HIV-positive woman, who accused him of rape. Subsequently, HIV/AIDS helplines report having a lot of difficulties in dealing with confused callers.

Figure 3: Risk Perception for Selected African Countries, Demographic Health Surveys, 1994–2000



want to reduce fertility rather than give birth to infected children. Again the field evidence on this channel is mixed where findings range from HIV positive women not changing their fertility to reducing it.[17]

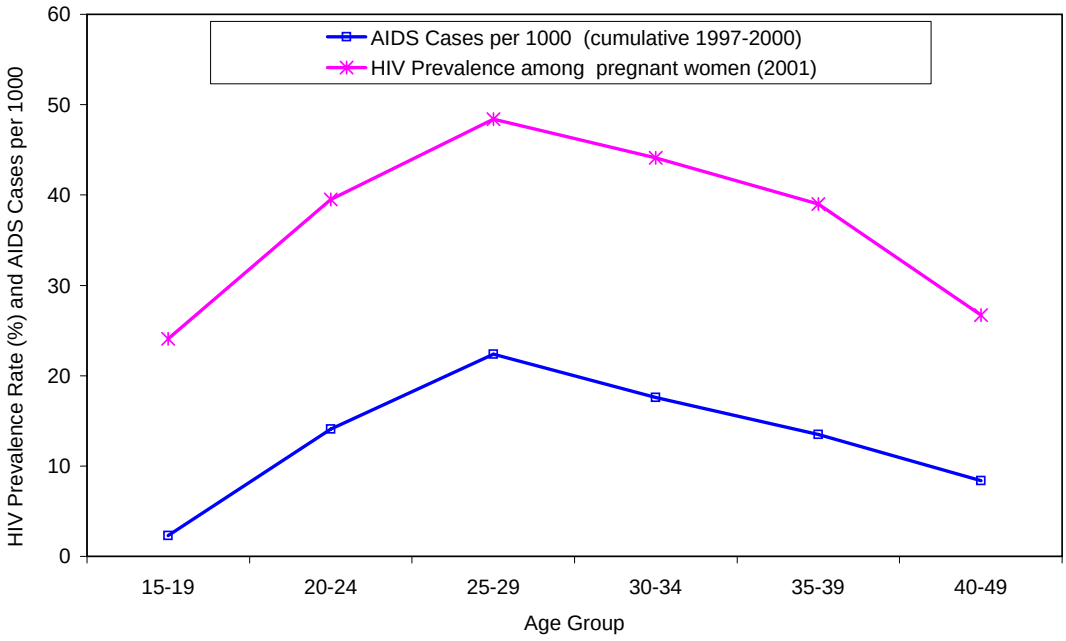
Uninfected people, and people who think they are not at risk, might behave differently. If they know that there is a high level of mortality in their surrounding population, they might reduce their risky sexual activity which will lead to lower fertility as a by-product, or they might increase their fertility along the lines of the outlined models in sections 2.1 and 2.2 since the epidemic causes a rise in adult and youth mortality. HIV/AIDS prevalence peaks around age 25−30 in general as shown in figure 4. Figure 5 shows the mortality profile for adults and children as a function of time since infection. In the absence of antiretroviral therapy, the median survival time for adults is 9 years. The estimates also imply that all infected children die by age 12. Figure 5 also shows estimates from Feeney (2001) for Zimbabawe. The probability of a 15 year old dying before age 50 shows a sharp increase since late 1980s, implying high mortality for young adults due to the epidemic during this time period.[18] The higher probabilities (around 50 percent) implied by the household reports might reflect the rapidly rising mortality that is captured in those surveys which are undertaken in 1997 (top x-axis) relative to others that are done earlier. These also reflect the subjective probabilities of the family members who experienced the deaths due to AIDS very closely.

Overall there might be various biological and behavioral responses of fertility to the HIV/AIDS epidemic which might lead to a reduction in fertility or an increase. A reduction

[17]Temmerman et al. (1990) find that in Nairobi a single session of counseling—which is common in most African countries—has no effect on the subsequent reproductive behavior of HIV-positive women. Allen et al. (1993) using cohort data from Kigali, Rwanda, find that in the first 2 years of follow-up after HIV testing, HIV-negative women were more likely to become pregnant than HIV-positive women. However, among HIV-positive women, those with no children were more likely to become pregnant than those with children and married women are more likely to become pregnant than unmarried women. The desire to have children among HIV-positive women altogether was 45 percent. On the other hand, Noel-Miller (2003) using panel data from Malawi shows that women who have higher subjective HIV risk perceptions for themselves were less likely to have children.

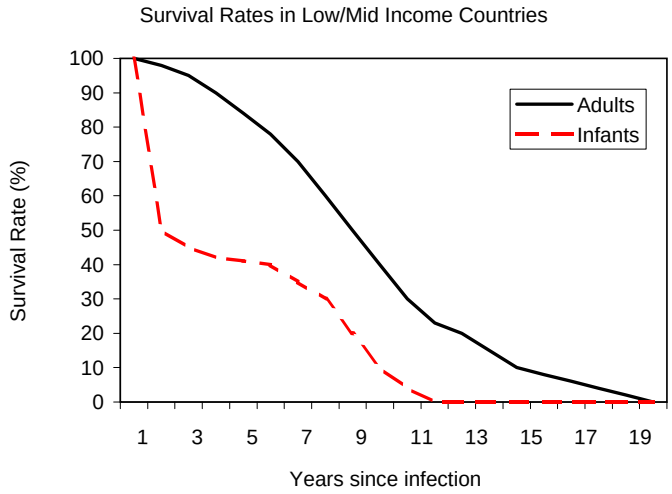
[18]This probability is defined as q_{15}^{35} in demographic terminology. Records from vital registration, reports from households and reports from surviving siblings all show an upward trend. Feeney (2001) argues the discrepancy between registered deaths and sibling reports comes from the fact that the former is adjusted for underreporting and the latter is not.

Figure 4: HIV/AIDS by age, Botswana

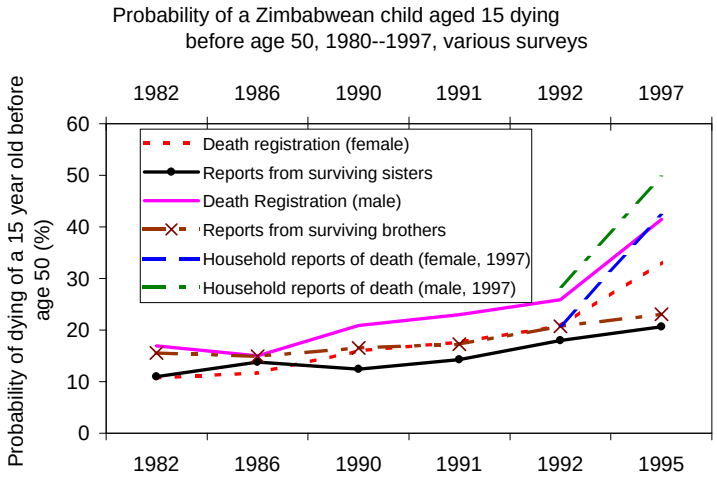


Data: Botswana 2001 HIV Sero-Prevalence Sentinel Survey among pregnant women.

Figure 5: Survival for Adolescents



Data: UNAIDS Reference Group, 2002.



Data: Feeney, 2001.

in age-specific fertility rates among HIV positive woman due to the biological and physiological responses may serve to reduce total fertility in a high HIV prevalence country in the absence on any behavioral response from the uninfected woman. Behavioral response from the infected women (if they know their own status) will also cause a reduction in fertility. Last but not least uninfected womens' fertility might also decrease due to a reduction in risky sexual behavior. Put it differently, for fertility to increase as a result of the epidemic any positive behavioral response of uninfected women have to overcome the negative biological and behavioral responses. The empirical analysis below aimed at testing which group dominates overall. I would like to stress that in the case of a negative finding we will not be able to tell whether this is due to a biological response or a behavioral response given the lack of data on *own* HIV status.

3 Data: Sources and Issues

3.1 Country Level

Fertility:

I use country level data on total fertility rates (TFR) both from World Bank, World Development Indicators and from Demographic Health Surveys (DHS).[19] WB data are available for 44 countries and ten years between 1985–2004. DHS data on fertility rates per woman ages 15–49 are available for 34 countries, where most countries has only one or two surveys. Only 3 out of 34 countries have four surveys, 10 countries have three surveys, 10 countries have two surveys and the remaining 11 countries has one survey. Survey years fall between 1986–2004. TFR is the sum of age-specific fertility rates; it is an approximation for the average lifetime fertility of women, i.e. for completed fertility. I also use data on desired fertility rate per woman ages 15–49, available for 34 countries, from DHS. There are two

[19]The World Bank uses UN World Population Prospects for every 2 years and update the UN data with the latest survey data such as DHS, MICS and so on. UN data comes from the countries vital registration system which is questionable quality in the case of Africa.

measures for desired fertility. The first is the wanted fertility rate and the second is the ideal number of children. Details of the variables and a full list of countries and survey years are provided in the appendix.

There might be a concern in the use of total fertility rates as the dependent variable. TFR uses *current* age specific fertility rates to project the number of children that a woman would have if she lived through all her childbearing years. In dynamic situations with evolving fertility environments it might be problematic to apply the current age-specific rates to project total fertility of women who have many years of childbearing. The alternative is to use individual level data on the recent behavior, such as births in the past year. The problem is that this is not a measure of completed fertility either and can also be problematic in the context of a dynamic epidemic. Different cohorts might have different behavioral responses such as younger cohorts increasing their fertility and older cohorts decreasing it. The paper at hand will be using both measures to get a better insight.

HIV/AIDS:

I use four different indicators for HIV/AIDS at the country level, none is perfect and all have different problems.[20] For AIDS, I use data that come from UNAIDS/WHO, Epidemiological Fact Sheets (2003). These are the number of reported AIDS cases available for each country in every year between 1985–2004.[21] I multiply the number of reported cases by 100,000 and divide by the country's population in each year, to obtain rate per 100,000 per country

[20]The correlation between different indicators is around 70 percent on average. The indicators in general suffer from different biases. Classical measurement error is one but there can also be other errors that are not classical. For example, since most of the indicators are based on estimates from antenatal clinics, the measurement error might be correlated with the population attending the clinics, which itself might be correlated with fertility. Hence it is not straightforward to gauge the direction of the resulting bias due to the measurement error in HIV/AIDS on the estimated coefficients.

[21]UNAIDS definition of AIDS (Acquired Immunodeficiency Syndrome) is that AIDS is the most severe manifestation of infection with the HIV (human immunodeficiency virus). The Centers for Disease Control and Prevention (CDC) lists numerous opportunistic infections and neoplasms (cancers) that, in the presence of HIV infection, constitute an AIDS diagnosis. In 1993, CDC expanded the criteria for an AIDS diagnosis to include CD4+ T-cell count at or below 200 cells per microliter in the presence of HIV infection. In persons (aged 5 and older) with normally functioning immune systems, CD4+ T-cell counts usually range from 500 to 1500 cells per microliter. Persons living with AIDS often have infections of the lungs, brain, eyes and other organs, and frequently suffer debilitating weight loss, diarrhoea, and a type of cancer called Kaposi's sarcoma.

per year. According to UNAIDS, AIDS case reports come from surveillance systems of varying quality. Reporting rates vary substantially from country to country and low reporting rates are common in developing countries due to weaknesses in the health care systems. Hence there can be systematic biases such as in countries with worse medical institutions (which is probably correlated with other country characteristics) underreporting will be worse. AIDS case reporting provides information on transmission patterns and levels of infection approximately 5-10 years in the past, limiting its usefulness for monitoring recent HIV infections. Despite these caveats, AIDS case reporting is useful in estimating the burden of HIV-related mortality.

For HIV, I use three different indicators. First, I use data on HIV prevalence rates among pregnant women that are from the U.S. Census Bureau, HIV Surveillance Database (2005). UNAIDS/WHO also provides similar data. This is the indicator that is used by most researchers.[22] Both U.S. Census and UNAIDS databases collect regional estimates of HIV/AIDS prevalence since the early 1980s. The main indicator for the epidemic is the percent HIV-1 incidence among pregnant women for each country and year.[23] However, these estimates are in general very high. Representativeness of these estimates for the general population is also debatable since they are based on pregnant women and high risk groups, which in turn is the main reason for these inflated estimates.[24] More recently, DHS started providing results from population-based HIV testing. These new estimates are much lower than the UNAIDS and U.S. Census estimates.[25] The new population based DHS estimates are only available for a limited set of countries for their latest survey though and hence do not provide enough information about variation over time, as the HIV estimates from the

[22]See Young (2005), Werker et al. (2006) and Oster (2005).

[23]HIV is the retrovirus isolated and recognized as the etiologic (i.e. causing or contributing to the cause of a disease) agent of AIDS. HIV-1 is classified as a lentivirus in a subgroup of retroviruses. Most viruses and all bacteria, plants, and animals have genetic codes made up of DNA, which uses RNA to build specific proteins. The genetic material of a retrovirus such as HIV is the RNA itself. HIV inserts its own RNA into the host cell's DNA, preventing the host cell from carrying out its natural functions and turning it into an HIV factory.

[24]See Timberg (2006) and McNeil (2007).

[25]See Oster (2006) and Juhn, Kalemli-Ozcan and Turan (2007) for a comparison of the various estimates.

Surveillance Database. On the other hand, the time series variation in these prevalence rates from the Surveillance Database of U.S. Census and UNAIDS is far from perfect. UNAIDS (2006) notes in its most recent report on the epidemic that it is not possible to use previous reports to compare prevalence over time. Using the U.S. Census HIV surveillance database suggests that HIV rates are flat or falling over the 1990s in virtually all countries in Africa, which seems inconsistent with the casual observation. A close inspection of these estimates shows that there is considerable year to year variation which calls into question the reliability of the time variation in these data. It has been suggested elsewhere that selection of locations that the estimates are collected are changing over time. Based on these problems, Oster (2006) develops a methodology to estimate HIV prevalence over time from mortality data. To avoid the problem of lack of official mortality statistics for Africa she takes advantage of sibling mortality histories in the DHS. She has HIV estimates for 9 countries since mid-1980s.[26] I use her estimates as a second indicator for HIV.

As a third indicator for HIV, I use the *projected* HIV from the U.S. Census Bureau, International Programs Center. The International Programs Center uses Estimation and Projection Package (EPP) from WHO/UNAIDS to project adult HIV prevalence among 15–49 year old from U.S. Census Surveillance data between 1985–2004.[27] While EPP can be used in all countries with sufficient surveillance data, it is specifically recommended for countries with "generalized epidemics." Generalized epidemics are those that have broken out into the general population or consistent HIV prevalence at over 1 percent in low risk individuals. The proxy for low risk individuals is women attending antenatal clinics. Thus, the inputs to EPP in countries with generalized epidemics are the same surveillance data on HIV prevalence among pregnant women. EPP estimates the trends over time of HIV prevalence by fitting an epidemiological model to data from urban and rural sites.[28] Although EPP model fits a

[26]As argued by Oster (2006), finding of a small effect of HIV on economic development is sensitive to the size of the estimates used. If the true HIV rates are only a third or half of the estimated HIV rates, the effects may be two or three times as large as estimated.

[27]These EPP data are used by Young (2005b).

[28]It chooses a set minimizing least squares and projects future course based on fitted parameters, such as a parameter for the start year of the epidemic; one for the force of infection (how explosive the epidemic is in

somewhat flexible curve to a not so long time series, the modeling is still an issue of concern given the dynamic nature of the epidemic. In addition, the input data are still the same HIV data that comes from the U.S. Census Bureau, HIV Surveillance Database (2005) and UNAIDS, where the representativeness and the quality of the time series variation are both questionable as argued above.

I also use data on *perceptions*, specifically the variable "know someone died of AIDS." The data on the percent female who know someone personally who has the virus that causes AIDS or has died of AIDS are from DHS. This is the ideal measure for the purpose of this paper however since this question has only been asked in the most recent surveys the data are available only for 22 countries whose survey years fall between 1993–2004.

Other Controls:

All other controls such as different measures of female schooling, child mortality, urbanization, contraception, marital status, and GDP per capita are taken from World Bank, WDI, and from DHS. The details of these data are provided in the appendix.

3.2 Regional Level

Fertility:

I use data on regional total fertility rates from DHS. They are available for 71 regions from 14 countries, whose surveys years fall between 1988–2004. A full list of regions is provided in the appendix.

HIV/AIDS:

The data for regional HIV rates come from U.S. Census Bureau, HIV Surveillance Database (2005) and available for 40 regions from 13 countries between 1985–1990. The overlap between the regional fertility rates and HIV rates give us 32 regions from 12 countries.

its initial stage); one for the fraction of new entrants to the population going into to the at-risk category (a parameter largely determines where the epidemic levels off); and one for the recruitment (a high value means people are brought into the at-risk population as people die of HIV, thus helping to sustain the epidemic at a higher level).

3.3 Individual Level

To replicate the results of Young (2005, 2005b) individual level data are used for 27 countries from Demographic Health Surveys. In the 27 country sample of Young only 2 countries have four surveys, 10 countries have three surveys, 8 countries have two surveys, and the remaining 7 countries have only one survey. A full list of survey countries and years are provided in the appendix.

Fertility:

Fertility is measured as the number of births in last year for each woman. Desired fertility, measured as the ideal number of children for each woman, is also used. Both measures are from DHS. For South Africa, where there is only one survey, a panel is constructed using each womans birth histories since age 12, including the period between 1961–1998. Retrorespective fertility for South Africa is then, the number of pregnancies of each woman in each year, including that were lost before term or resulted in stillbirths.

HIV/AIDS:

Following Young (2005, 2005b) country level *projected* HIV-EPP rates are used from the U.S. Census Bureau, International Programs Center. For South Africa, Young (2005) uses historical antenatal clinic HIV seroprevalence rate, from South Africa Department of Health.[29] The same source is cited in the U.S. Census Bureau's HIV/AIDS surveillance database (2005) and hence I use South Africa HIV prevalence rates among pregnant women from this database.

[29] In spite of many emails and calls these data could not be obtained from South Africa Department of Health.

4 Econometric Framework, Identification Strategy and Empirical Analysis

4.1 Framework and Identification

Theoretical models of the demand for fertility have the following empirical predictions: 1) increased education of women raises the cost of childbearing and reduces fertility; 2) reduced child mortality, assuming the demand for surviving children is price inelastic, is associated with a decline in fertility;[30] 3) increased income per capita increases demand for children since they are normal goods; 4) the net cost of child bearing is greater for parents in urban than in agricultural settings. Thus, I control for these determinants, that are shown to be significant in the other empirical studies,[31] in a regression of total fertility rate on the indicators of HIV/AIDS. I estimate Ordinary Least Squares (OLS) regressions of the following form, using both country and regional level data:

$$TFR_i = \alpha + \beta HIV/AIDS_i + \mathbf{X}_i'\gamma + \epsilon_i, \tag{8}$$

where TFR_i is the total fertility rate for country i, $HIV/AIDS_i$ is the indicator for HIV or AIDS for country i, $\mathbf{X}_i$ is a vector of other covariates, and ϵ_i is a random error term.[32] The coefficient of interest is β, the effect of the epidemic on fertility. Recall that four different indicators for HIV/AIDS is used: I use AIDS cases per 100,000 per country per year from UNAIDS. I will call this variable "AIDS." Next, I use the HIV prevalence rates among pregnant women that are from the U.S. Census Bureau. I will call this "HIV." I also use Oster (2006) estimates, which I will call "HIV-Oster." Finally, I use the projections of the U.S. Census Bureau, which I will call "HIV-EPP."

Notice that the regression presented in equation (8) only exploits variation *between* coun-

[30] Notice that inelastic demand ceases to be a necessary condition once you introduce uncertainty about child survival into the model. See Kalemli-Ozcan (2003).

[31] See Schultz (1997) for an example.

[32] This regressions is also run at the regional level with country dummies included, i.e., for region r: $TFR_r = \alpha_i + \beta HIV/AIDS_r + \mathbf{X}_r'\gamma + \epsilon_r$, where α_i is the country dummy.

tries, i.e., it is a "between regression." A *within* regression framework to identify the parameters using only within-country variation over time is preferable since this framework will do a better job in controlling unobserved country heterogeneity. However, as summarized above we have no good information on the time variation of the epidemic hence I am hesitant to rely solely on within country time variation by using first differences or country fixed effects. I will present results for both frameworks to get a better insight. Total fertility rates were falling in almost all the African countries before the HIV/AIDS epidemic. Thus, I run a panel regression with country and time fixed effects. I also run the same regressions with a general time trend and country specific time trends. The "within regressions" are of the form:

$$TFR_{it} = \mu_i + \lambda_t + \psi HIV/AIDS_{it} + \mathbf{X}'_{it}\theta + \varepsilon_{it}, \tag{9}$$

where TFR_{it} is the total fertility rate for country i at time t, μ_i is the country fixed effect, λ_t is the time fixed effect, $HIV/AIDS_{it}$ is one of the four indicators for HIV/AIDS, $\mathbf{X}_{it}$ is a vector of other covariates, and ε_i is a random error term.

The econometric framework presented in equations (8) and (9) posits an endogeneity problem since HIV/AIDS is related to sexual behavior and marriage markets, both of which are independently related to fertility. Areas with initially higher levels of sexual behavior will have higher HIV rates and they may also have higher rates of fertility. Also there are compelling reasons to believe that HIV infection is higher in areas with greater population density and economic activity. Then, country level HIV rates suffer from an omitted variables bias since countries that are the most economically active may have both higher infection rates and lower fertility, the latter being due to possibly the higher cost of women's time. Failing to control any variable that is negatively correlated with the epidemic such as female education will cause a downward bias. There might also be a bias due to unobservable factors such as culture, prudence of the parents, and governments' response to the epidemic. These factors may determine HIV/AIDS and fertility simultaneously.

The conditioning variables should take care of the large part of the effect of the differential development levels. Nevertheless, I will undertake a simple falsification exercise that

investigates the relationship between pre-AIDS fertility and current HIV. The "within regressions" are immune to the unobservable factors that are time-invariant such as religion, climate and culture. However, individuals may start taking less risks as a result of the epidemic over time or across places, which will bias both "between" and "within" estimates. As argued in the previous section the evidence from micro studies are such that there is limited change in sexual behavior as a result of the epidemic. Even if there is a decrease in risky sexual behavior and people start taking less risks (more condoms, fewer partners or abstaining) because of HIV/AIDS then fertility will decrease as a by-product, and hence a negative relation between fertility and HIV/AIDS will be the result. This would be true assuming that despite changes in sexual activity HIV rates remain high. Ultimately, it is plausible that, societies which lower their level of risky sexual activity are likely to experience declines in HIV rates and in fertility levels. Hence the initial negative relationship between HIV and fertility as a result of changes in risky sexual behavior can turn into a positive one in the long run. I do not expect the long-run effect to dominate the initial effects for the time period that this paper is concerned with.[33] Hence the results of this paper constitute a lower limit for the effect of HIV/AIDS on fertility.

To further deal with the problem of endogenity, I will follow Oster (2007) and instrument HIV/AIDS by the distance to the origin of the epidemic, which is Democratic Republic of Congo. Oster (2007) argues that two factors determine HIV prevalence within a given area are the speed at which the prevalence increases and the date at which the virus is introduced. The speed of increase, in turn, is determined by sexual behavior and the viral transmission rate. Hence the viral transmission rate or the arrival date of the virus, are potentially plausible instruments. She focuses on the virus arrival date. However, she also argues that using date directly is problematic since testing early in the epidemic is very limited and hence it is likely that the first date that the virus is observed is correlated with sexual behavior, which is also related to fertility for my case. She uses distance as an alternative since she

[33]If sexual behavior declines for some other reason than HIV/AIDS, then this will lead a positive association between fertility and the epidemic since both will decline as a result. One cannot rule this out. But this is very unlikely.

argues that if the virus takes time to travel, moving from person to person, areas further from its origin should have lower prevalence on average. She uses the longitude and latitude of each DHS survey cluster to calculate the distance of cluster to the center of Congo (middle of the country) since the virus is originally observed on both sides. In a similar fashion, I use the distance from the capital city of each country in my sample to the capital city of the Democratic Republic of Congo, which is measured as the distance between the center of the capital cities.[34] In the regional regressions, I use the distance of each region to the center of the epidemic as calculated by Oster (2007).

For the instrument to be valid, it must be correlated with the HIV/AIDS but uncorrelated with the fertility rate, except through the variable of interest that is included in equation explaining fertility. The most obvious way in which distance to the origin of the epidemic might systematically affect the fertility rate—other than via HIV/AIDS— is through its correlation with geographic and/or socioeconomic variables. Having controlled for these factors, as will be shown in detail below, it seems plausible to argue that distance to the origin of the epidemic will be otherwise unrelated to fertility.

4.2 Descriptive Statistics

Table 2 shows the mean, maximum, minimum, and standard deviation of the dependent and independent variables. Fertility rates vary from 2 children to 8 children depending on the source of the data, where desired fertility rates have lower means. Gross school enrollment rates show large variation with the enrollment rate being 5 times higher in the country with the highest enrollment rates than in the country with the lowest enrollment rates. For AIDS, the most affected country has prevalence that is 8000 times higher than that of the least affected country. The difference in the HIV prevalence between the highest and lowest prevalence country is 250 times in UNAIDS and U.S. Census data but only 10 times in the

[34] I also use the alternative instrument of circumcision as used by Werker et al. (2006). Circumcision might be less appropriate since it is highly correlated with ethnic group, which is likely to be correlated in turn with the fertility behavior.

Oster (2006) estimates. GDP per capita moves between 100 and 6000 dollars. The remaining variables also show extensive variation.

4.3 AIDS, HIV and Fertility: Between Regressions

Table 3 reports the results of the OLS estimation of equation (8).[35] Columns (1)-(3) of table 3 uses the average values of dependent and independent variables over 1985−2000 and show that the first two indicators of the epidemic, namely, AIDS and HIV, are positively significant at 1 percent and at 5 percent level respectively, whereas the other indicator, i.e., the HIV-EPP is not statistically significant.[36] Female schooling measured as secondary school enrollment is negative and significant at 1 percent level, while GDP per capita is insignificant, a result which is probably due to the high correlation between GDP per capita and female schooling.[37] Another important control is infant and child mortality, which is positive and significant at 1 percent level. This control also partly accounts for the variation in fertility rates due to mortality risk from other competing diseases such as Malaria.

Figure 6 shows the partial correlation plot for the regression shown in column (1), hence the slope of the solid blue line is 0.14. If I omit Congo, Rep. the coefficient goes down to 0.10 but stays statistically significant at 5 percent level as shown by the dashed red line. To interpret the coefficient on HIV/AIDS, I perform the following thought experiment: imagine going from the country with the lowest level of HIV/AIDS (Senegal) to the country with the highest level of HIV/AIDS (Zambia). This predicts an increase of 0.6 children per woman.

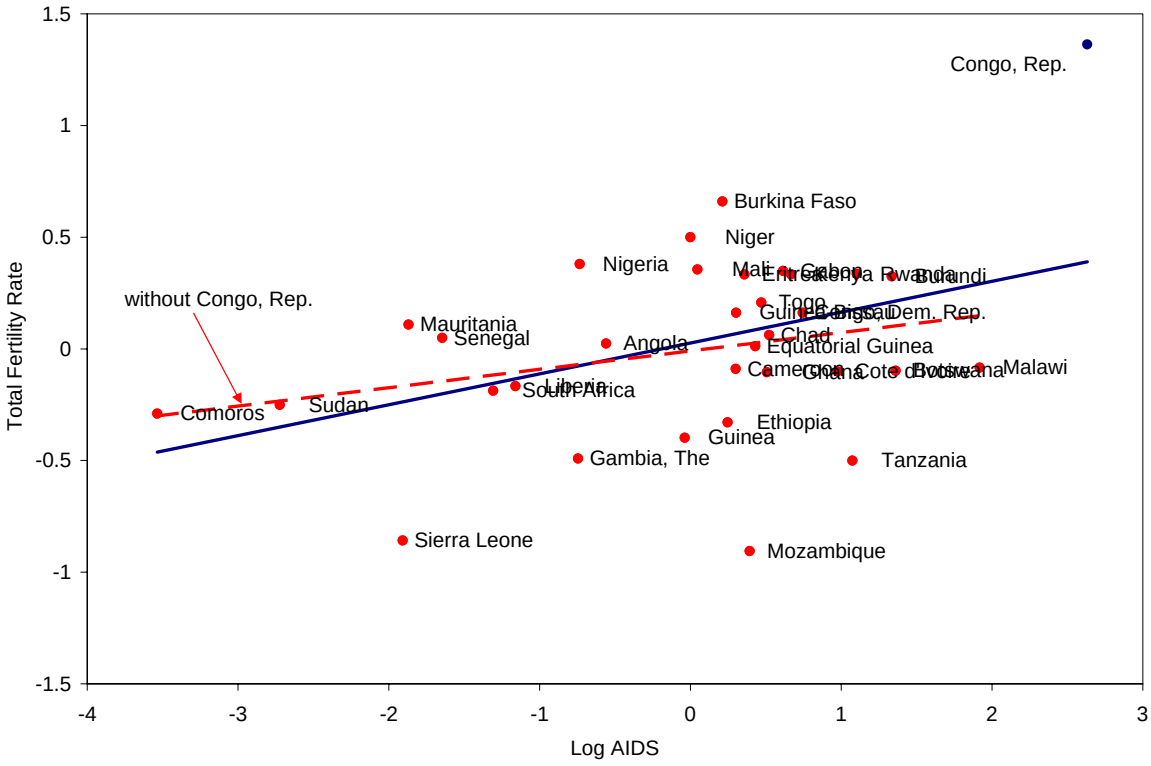
Columns (4)-(6) shows similar regressions using data on fertility from DHS. I also used desired fertility from DHS obtaining very similar results and hence I do not report them but they are available upon request. The point estimates for HIV and AIDS are larger and significant at 1 percent level for fertility rate. The 0.20 coefficient predicts an increase of

[35] I match the years of TFR data, i.e., 1985, 1987, 1990, 1992, 1995, 1997, 2000 to that of HIV/AIDS indicators before averaging.

[36] Using both indicators AIDS and HIV in a horse race, leads a positive significant coefficient on AIDS (0.23 with a standard error of 0.09) and a negative insignificant one on HIV.

[37] Using other measures of female schooling yield similar results.

Figure 6: Partial Correlation Plot for AIDS and Fertility



0.9 children per woman if we go from the country with the lowest level of HIV/AIDS to the country with the highest level of HIV/AIDS. The coefficient estimates for HIV-EPP are insignificant as before and hence not reported. I also use Oster (2006) HIV estimates (not shown due to space limitations) that deliver a positive but borderline significant coefficient. However there is only 9 countries in this estimation and hence the large standard errors are not surprising.

Finally column (6) uses data on perceptions about the epidemic instead of the actual prevalence rates.[38] The data on perceptions are much more relevant for the purpose of this paper, however the availability is only for 22 countries given the fact that this question has only been asked in the most recent surveys. Women who know someone who died of AIDS, are the ones who should react most by changing their fertility behavior. As shown this is indeed the case. The data on the percent female who know someone personally who has the virus that causes AIDS or has died of AIDS are from DHS and averaged according to the available survey years. In spite of the limited number of countries there is a strong positive association between perceptions about the epidemic and the fertility behavior as also shown in figure 7.[39] The variable "know someone who died of AIDS" can by itself explain 20 percent of the cross-country variation in the fertility behavior.[40] The estimated OLS coefficient of 0.02 implies that going from a country where 17 percent of people know someone who died of AIDS to a country where 90 percent of people know someone who died of AIDS predicts an increase of 1.2 children. The actual average decline in the total fertility rate from 1985 to 2004 is 1.12 children, hence these estimates overall are economically significant in the sense that HIV/AIDS can cause a stall in the demographic transition.

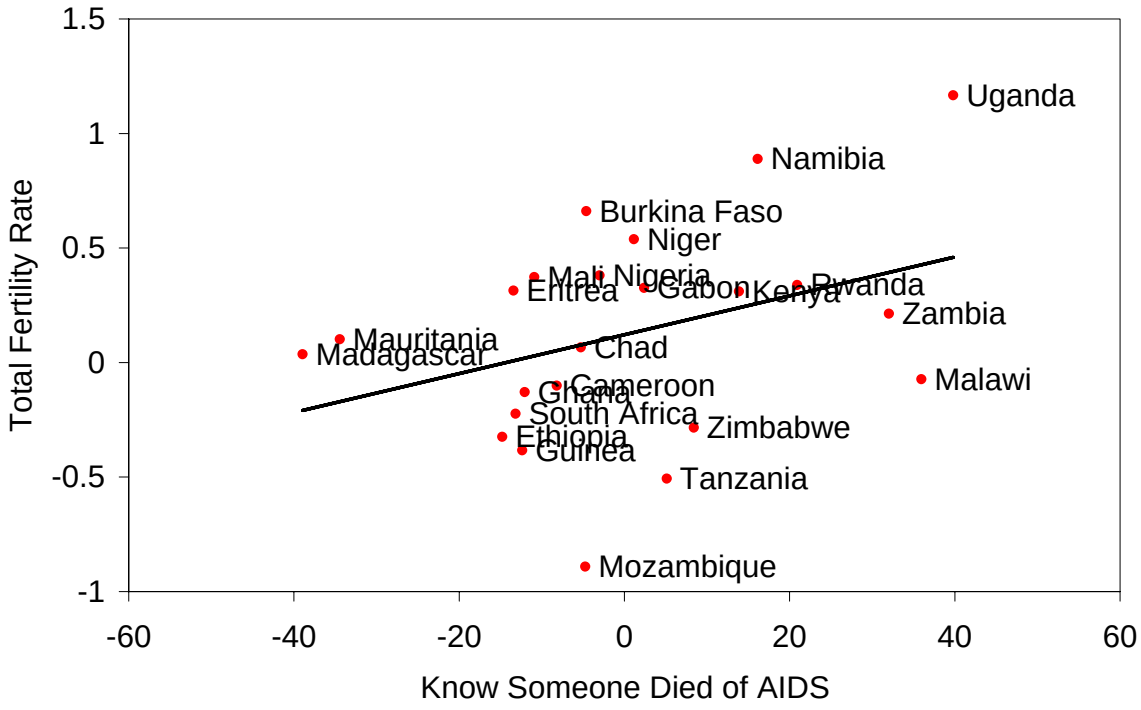
The indicators of HIV/AIDS are used in logs following Oster (2007). She argues that in her first stage estimations, where she regresses HIV on distance, log HIV prevalence on

[38]Francis (2006) shows that having a relative with AIDS changed the sexual behavior, desire and the self-reported identity of homosexual man in the U.S.

[39]I have also tried interacting the perception variables with the actual prevalence rates. However due to the high correlation between the HIV/AIDS prevalence rates and the perception variables and also due to the fact that I have limited number of countries the results of those interaction regressions are weaker.

[40]This is the partial R^2.

Figure 7: Partial Correlation Plot for Perceptions of the Epidemic and Fertility

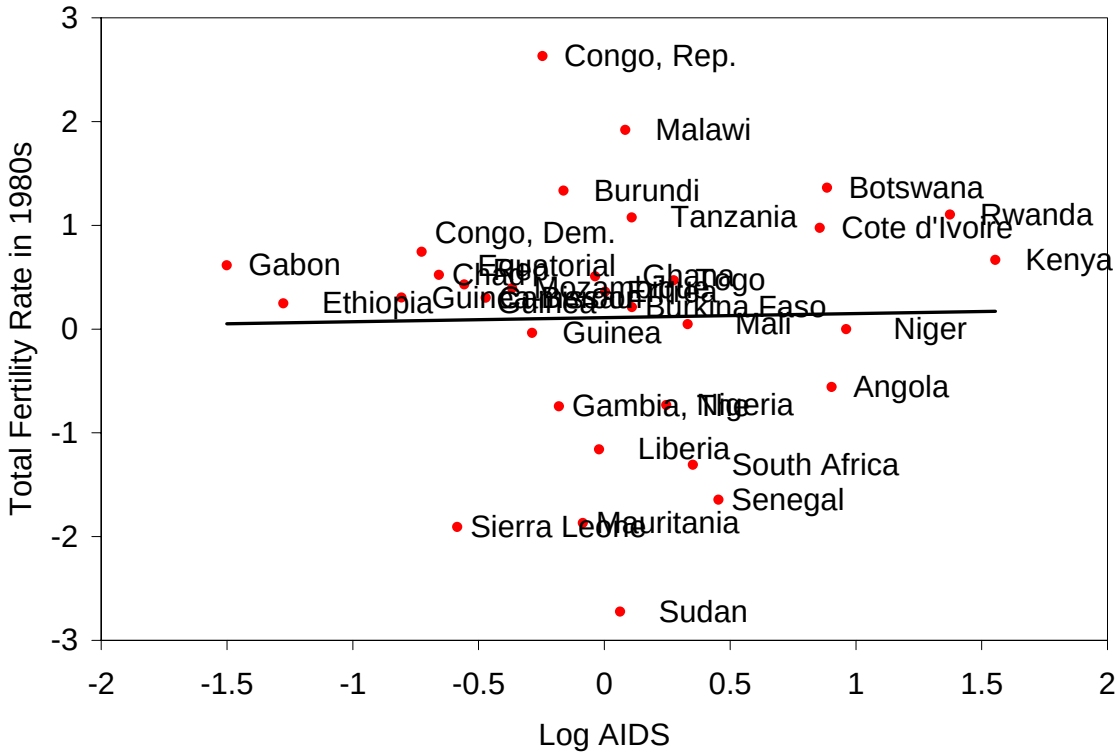


linear distance provides the best fit (most linear) as shown by simulations. I run similar first stage regressions, where log HIV provides a better fit. Hence the non-IV OLS is also used in this functional form. Although using the log of HIV/AIDS makes the quantitative interpretation harder, it has also several econometric advantages such as dampening the outliers and making the estimated coefficient immune to the scale effect due to underreporting, assuming underreporting is similar across countries. There might be a concern in using the log form though since log specification in principle compares the countries that have any AIDS to those that don't. I would argue that this is not a serious concern in the case of Africa. First of all due to averaging over time, I do not have any zeros in HIV/AIDS; the only zeros for the initial years of epidemic for few countries are averaged out. Second of all, the sample I am using are composed of countries that are classified as "generalized epidemic countries" with the exception of Comoros, Madagascar, Mauritania, and Sudan. The results are robust to excluding these four countries. The results are also robust to, even stronger, using the non-logged proxies for HIV/AIDS and available from the author upon request.

Given the fact that between regressions fail to control country fixed effects I proceed as follows. First I will undertake a falsification exercise and second I will run a panel regression. Table 4 represents the results from a regression, where I regress fertility rates from 1980s, on the current HIV/AIDS, averaged over 1995–2004 and the other controls. There is no statistically significant relationship between current HIV/AIDS and fertility in 1980s as shown in columns (1) and (2) and further in figure 8.[41] This exercise suggests that time invariant unobserved country heterogeneity is not driving the results. I also run a regression similar to a diff-in-diff specification such as I regress change in fertility from 1990 to 2004 on the change in HIV/AIDS from 1990 to 2004, obtaining similar results to that of between regressions.

[41]Fertility Rates in 1980s is the average of rates in 1980, 1982, 1987.

Figure 8: Falsification Plot for AIDS and Fertility



4.4 Country Level HIV and Fertility: Within Regressions

Table 5 reports results of the OLS estimation of equation (9) using country level data.[42] Standard errors are clustered at the country level to deal with the possible serial correlation among residuals. I also include regional dummies for East, West and Southern Africa. Results were similar. For each of the three indicator of HIV/AIDS both pooled regression results with a common time trend (that captures the declining trend of fertility in the absence of HIV/AIDS) and "within" regression results with both country and time fixed effects are shown in columns (1)-(6).[43] AIDS is positively significant both in the pooled and in the fixed effects regressions, HIV is only significant in the pooled regressions, and as before HIV-EPP is not significant. These specifications are similar to Young (2005b) specifications except they use country level TFR as the dependent variable instead of individual level fertility pooled from different countries as in Young (2005b). Restricting the sample to that of Young (2005b) of 27 countries does not change the results (not shown due to space considerations). All other control variables yield similar results as before.

The use of within country time variation as the main identifier for the effect of the epidemic is of suspect given the questionable quality of the time variation in these data as argued before. Indeed figures 9, 10, and 11 show time series path of HIV data from three countries, where there is so much noise.

Following Young (2005b) I used country-specific time trends as shown in the last 3 columns of table 5. The first two indicators gave insignificant results, however the third indicator, that is HIV-EPP yields a negative significant result as in Young (2005b).[44] Recall that this variable never turned out to be statistically significant up until this specification. Given the nature of these *projected* HIV-EPP rates, I would argue that the analysis of the effect of the epidemic on fertility should not focus only on these data used in a within regres-

[42] I also perform Weighted Least Squares (WLS) panel regressions; where all observations are weighted in the second step with the inverse of the estimated standard deviations from the first step. Weighting by country's population or log population yields similar results.

[43] I also experimented with a common non-linear quadratic and cubic trend obtaining similar results.

[44] Recall that Young (2005b) uses solely this indicator for HIV.

Figure 9: Ethopia HIV

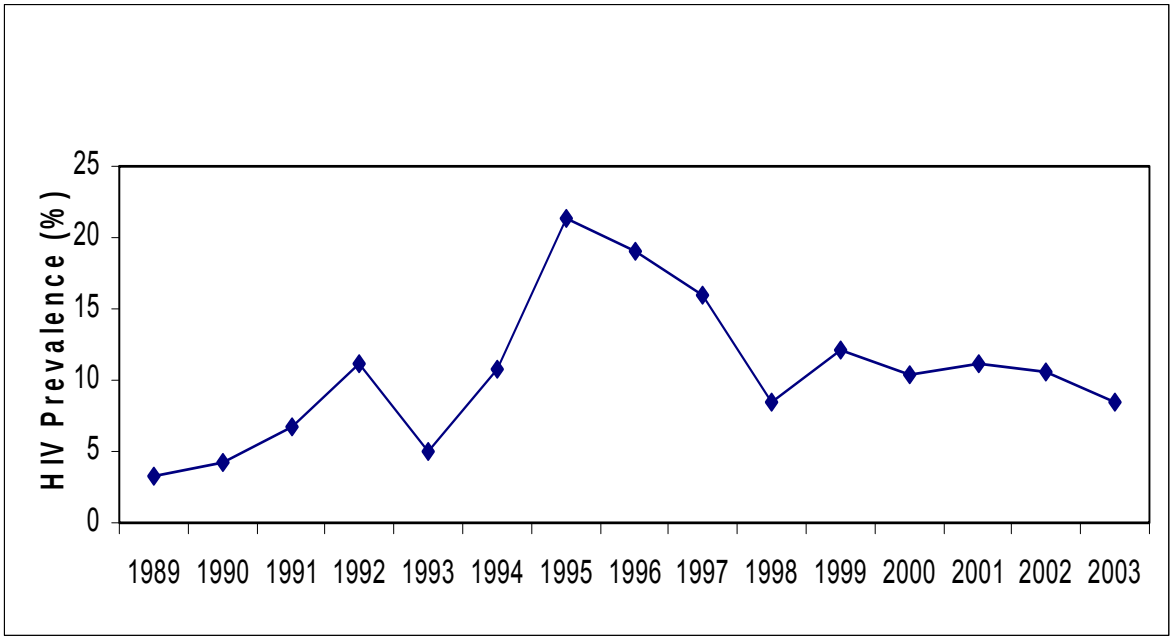


Figure 10: Kenya HIV

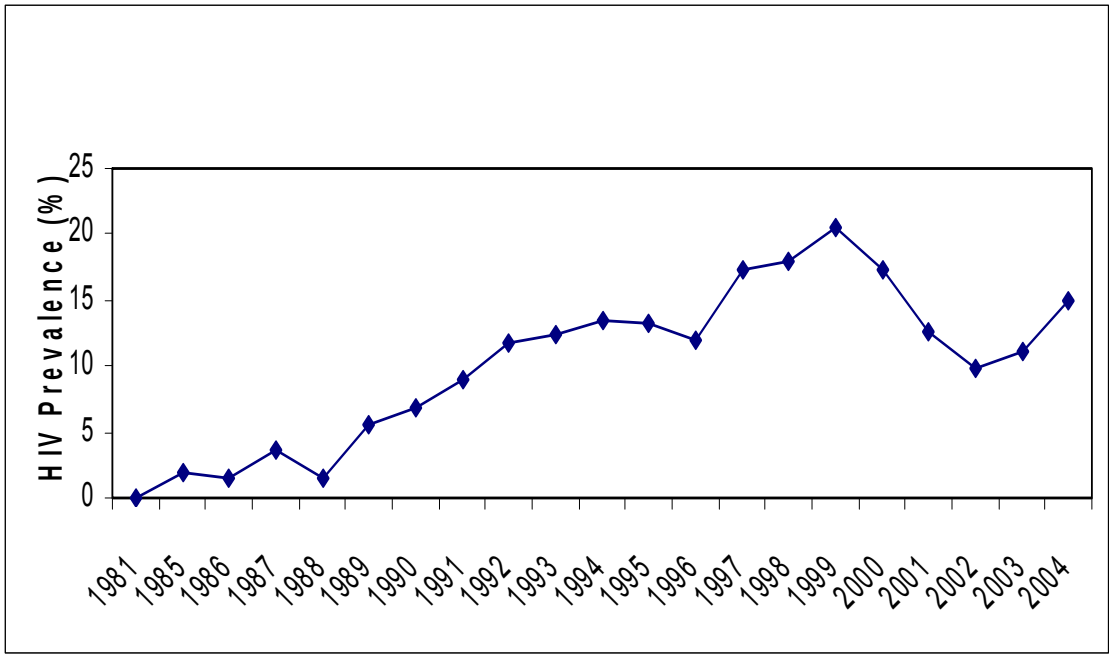
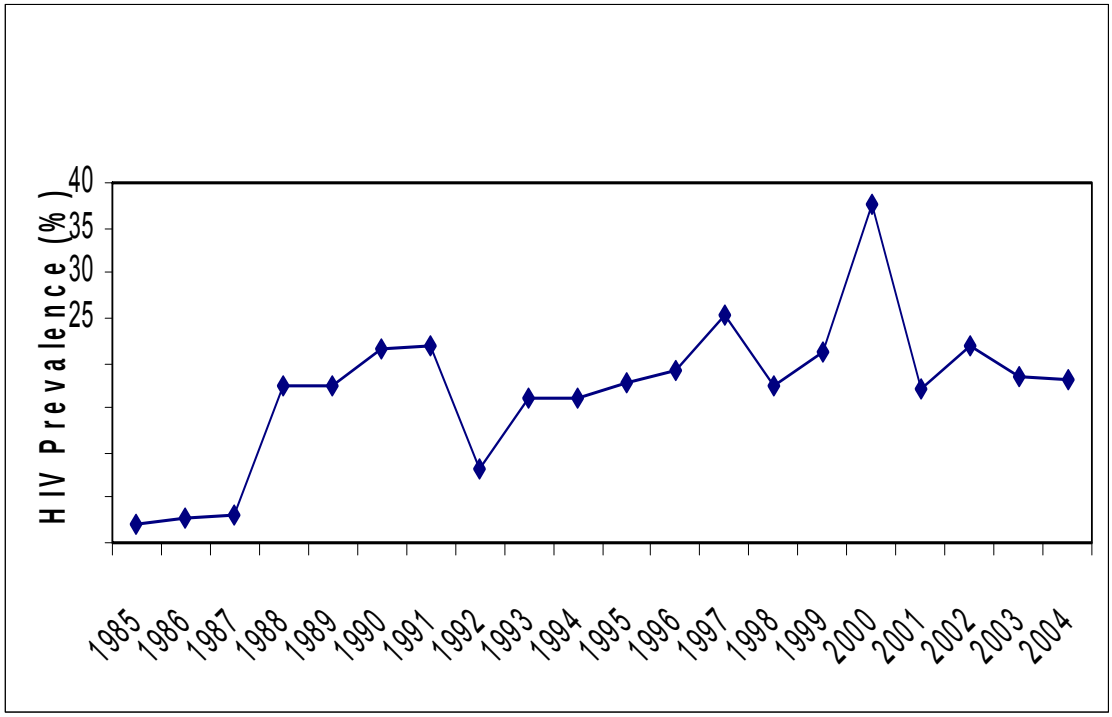


Figure 11: Malawi HIV



sion framework with country specific trends with limited degrees of freedom. Although EPP model fits a somewhat flexible curve to a not so long time series, the modeling is still an issue of concern given the dynamic nature of the epidemic. In addition, the input data is still the same HIV data that comes from the U.S. Census Bureau, HIV Surveillance Database (2005) and UNAIDS, where the representativeness and the quality of the time series variation are both questionable as argued above. If the time-series trend in HIV/AIDS is similar to the country specific trend including these trends should not add much in principle.[45]

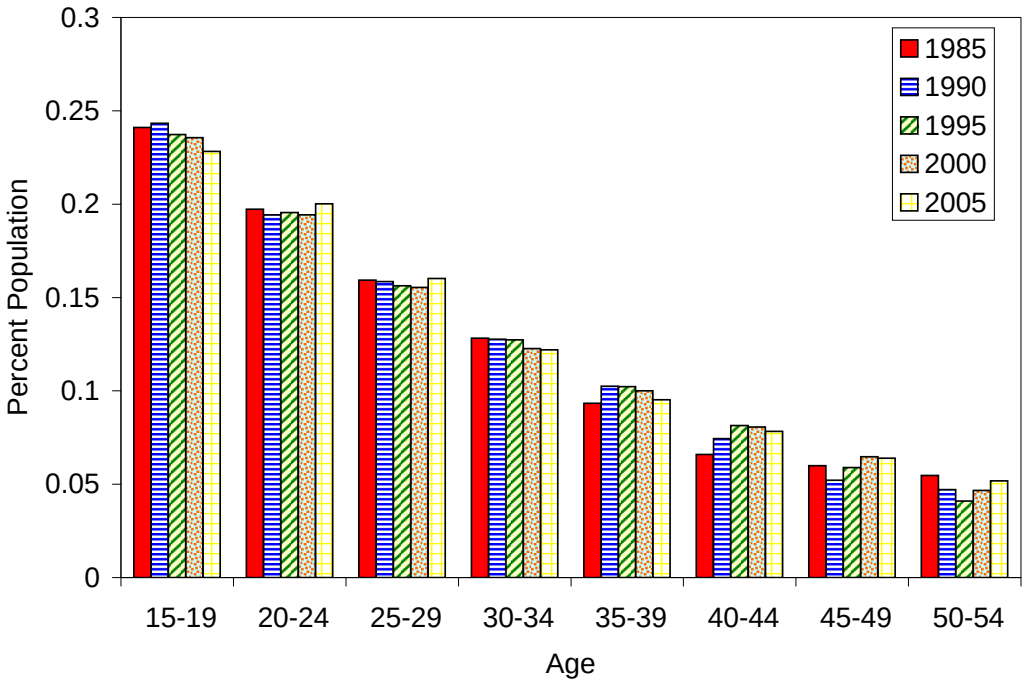
4.5 Alternative Explanations and Robustness

One story that comes into mind, which might lead to a positive association between HIV and fertility, is a shift in the population age distribution. If older women are dying because of the epidemic, the total fertility rate will increase simply due to the fact that younger women have more children. Figure 12 takes a stab at this by looking at the data for Kenya—a high prevalence country. The data show that, in spite of the high mortality, the population age distribution has not shifted much. The population age distribution for Africa as a whole also tells a similar story. Hence it is unlikely that the positive effect we have found so far is due to the shifts in the population age distribution.

Another alternative story rests on the question that what is driving the results in terms of sub-groups. Given the fact that total fertility rates is a summation over the age-specific fertility rates, the following scenario is also plausible. People might marry early, settle down sooner and start having children, due to the epidemic. This would lead to a shift in the timing of fertility, where people have children at younger ages. Thus even if each woman were to have no more children with HIV/AIDS than without HIV/AIDS, one might observe a gain in measured total fertility rates because in a given period two generations of women would be bearing children, a previous generation whose schedule of childbearing had not

[45]Given the limited survey years the coefficient in Young (2005b) is identified of off 15 countries, 3 years at most. Also in some of the Young (2005b) specifications with country specific trends HIV turned out to be insignificant.

Figure 12: Population age Distribution in Kenya, 1985–2005



been affected by AIDS and a new generation who had decided to have children earlier.[46]

A first step would be looking at age-specific fertility rates from Kenya as shown in figure 13. These rates show that it does not seem to be the case that age-specific fertility rates are changing disproportionately, if anything they all have increased in the last survey consistent with the increase in the total fertility rate in Kenya, shown before in figure 1. One must caution though since there is also ample evidence—as reviewed in section 2.3—from clinic and cohort based studies that HIV positive woman have lower fertility and childbearing odds. Hence age-specific fertility rates can also be lower for high HIV countries as a result of the epidemic. Indeed, for Kenya the decline in the age-specific fertility rates from 1993 to 1998 survey can easily be due to this biological effect of the epidemic. The point I am trying to make is that it seems not to be the case, at least for now, that there is a disproportionate change in age-specific fertility rates.

For robustness analysis, I have investigated the effect of other control variables such as the use of contraception. The contraception data (defined as any form but mostly constitutes condom use) are available for 34 DHS survey countries for 3 years at most from World Bank and DHS. Oster (2007) argues that reliable data on condom use is hard to get. The data on condom use does not capture consistency of use; the DHS question is such that it involves reporting condom use during the last sexual encounter. Oster (2007) finds no statistically significant effects of HIV on condom usage. Juhn, Kalemli-Ozcan, and Turan (2007) also finds a statistically insignificant effect of condom use on the individual's fertility behavior.

Figure 14 presents the data on contraception prevalence from WB and DHS. Each countries survey year is on or around the dates shown on the x-axis. After an initial increase the use of contraception came to a halt in the latest surveys, as shown for Kenya and Cameroon. For some countries it has been constant throughout such as South Africa and Eritrea. And for some others, the use of contraception seems to have decreased in the latest years such as Chad, Nigeria, and Rwanda. There are yet other Africa countries where the use of con-

[46]In response to HIV/AIDS, members of a couple might have less unprotected extramarital sex, but more unprotected marital sex, which will raise fertility.

Figure 13: Age-Specific Fertility in Kenya, 1989, 1993, 1998, 2003

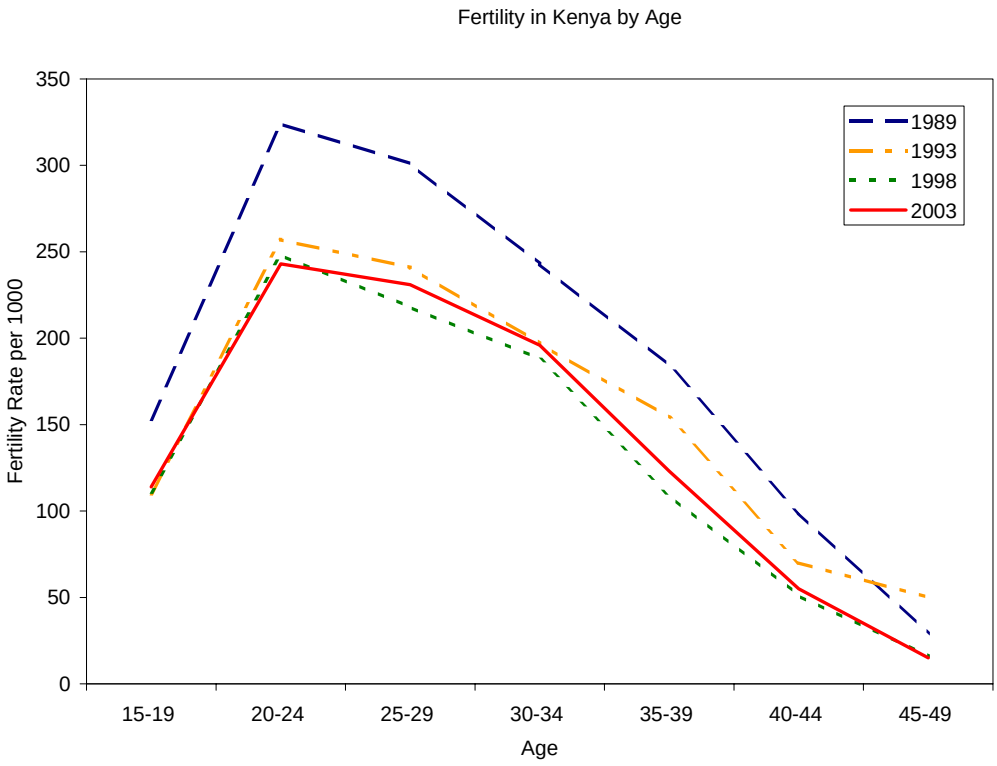
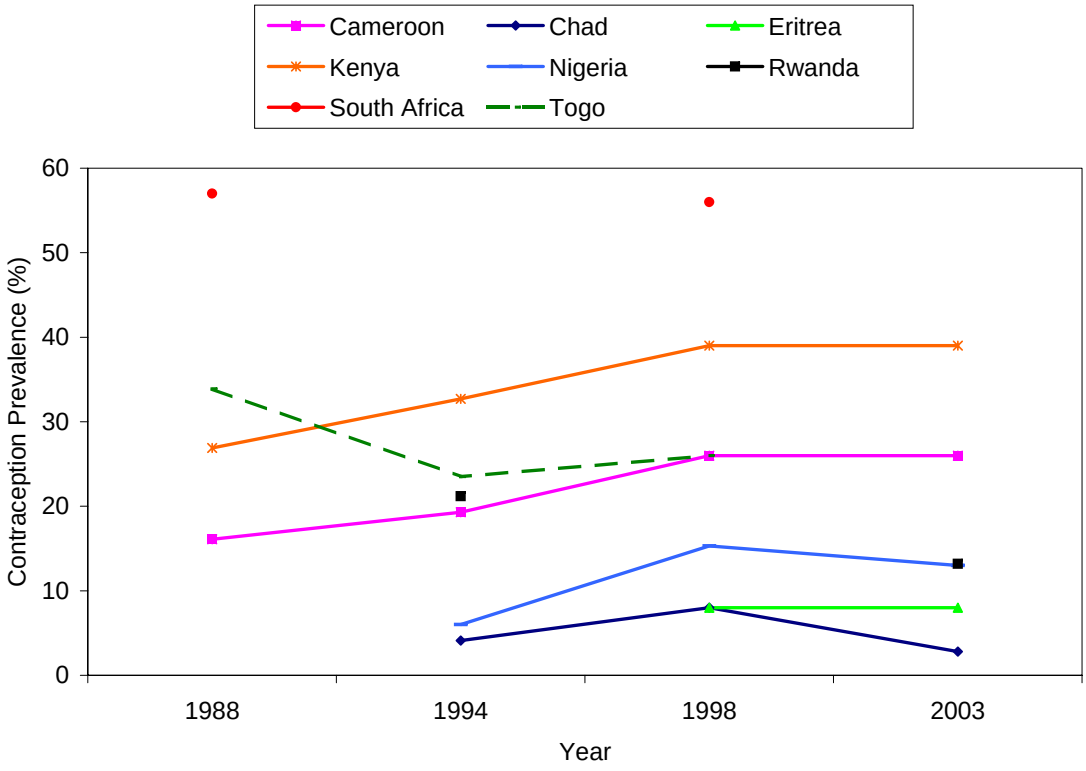


Figure 14: Contraception Prevalence in Africa



traception is on a steady rise. The bottomline is that it is hard to conclude sexual behavior and/or fertility is changing in any direction based on the available contraception data since there seems to be no definite pattern.

Including the use of contraception in the between regressions does not change the main result: HIV/AIDS prevalence is still positive and significant. Contraception use is negative and significant at 10 percent level. Another interesting issue about the use of contraception as shown in figure 14 is that it has been very low, below 50 percent for most countries. Researchers have argued that family planning programs explain only 10-40 percent of the

decline in developing countries and the rest of the decline is explained by the changes in desired fertility, i.e., the number of children families want to have. Weil (2003) and Pritchett (1994), who shows that the relation between actual fertility and desired fertility among a cross-section of developing countries are very close and thus there is little scope for reducing fertility in many countries through better provision of contraceptives.[47]

I have also included the population age structure, male schooling, urbanization, and regional dummies for East, West and Southern Africa, as controls, and they mostly turn out to be insignificant. To add marital status as a control, I use data on the percent female (15-49) who are married at least 1 year prior to the survey year and still married from DHS, which is available for 30 countries. The variable percent married enters positively but not significantly.

4.6 IV Regressions

As discussed before to deal with the issue of endogeneity I will pursue an instrumental variable strategy. Following Oster (2007), I instrument HIV/AIDS by the distance to the origin of the epidemic, which is Democratic Republic of Congo. Given the time invariant nature of the instrument I will focus on the between estimates. For the instrument to be valid, it must be correlated with the HIV/AIDS but uncorrelated with the fertility rate, except through variable of interest that is included in the equation explaining fertility. The most obvious way in which distance to the origin of the epidemic might systematically affect the fertility rate—other than via HIV/AIDS— is through its correlation with geographic and/or socioeconomic variables. Hence these factors will be controlled in various ways as detailed out below.

The first stage regression will be of the form:

$$HIV/AIDS_i = \varsigma + \chi Distance_i + \mathbf{X}_i'\pi + \xi, \tag{10}$$

[47]The correlation between desired and actual fertility in my sample is 89 percent.

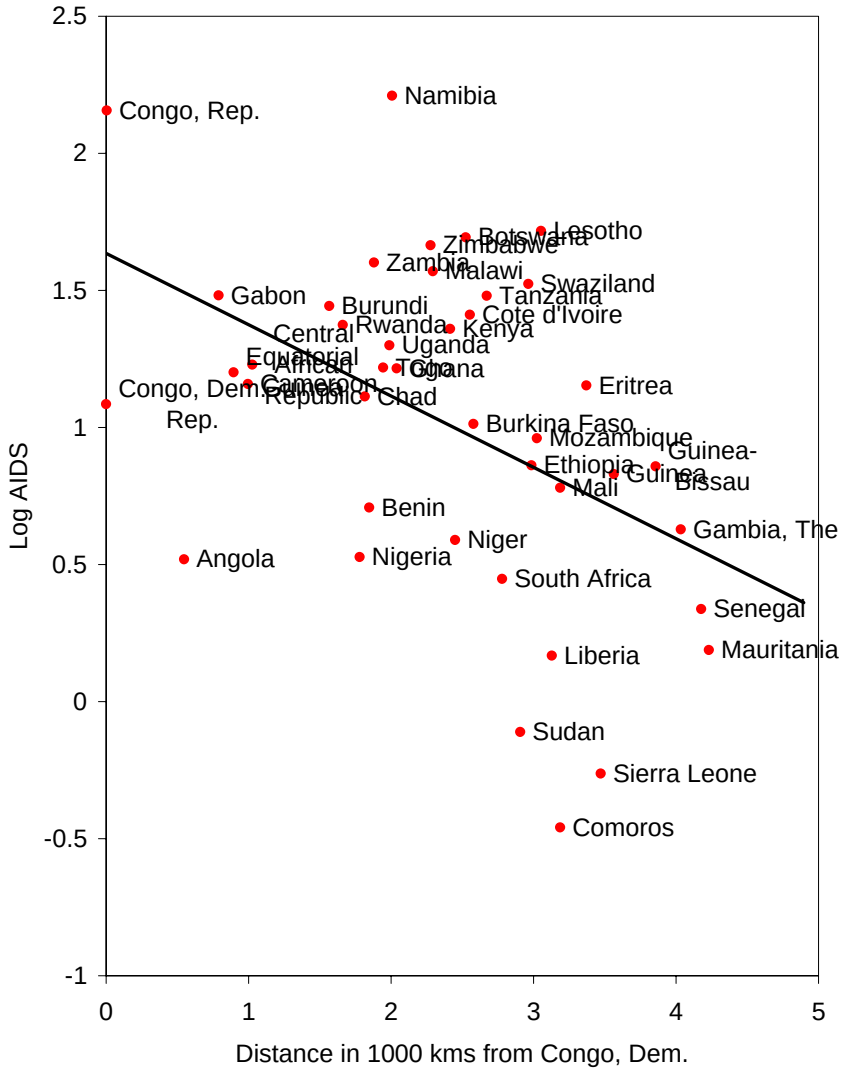
where distance is measured as the distance from the capital city of each country to the capital city of the Democratic Republic of Congo, which is measured as the distance between the center of the capital cities. **X** matrix represents the controls that are used in the second stage regression and some other geographic controls. Figure 15 shows the relationship between log HIV prevalence and distance to the center of the Democratic Republic of the Congo. The relationship is clearly downward sloping and approximately linear as in Oster (2007).

Table 6 estimates the relationship between HIV/AIDS and distance adding the controls instead of the simple scatter plot shown in figure 15. Column (1) and column (3) show the relationship between distance and AIDS and HIV without any controls, respectively. However if there are regional differences among fertility rates and this is correlated with distance to Congo, then this will bias the results. Hence I include regional dummy variables for East, West and Southern Africa in column (2) and (4). Columns (2) and (4) also include the other socioeconomic controls such as female schooling, GDP per capita, and infant mortality, which are also controls in the second stage. The coefficient on distance stays negative significant and similar across all these specifications.[48]

Table 7 reports the 2SLS estimates for the effect of HIV/AIDS on fertility. The coefficients in columns (1) and (2) is positive and significant at 1 percent level. The coefficient for the projected HIV-EPP is insignificant as before and hence not reported. Notice that the coefficients for HIV and AIDS are higher than the corresponding OLS coefficients. This is probably due to the fact that the IV regression corrects for the measurement error which leads to attenuation bias in the OLS regression. To interpret the coefficient of 0.28, I perform the same thought experiment as before: going from the country with the lowest level of AIDS prevalence to the country with the highest level of AIDS prevalence predicts an increase of 1.1 children per woman. Doing the same exercise with HIV predicts and increase of 1.2 children. These effects are higher than OLS but consistent with the ones I have obtained from the perceptions data.

[48]Oster's (2007) first stage estimates vary from –0.3 to –0.7 depending on the controls and country fixed effects.

Figure 15: Distance to the Center of the Epidemic and HIV/AIDS



It is unlikely that distance drives fertility directly. There might still be a concern that distance might be correlated with fertility in the pre-HIV/AIDS period.[49] Table 8 shows the falsification exercise for the exclusion restriction. Column (1) regresses fertility in 1980s on distance, obtaining an insignificant coefficient. The same result follows for fertility in 1970s. Columns (2)-(4) checks whether or not distance is directly correlated with socioeconomic factors such as GDP per capita, infant mortality, and child mortality. Oster (2007) further shows that there is no correlation between distance and malaria and life expectancy. Overall these results suggest that distance seems to be unrelated to pre-epidemic fertility and to various socioeconomic factors. These exercises provide confidence in the validity of instrument and it seems plausible to argue that distance to the origin of the epidemic will be unrelated to fertility.

4.7 Regional Evidence

This section provides additional supporting evidence for the positive effect of the epidemic on fertility using regional data on fertility and HIV from the African countries. I have data on 32 regions from 12 countries. These are the regions with overlapping data on regional total fertility rates and HIV prevalence rates. Each country's survey year falls between 1998–2004. If there is more than 1 survey year during this period, than the data on the total fertility rate are averaged. I regress the regional total fertility rates on the logarithm of regional HIV prevalence rates among pregnant women averaged over 1990–1995, including country dummies. Unfortunately, the other controls are not available at the regional level.

Table 9 shows the results of the OLS regressions. Both columns show that results are positive and significant at 1 percent level. To deal with the potential serial correlation across residuals given the regional data, I cluster at the country level, which raises the standard errors as shown in column (2). Though the results are still significant at 1 percent level. I also tried a "Weighted Least Sqaures (WLS)" specification, where in order to limit the influence

[49] Oster (2007) deals with a similar concern in the case of sexual behavior and she checks this by looking at the relation between old people's sexual behavior and distance.

of small regions, I weighted by the population and also alternatively by the logarithm of regional population from DHS, averaged over the survey years. Results were similar and hence not reported.

The size of the estimate and the quantitative impact is similar to that found in the country regressions. Going from the region with the lowest level of HIV in this sample (Western Cape Province of South Africa) to the region with the highest level of HIV (Manicaland region of Zimbabwe), implies 1.1 more births per woman. The results requires some caution since there might still be heterogeneity in spite of the country effects and clustering the standard errors at the country level, due to variation in proximity to road networks and urban-rural differences.[50]

5 Comparison to Young's Results

The careful work of Alwyn Young shows that there is a negative effect of country-wide HIV prevalence on individual fertility in South Africa (Young, 2005) and this conclusion can be generalized to other 27 African countries (Young, 2005b). In this section I replicate both papers of Young (2005, 2005b) using the country level HIV rates and individual level fertility rates both for South Africa and for 27 African countries using data from Demographic Health Surveys. I show that the results of the current paper are consistent with his findings.

5.0.1 South Africa: Young (2005)

Table 10 shows the replication results of Young (2005). A Poisson count model is used. Wage index is the estimated wage for each individual depending on own education, where wage regressions are done using South Africa 1995 October Household Survey (OHS) following Young's methodology.[51]

[50] I also run IV regressions for a smaller sub-sample. In spite of a strong first stage the second stage regressions gave statistically insignificant results.

[51] Using both South Africa DHS (1998) and OHS (1995) wage is estimated as a function of age, sex, and education. Individual education levels reported in OHS and DHS are converted to standardized years of education, as shown in appendix. These educational categories are then used to construct wage index as

The fertility data are from South Africa 1998 Demographic Health Survey. Following Young (2005), only women who are 25 or older are used in the study, and there are 7276 of them in the data set. The panel is constructed using each womans birth histories since age 12, and it includes the period between 1961–1998. Retrorespective fertility is the number of pregnancies of each woman in each year, including that were lost before term or resulted in stillbirths. Note that this is not a measure of completed fertility since different women are observed at different points in time.

Young's (2005) identification comes from variation in HIV exposure by age. He controls for the effects of age and cohort using linear (and sometimes polynomial) trends in birth year and age, all of which will control for a smooth trend.[52] Young (2005) uses the historical antenatal clinic sero-prevalence rate for each woman's age group at that time of their life from South Africa Department of Health summary reports. These rates are available since 1990. Hence, he assumes a 0 rate of HIV for the part of the panel before 1990. In spite of many emails these data could not be obtained from South Africa Department of Health. The same source is cited in the U.S. Census Bureau's HIV/AIDS Surveillance Database (2005) and hence I use South Africa HIV prevalence rates among pregnant women by age group from the U.S. Census Bureau, HIV Surveillance Database (2005). I match Young (2005) number of observations exactly, i.e., 171206. Young (2005) finds the coefficient on HIV as –1.63.[53] I find a coefficient of –1.36 as shown in column (1), which is statistically significant. By reading the data off of the printed South Africa Department of Health Report (2007), I find a coefficient of –1.48. Following Young's methodology, column (1) assumes a HIV prevalence rate of 0 before 1990. In column (2) only part of the panel that is after 1990 is used in regressions. Columns (3) and (4) repeat the same exercise by using OLS estimation.

in Young (2005). Wage index is calculated as: $B_E * E_i + B_{E_2} * E_i^2$, where B_E and B_{E^2} are the return to education coefficients coming from the regression of wages on age, education and their squares.

[52]In a country like South Africa one can imagine the existence of more complicated trends due to the abolition of apartheid, which is a discrete change.

[53]Young (2005b) notes that if he had used the entire South African sample of women aged 15 and above instead of 25 and above, the estimated coefficient on HIV would have been only –0.86 instead of –1.63. 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.

The coefficient of interest changes sign and becomes positive significant when post-1990 part of the panel used.

The explanation for this finding is straightforward. Figure 16 plots the mean residuals (blue circle line) from a regression of fertility on other control variables except for HIV against the mean residuals (red triangle line) from a regression of HIV on other control variables for every year; an exercise that yields the same OLS coefficient as in column (3). Assuming zeros for HIV for every women before 1990 creates an artificial trend in the residuals as opposed to the case where we use only the part of the panel after 1990, creating a negative association between HIV and fertility.[54]

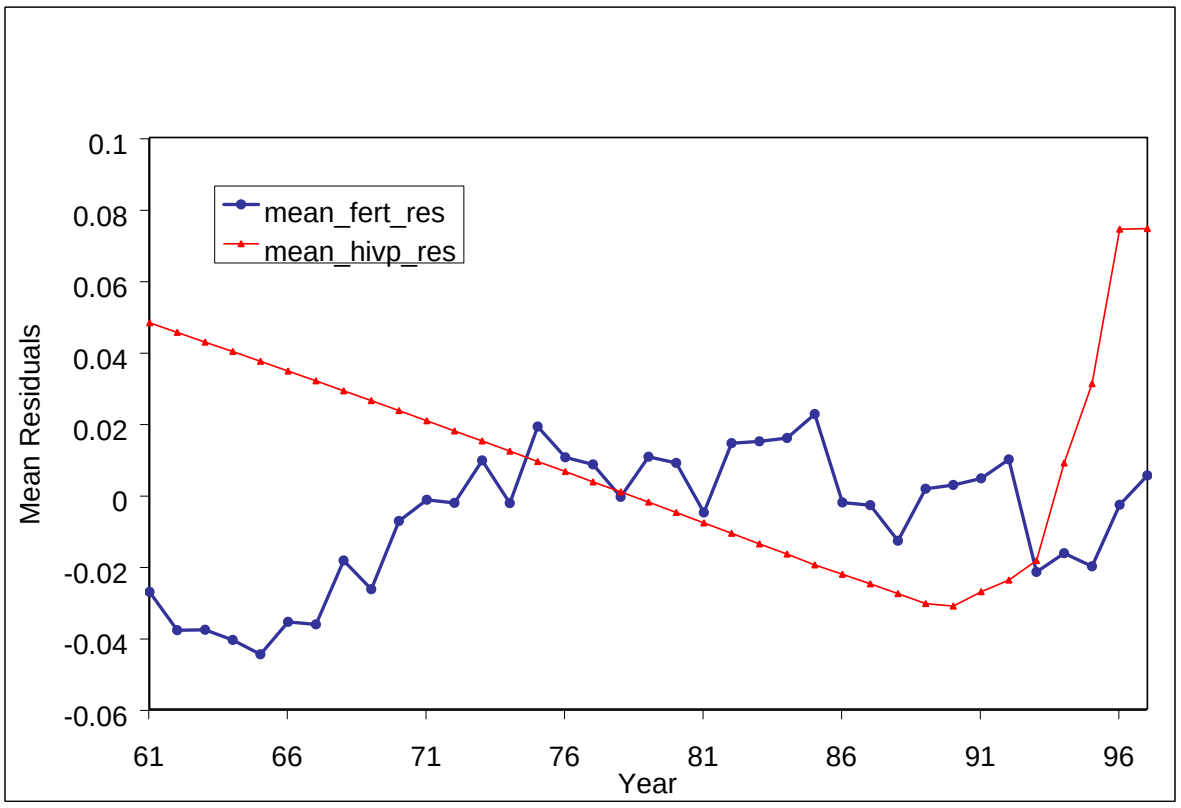
5.0.2 Other African Countries: Young (2005b)

In a similar fashion, Young (2005b) extends the results of Young (2005) to 27 African countries by using DHS surveys. He uses the country level HIV-EPP projections and exploits within country variation over time as I did in table 5. However for this 27 country sample only 2 countries have four surveys, 10 countries have three surveys, 8 countries have two surveys, and the remaining 7 countries have only one survey. Thus the identification rests on limited observations. Table 11 shows the replication of his results. Every regression has controls for age and education and also country and year effects. Columns (1)-(4) undertake a Poisson estimation of past year births following Young (2005b) and columns (5)-(8) perform an OLS exercise.[55] The difference between column (1) and (2) and similarly the other column pairs is the inclusion of control variables, which are marital status (never, currently or formerly married), urban/rural location, the number and square of born and living children, and the presence of a radio, television, refrigerator or bicycle (each entered separately) in the household.

[54]This can simply be expressed as follows: Figure 16 plots $residuals_{TFR}$ against $residuals_{HIV}$, where $residuals_{TFR}$ come from a regression of fertility on other control variables and hence $residuals_{TFR} = TFR - \hat{a}X$. $residuals_{HIV}$ come from a regression of HIV on other control variables and hence $residual_{HIV} = HIV - \hat{b}X$. Given $HIV = 0$ before 1990 and X is increasing in years $residual_{HIV}$ has a downward trend.

[55]Using desired fertility measured as the ideal number of children gives similar results and available upon request.

Figure 16: Partial Correlation Plot by Years for HIV and Fertility



Columns (1), (2) replicates regressions of Young (2005b) by clustering at individual level (on case-id). Not clustering at all give very close results. Columns (1) is an exact match to Young (2005b) and column (2) is close. However, when I cluster at country level in columns (3) and (4), I get insignificant results since clustering at country level leads to large standard errors. The same story goes for the OLS estimation as shown in columns (4)-(8). Given the fact that the treatment is at the country-time level, it is preferable to cluster either at country-time or at country, to deal with possible serial correlation.[56]

Peterson (2007) and Bertrand at el. (2003) argue, it is common practice that researchers does not adjust the standard errors for possible dependence in residuals in the panel data sets. Peterson (2007) reports that 42 percent of the papers published in the last five years in finance using panel data by firm and by time does not adjust standard errors at all. He shows that the true standard error is 11 times the estimated and 81 percent of the time t-statistics are falsely significant at 1 percent. Bertrand et al. (2004) have drawn attention to robust standard error estimation in the context of a special fixed effect model, that is "Differences-in-Differences (DD)," where they show 65 percent of the time, there is false significance because of non-clustering. Out of 92 DD papers only 36 deal with the issue. Peterson (2007) and Bertrand et al. (2004) both show using simulations that only clustered standard errors are unbiased as they account for the residual dependence created by the state or firm or country effect. None of the other corrections can alleviate the serial correlation.[57]

A priori one might think that inclusion of country fixed effects would address some of the heterogeneity by country. Indeed Peterson (2007) says 29 percent of the papers prefer this method, however fixed effects will not be addressing the problem fully, as we see here. In the

[56]Since the autocorrelations can be positive or negative it is possible for the non-clustered standard error to under or over-estimate the true standard error. In the case of HIV, positively serially correlated residuals lead to underestimated standard errors and hence false significance.

[57]Peterson (2007) shows this for the standard OLS regression but he reports that his results generalizes to non-linear models too. Bertrand et al. (2004) focuses on a DD model such as; $Y_{ist} = A_s + B_t + cX_{ist} + \beta I_{st} + \epsilon_{ist}$, for individual i, state s, and time t. They also show simple parametric corrections, such as fitting an AR1 process for the error structure, or non parametric corrections, such as block bootstrap, only works with large number of states. They show that clustering at state level not just at state-year cell is the best solution.

case where there are both country and time fixed effects the best practice is to cluster at both levels or if the number of clusters is small in one dimension, like the time dimension here, than use a fixed effect for that dimension and cluster on the other dimension, where more clusters are available.[58] Kezdi (2004) shows clustered standard errors can be too large in a fixed effects model but he also shows only clustered standard errors are unbiased irrespective of having a country effect, as also shown by Peterson (2007).[59] Kezdi (2004) shows that the general robust standard error estimator known as the cluster estimator is not only consistent in general but it behaves well in finite samples.[60]

6 Conclusion

Theoretical models of demographic transition imply that fertility decreases as a result of a decline in mortality, either by increasing the returns to human capital or by reducing the precautionary demand for children. The HIV/AIDS epidemic is a positive shock to mortality that implies a reverse path. Using both macro and micro data from a panel of African countries during 1985−2000, this paper mostly shows a positive effect of the epidemic on fertility. The main results can be summarized as follows: 1) Between estimates based on country data suggest a strong positive effect of HIV/AIDS on fertility both in OLS and in IV. These estimates predict that a country with a high level of HIV/AIDS prevalence, such as Zambia, have 1 more child per woman on average compared to a country with a low level of HIV/AIDS prevalence, such as Senegal. 2) Within country estimates show mixed results due to their sensitivity to different time trends. 3) Within estimates based on individual data from South Africa suggest a positive effect of HIV/AIDS in the 1990s, whereas pooling

[58] If the number of clusters are less than 10, clustered standard errors might underestimate the true standard errors. Peterson (2007) shows that one does not need to cluster both at country and time in the presence of fixed effects for both.

[59] Peterson (2007) shows the generalization of the results for the GLS case too.

[60] His Monte Carlo simulations shows that only cluster estimator gives unbiased results even in small cross-sectional samples. He shows in a fixed effect model with short time series, serial correlation in the error process and the right hand side variables induce severe bias in conventional standard errors. Clustered estimator applied to mean-differenced data is consistent and behaves well in finite sample and it does not get biased with high T or small N.

the survey data for 27 African countries suggests a zero effect.

The quantitative effects may not seem big enough to cause a "reversal" in the fertility transition. However given the average decline of 1.2 children in the total fertility rate across Africa in the last 20 years and the mean of 5.8 children in the sample, the quantitative effects are in the ballpark for a stall in the demographic transition. Looking at the data from the latest rounds of the DHS surveys show a widespread stall in the demographic transition in Africa, which is inconsistent with declining fertility across Africa found in Young (2005b) as a result of the epidemic.

There can also be a countering effect due to the introduction of ARVs. Some recent studies have argued that AIDS crisis has not materialized in many countries and appears to have reached a plateau in others, due to partly because of the treatment. There is great optimism that the introduction of ARVs to Africa will increase people's knowledge about their HIV status and quell the epidemic. However the validity of these predictions came into question in other research. First of all introduction of ARV may lead to more risk taking.[61] Second, ARVs are generally introduced when one progresses to AIDS due to budgetary considerations in Africa, which will not help that much given the median survival time to be 5 to 9 months once you develop AIDS. Finally, only 18 percent of those who need treatment in Africa are receiving ARV.[62]

I close by stressing that divergence in results among similar studies indicates that the issue of the fertility effects of HIV/AIDS is far from settled. The relationship between fertility and the HIV/AIDS epidemic is one of the most important missing pieces in the puzzle of AIDS and development. The association between fertility and the epidemic is not easy to detect given the difficulty of separating biological and behavioral responses and the problems of the aggregate HIV data. To investigate this association using the newly available HIV data based on individual testing will be an important avenue for future research.

[61] See Kremer (1996) for a model that produces multiple equilibria for the behavior change. See Over (2007) for evidence.

[62] See Canning (2006).

Appendix

Country Level Data:

Countries: Angola, Benin**, Botswana, Burkina Faso**, Burundi*, Cameroon**, Central African Republic*, Chad**, Comoros, Congo Democratic Republic, Congo Republic, Cote D'Ivoire**, Equatorial Guinea, Eritrea**, Ethiopia*, Gabon*, Gambia, Ghana**, Guinea*, Guinea-Bissau, Kenya**, Lesotho, Liberia*, Madagascar**, Malawi**, Mali**, Mauritania, Mauritius, Mozambique**, Namibia**, Niger**, Nigeria**, Rwanda** Senegal**, Seychelles, Sierra Leone, South Africa*, Sudan, Swaziland, Tanzania**, Togo**, Uganda**, Zambia**, Zimbabwe**.

Countries with a * has at least one DHS survey, and countries with ** has more than one DHS survey.

- *AIDS*: The AIDS data come from UNAIDS/WHO, Epidemiological Fact Sheets (2003) and US Census Bureau HIV/AIDS Surveillance Database (2005). These are the number of reported AIDS cases for each country in every year and available for 44 African countries for 1985–2004. I multiply these number of reported incidents by 100,000 and divide by the country's population in each year, converting them to incidence per 100,000 per country per year. WHO-UNAIDS definition of AIDS (Acquired Immunodeficiency Syndrome) is that AIDS is the most severe manifestation of infection with the HIV (human immunodeficiency virus). The Centers for Disease Control and Prevention (CDC) lists numerous opportunistic infections and neoplasms (cancers) that, in the presence of HIV infection, constitute an AIDS diagnosis. In 1993, CDC expanded the criteria for an AIDS diagnosis to include CD4+ T-cell count at or below 200 cells per microliter in the presence of HIV infection. In persons (aged 5 and older) with normally functioning immune systems, CD4+ T-cell counts usually range from 500 to 1500 cells per microliter. Persons living with AIDS often have infections of the lungs, brain, eyes and other organs, and frequently suffer debilitating weight loss, diarrhoea,

and a type of cancer called Kaposi's sarcoma.

- *Contraceptive Prevalence*: Data on the percentage of women aged 15-49 who are using, or whose partners are using, any form of contraception, whether modern or traditional are available from World Bank, World Development Indicators (2006). The data are available only for 34 countries and some years between 1985–2004.

- *Desired Fertility Rate*: The data are available for 34 countries whose survey years fall between 1986–2004. Two measures are used; both from DHS, www.measuredhs.com, MEASURE DHS, Macro International Inc., for ages 15–49. 1) Wanted fertility: The survey question is as follows: At the time you became pregnant with (NAME), did you want to become pregnant then, did you want to wait until later, or did you not want to have any (more) children at all? Respondents can answer as: Then, Later, Unwanted. This question permits estimation of what the level of fertility would be if only wanted births had occurred. 2) The ideal number of children for each woman: This question asks one number from each women which would be their ideal number of children.

- *Distance to Democratic Republic of Congo in kms:* Pair-wise distance is taken from Arcview 3.x software, where each country's distance to Congo is measured as the distance from its capital to the capital of Congo.

- *Enrollment Rates*: Gross school enrollment rates are from World Bank, Word Development Indicators (2006). They are available for 35 countries and years between 1985–2004.

- *GDP per capita*: GDP per capita (PPP 2000 $s) is from World Bank, World Development Indicators (2006).

- *HIV*: HIV prevalence rates among pregnant women are from the U.S. Census Bureau, HIV Surveillance Database (2003). UNAIDS/WHO also provides similar data. Both Census and UNAIDS databases collect all studies and estimates of HIV/AIDS prevalence since the early 1980s. They provide information on prevalence, population and

other factors and also provide regional estimates. The main indicator for the epidemic is the percent HIV-1 incidence among pregnant women for each country and year. HIV is the retrovirus isolated and recognized as the etiologic (i.e. causing or contributing to the cause of a disease) agent of AIDS. HIV-1 is classified as a lentivirus in a subgroup of retroviruses. Most viruses and all bacteria, plants, and animals have genetic codes made up of DNA, which uses RNA to build specific proteins. The genetic material of a retrovirus such as HIV is the RNA itself. HIV inserts its own RNA into the host cell's DNA, preventing the host cell from carrying out its natural functions and turning it into an HIV factory. HIV-2 is a virus closely related to HIV-1 that has also been found to cause AIDS. It was first isolated in West Africa. Although HIV-1 and HIV-2 are similar in their viral structure, modes of transmission, and resulting opportunistic infections, they have differed in their geographical patterns of infection.

- *HIV-EPP*: The International Programs Center of the Census Bureau uses Estimation and Projection Package (EPP) from WHO/UNAIDS to estimate and project adult HIV prevalence among 15–49 year old from surveillance data between 1985–2004. While EPP can be used in all countries with sufficient surveillance data, it is specifically recommended for countries with generalized epidemics. Generalized epidemics are those that have broken out into the general population or consistent HIV prevalence at over 1 percent in low risk individuals. The proxy for low risk individuals is women attending antenatal clinics. The input to EPP in countries with generalized epidemics is surveillance data from various sites and years showing HIV prevalence among pregnant women, as well as data from national population-based surveys. EPP estimates the trends over time of HIV prevalence by fitting an epidemiological model to data from urban and rural sites. It tests possible epidemiological parameters, chooses a set minimizing least squares and projects future course based on fitted parameters, such as a parameter for the start year of the epidemic; one for the force of infection (how explosive the epidemic is in its initial stage); one for the fraction of new entrants to the population going into to the at-risk category (a parameter largely determines where

the epidemic levels off); and one for the recruitment (a high value means people are brought into the at-risk population as people die of HIV, thus helping to sustain the epidemic at a higher level).

- *HIV-Oster*: Oster (2006) estimates HIV rates for 9 countries between 1985–2000 for 15–35 year old individuals of both genders. She develops a methodology to estimate HIV prevalence over time from mortality data. To avoid the problem of lack of official mortality statistics for Africa she takes advantage of sibling mortality histories in the DHS.

- *Know Someone Died of AIDS*: The data on the percent female who know someone personally who has the virus that causes AIDS or has died of AIDS are from DHS, www.measuredhs.com, MEASURE DHS, Macro International Inc. The data are available for 23 countries whose survey years fall between 1993–2004.

- *Literacy*: The data on percent of people ages 15 and above are from World Bank, World Development Indicators (2006). The data are available for 42 countries and 2 years only (1990, 2004).

- *Mortality*:

  Infant Mortality: Infant mortality is the rate per 1000 live births and from World Bank, World Development Indicators (2006). The data are available for 8 years (1985, 1987, 1990, 1992, 1995, 1997, 2000,2004).

  Age 5 Mortality: Age 5 mortality is the rate per 1000 children under age 5 and from World Bank, World Development Indicators (2006). The data are available for 5 years (1985, 1990, 1995, 2000, 2004).

- *Total Fertility Rate*: Data on total fertility rates are from World Bank, World Development Indicators (2006) and available for 10 years (1985, 1987, 1990, 1992, 1995, 1997, 2000, 2002, 2003, 2004) and 44 countries. DHS data on total fertility rate per woman ages 15–49 are from DHS, www.measuredhs.com, MEASURE DHS, Macro Interna-

tional Inc. The data are available for 34 countries whose survey years fall between 1986–2004.

- *Total Fertility Rate in 1960s and in 1970s*: 1960s rate is the average of rates in 1960, 1962, 1967 and for 1970s it is the average rates in 1970, 1972, 1977.

- *Urbanization*: Urbanization is the percent of urban population in total population and from World Bank, World Development Indicators (2006).

Regional Level Data:

Regions:

Benin: Atacora Province, Atlantique Province, Borgou Province, Mono Province, Oueme Province, Zou Province.

Ethiopia: Addis Ababa, Dire Dawa, Gambella, Harari.

Ghana: Accra, Northern region, Upper East region, Upper West region.

Lesotho: Maseru, Leribe district, Mafeteng district, Quthing district, Mokhotlong.

Madagascar: Antananarivo, Antsiranana, Fianarantsoa, Mahajanga, Toamasina, Toliary.

Malawi: Lilongwe, Blantyre, Mangochi, Mulanje, Mzimba, Thyolo.

Mali:Bamako, Koulikoro, Mopti, Sikasso.

Niger:Dosso, Maradi, Niamey, Tahoua, Zinder.

Nigeria: North East zone, North West zone, South East zone, South West zone.

Rwanda: Butare, Byumba, Gisenyi, Kigali, Ruhengeri.

South Africa: Eastern Cape Province, Free State Province, Gauteng Province, Mpumalanga Province, Northern Cape Province, Northern Province, North-West Province, Western Cape Province.

Tanzania: Dar es Salaam, Rukwa region, Arusha region, Zanzibar area.

Togo: Kara, Plateaux, Savanes.

Zimbabwe: Harare, Bulawayo, Manicaland, Masvingo, Mashonaland West Province, Matabeleland South.

- *Fertility Rates*: Regional fertility rates are from DHS, www.measuredhs.com, MEASURE DHS, Macro International Inc., and available for 14 countries, whose surveys years fall between 1988–2004.

- *Distance to Democratic Republic of Congo in kms:* Distance from center of every region to the center of Congo is provided by Emily Oster.

- *HIV Rates-US Census*: Regional HIV data come from U.S. Census Bureau, HIV Surveillance Database (2005) and available for 14 African countries. The data are available for 1985–1990 and also for later years for a smaller number of regions.

Individual Level Data:

To replicate the results of Young (2005, 2007) individual level data are used for 27 countries from 57 Demographic Health Surveys: Benin (1996, 2001), Burkina Faso (1992/1993, 1998/1999, 2003), Burundi (1987), Cameroon (1991, 1998), Central Republic of Africa(1994/1995), Chad (1996/1997), Cote D'Ivoire (1994, 1998), Ethiopia (2000), Gabon (2000), Ghana (1988, 1993, 1998, 2003), Guinea (1999), Kenya (1989, 1993, 1998, 2003), Liberia (1986), Malawi (1992, 2000), Mali (1987, 1995/1996, 2001), Mozambique (1997), Namibia (1992, 2000), Niger (1992, 1998), Nigeria (1990, 1999, 2003), Rwanda (1992, 2000), Senegal (1986, 1992/1993, 1997), South Africa (1998), Tanzania (1992, 1996, 1999), Togo (1988, 1998), Uganda (1988, 1995, 2000/2001), Zambia (1992, 1996, 2001/2002), Zimbabwe (1988, 1994, 1999).

- *Educational Attainment*: This is a categorical variable for woman's educational attainment level. Categories are "No Education", "Primary Education", "Secondary Education", "Tertiary Education" (v106). Note that v106 combines those with incomplete schooling with those who have completed school both for primary and secondary. A better variable to use might be v 149 with greater specificity. However since the individual data is used in replicating Young's work, I stick to the variables he used.

- *Fertility:* Measured as number of births in last year for each woman (v209).
- *Desired Fertility*: The ideal number of children for each woman (v613): This question asks one number from each women which would be their ideal number of children.
- *Other Controls*: Other control variables from are: Age (v121), year of survey (v007), presence of radio in the household (v120), presence of television in the household (v121), presence of refrigerator in the household (v122), presence of bicycle in the household (v123), urban/rural (v102), number of born children (v201), number of living children (v201-v206-v207).

For South Africa, South Africa October Household Survey (OHS) (1995) is also used, available at www.statssa.gov.za. Using both South Africa DHS (1998) and OHS (1995) income is estimated as a function of age, sex, and education. Individual education levels reported in OHS and DHS are converted to standardized years of education: (i) No schooling or less than one year completed = 0 years, (ii) Sub A/sub B/grade 1/grade 2/Std 1 = 2 years, (iii) Standards 2-10 = standard year + 2, (iv) Diploma/certificate with Std 9 or lower or further studies incomplete = 13 years, (v) Diploma/certificate with Std 10 or diploma/other post-school complete = 14 years, (vi) Degree or further degree complete = 16 years. These educational categories are then used to construct wage index in replicating Young (2005). Wage index is calculated as: $B_E * E_i + B_{E_2} * E_i^2$, where B_E and B_{E^2} are the return to education coefficients coming from the regression of after-tax wages on age, education and their squares.

References

Acemoglu, Daron and Simon Johnson, "Disease and Development: The Effect of Life Expectancy on Economic Growth," MIT, mimeo (2006).

Allen, S., Serufilira, V., Gruber, S., Kegeles, P., Van de Perre, M., Carael , and T.J. Coates, "Pregnancy and Contraception Use among Urban Rwandan Women after HIV Testing and Counseling," *American Journal of Public Health* 83 (5) (1993), 705–710.

Becker, G. S., R. J. Barro. (1988). "A Reformulation of the Economic Theory of Fertility," *Quarterly Journal of Economics* 103, 1–25.

Bell Clive, Devarajan Shantayanan, and Hans Gersbach, "The long-run economic costs of AIDS: theory and application to South Africa," June draft. University of Heidelberg, World Bank (2003).

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan, "How Much Should We Trust Difference in Differences Estimates?," *Quarterly Journal of Economics* 119 (1) (2004): 249–275.

Bloom, David and Ajay Mahal, "Does the AIDS Epidemic threaten Economic Growth" *Journal of Econometrics* 77 (1997), 105–124.

Bloom, Shelah, Charles Banda, Gloria Songolo, Samantha Mulendema, Amy Cunningham, and J. Ties Boerma, "Looking for Change in Response to the AIDS Epidemic: Trends in AIDS Knowledge and Sexual Behavior in Zambia, 1990 through 1998," *Journal of Acquired Immune Deficiency Syndrome* XXV (2000), 77–85.

Boldrin, Michele and Larry E. Jones, "Mortality, Fertility, and Saving in a Malthusian Economy," *Review of Economic Dynamics* 5 (4) (2002), 775–814.

Bonnel, Rene, "HIV/AIDS: Does it Increase or Decrease Growth in Africa?" World Bank, mimeo (2000).

Caldwell, John, I.O. Orubuloyeb, and Pat Caldwell, "Resistances to Behavioural Change to Reduce HIV/AIDS Infection in Predominantly Heterosexual Epidemics in Third World Countries," Health Transition Center, (1999).

Canning, David, "The Economics of HIV/AIDS in Low-Income Countries: The Case for Prevention", *Journal of Economic Perspectives,* 20 (3) (2006), 121–142.

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

Cervellati and Sunde, "Human Capital, Mortality, and Fertility: A Unified Theory of Demographic Transition," IZA Working Paper 2905.

Cheluget, B., G. Baltazar, P. Orege, M. Ibrahim, L. H. Marum, J. Stover, "Evidence for Population Level Declines in Adult HIV Prevalence in Kenya," *Sexually Transmitted Infections* 82, Supplement 1 (2006), 121–126.

Corrigan, Paul, Gerhard Gloom, and Fabio Mendez, "AIDS Crisis and Growth," *Journal of Development Economics* 43, (2005).

Economist, "The Shower Method, " April 22 (2006).

Ehrlich, Isaac, and Francis T. Lui, "Intergenerational Trade, Longevity, Intrafamily Transfers and Economic Growth," *Journal of Political Economy* 99 (1991), 1029–1059.

Feeney, Griffith, "The Impact of HIV/AIDS on Adult Mortality in Zimbabwe," *Population and Development Review* 27–4 (2001), 771–780.

Fortson, Jane, "The Effect of HIV and AIDS on Educational Attainment," University of Chicago, mimeo, (2007).

Francis, Andrew, M., "The Economics of Sexuality: The Effect of HIV/AIDS on Homosexual Behavior, Desire, and Identity in the United States," University of Chicago, mimeo, (2006).

Fylkesnes K., Musonda R. M., Sichone, M. Ndhlovu, Z. Tembo, F. Monze, "Declining HIV Prevalence and and Risk Behaviors in Zambia: Evidence from Surveillance and Population-Based Surveys," *AIDS* 15, 7 (2001), 907–916.

Galor, O. and D. N. Weil. (1999). "From Malthusian Stagnation to Modern Growth," *American Economic Review* 89, 150–154.

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

Hallett T. B., Aberle-Grasse J., "Declines in HIV Prevalence can be Associated with changing sexual behavior in Uganda," *Sexually Transmitted Infections* 82 Supplement 1 (2006), 1–8.

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

Juhn, Chinhui, Sebnem Kalemli-Ozcan, and Belgi Turan, "HIV and Fertility in Africa: First Evidence from Population Based Surveys," University of Houston, mimeo, (2007).

Kalemli-Ozcan, Sebnem, "A Stochastic Model of Mortality, Fertility and Human Capital Investment," *Journal of Development Economics* 62 (2003).

Kalemli-Ozcan, Sebnem, "Does Mortality Decline Promote Economic Growth?" *Journal of Economic Growth* 7 (2002).

Kalemli-Ozcan, Sebnem, "AIDS, Reversal of the Demographic Transition and Economic Development: Evidence from Africa," CDRLL Working Paper, Stanford University, May (2005), http://www.stanford.edu/ wacziarg/demogworkshop.html.

Kezdi, Gabor, "Robust Standard Error Estimation in Fixed-Effects Panel Models," University of Budapest, mimeo, (2004).

Kremer, Michael, "Integrating Behavioral Choice into Epidemiological Models of the AIDS Epidemic, " *Quarterly Journal of Economics* 111–2 (1996), 549–573.

Lagarde, Gilles Pison Emmanual and Catherine Enel, "Knowledge, attitudes and perception of AIDS in rural Senegal: relationship to sexual behaviour and behaviour change," *AIDS,* March (1996), 10 (3), 327-34.

Lindan, C, S Allen, M Carael, F Nsengumuremyi, P Van de Perre, A Serufilira, J Tice, D Black, "Knowledge, attitudes, and perceived risk of AIDS among urban Rwandan women: relationship to HIV infection and behavior change," *AIDS,* (1991), 5 (8), 993-1002.

Lorentzen, Peter, John McMillan, and Romain Wacziarg, "Death and Development," Stanford University, mimeo (2007).

Lucas, R. E. (2002). *Lectures on Economic Growth.* Cambridge: Harvard.

Luke, Nancy and Kaivan Munshi, "New Roles for Marriage in Urban Africa: Sexual Activity and Labor Market Outcomes in Kisumu", Brown University, mimeo (2004).

McNeill, Donald, "U.N. to Say It Overestimated AIDS Cases by Millions," New York Times, November (2007).

Meltzer, David, "Mortality Decline, the Demographic Transition and Economic Growth," Ph.D Dissertation, University of Chicago (1992).

Mwaluko, Gabriel, Mark Urassa, Raphael Isingo, Basia Zaba, and J. Ties Boerma, "Trends in HIV and Sexual Behavior in a Longitudinal Study in a Rural Population in Tanzania, 1994–2000," *AIDS* XVII (2003), 2645–2651.

Ng'weshemi, JZ, JT Boerma, R Pool, L Barongo, K Senkoro, M Maswe M, R Isingo, D Schapink, "Changes in male sexual behaviour in response to the AIDS epidemic: evidence from a cohort study in urban Tanzania," *AIDS,* October (1996), 10 (12), 1415-20.

Noel-Miller, C. M., "Concern Regarding the HIV/AIDS Epidemic and Individual Childbearing, Evidence from Rural Malawi," *Demographic Research, online (2003).*

Oster, Emily, "Sexually Transmitted Infections, Sexual Behavior, and the HIV/AIDS Epidemic, " *Quarterly Journal of Economics* May (2005), 467–514.

Oster, Emily, "Routes of Infection: Exports and HIV Incidence in Sub-Saharan Africa", University of Chicago, mimeo (2006).

Oster, Emily, "HIV and Sexual Behavior Change: Why Not Africa?", University of Chicago, mimeo (2007).

Over, Mead, "The Macroeconomic Impact of AIDS in Sub-Saharan Africa", World Bank Working Paper (1992).

Papageorgiou, Chris and Petia Stoytcheva "What Do We Know About the Impact of AIDS on Cross-Country Income So Far?," LSU, mimeo (2007).

Peterson, Mitchell, "Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches," Northwestern University, mimeo, (2007).

Pritchett, Lant H., "Desired Fertility and the Impact of Population Policies," *Population and Development Review* 1 (1994).

Sah, K. Raaj, "The Effects of Child Mortality Changes on Fertility Choice and Parental Welfare," *Journal of Political Economy* 99 (1991), 582–606.

Schultz, T. Paul, "Demand for Children in Low Income Countries." In Mark R. Rosenzweig and Oded Stark (eds), *Handbook of Population and Family Economics*, (1997) Amsterdam: Elsevier Science.

Soares, Rodrigo, "Mortality Reductions, Educational Attainment, and Fertility Choice," *American Economic Review* 95(3), June 2005, 580-601.

South Africa Department of Health Report, (2007).

Stoneburner, Rand and Daniel Low-Beer, "Population-Level HIV Declines and Behavioral Risk Avoidance in Uganda," *Science,* April (2004), 304, 714-718.

Tamura, R. (2006). "Human Capital and Economic Development." *Journal of Development Economics,* 79, (2006), 26–72.

Temmerman, M. S. Moses, D. Kiragu, S. Fusallah, I. A. Wamola, and P. Piot, "Impact of single session post-partum counselling of HIV infected women on their subsequent reproductive behavior," *AIDS* 2 (1990), 247–252.

Timberg, Craig, "How AIDS in Africa Was Overstated," Washington Post, April 6 (2006).

Thornton, Rebecca, "The Demand for and Impact of Learning HIV Status: Evidence from a Field Experiment," Harvard University, mimeo (2006).

UNAIDS/WHO, *Global Report (2006).*

UNAIDS/WHO, *Epidemiological Fact Sheets (2003).*

U.S. Census Bureau, *HIV Surveillance DataBase (2005).*

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

Weil, David N., *Economic Growth* (2003), Addison-Wesley.

Werker, Eric, Amrita Ahuja, and Brian Wendell, "Male Circumcision and the Economic Impact of AIDS in Africa, " Harvard Business School, mimeo (2006).

Westoff, Charles, and Anne Cross, "The Stall in the Fertility Transition in Kenya," DHS Analytical Studies No. 9, (2006).

Williams, Brian, Dirk Taljaard, Catherine Campbell, Eleanor Gouws, Lewis Ndhlovu, "Changing Patterns of Knowledge, Reported Behavior and Sexually Transmitted Infections in a South African Gold Mining Community," *AIDS* XVII (2003), 2099–2107.

World Bank, *World Development Indicators (2006).*

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

Young, Alwyn, "In Sorrow to Bring Forth Children: Fertility Amidst the Plaque of HIV, " University of Chicago, October, mimeo (2005b).

Zaba, B. and S. Gregson, "Measuring the Impact of HIV on Fertility in Africa," *AIDS* 12 Supplement 1 (1998), 41–50.

Table 1: HIV Antenatal Clinic (ANC) Testing Statistics, 1988–2004

	Female Population requesting a HIV test, receiving a test, and receiving test results	Female Population receiving a test	Female Population receiving test results	Pregnant counseled for HIV during ANC visit	Pregnant tested for HIV during ANC visit
Botswana	9				
Burundi	2				
Cameroon	18	21	5	36	9
Cote d'Ivoire	5				
Gambia	6				
Guinea-Bissau	4	15	7		
Kenya	8.5			29	3
Lesotho	9				
Mozambique	3	4	2	51	1
Niger	1				
Nigeria	4	7	3	24	1
Senegal	3				
Sierra Leone	2				
Togo	3				
Zambia	8			43	22

Notes: Statistics are reported in percent and taken from DHS, www.measuredhs.com, MEASURE DHS, Macro International Inc. The data for Gambia, Guinea-Bissau, and Sierra Leone are from Multi Indicator Cluster Surveys (MICS) for 2000. For other countries DHS survey years fall between 1988–2004.

Table 2: **Descriptive Statistics, 1985–2000**

	No. of Countries	Mean	Std.dev.	Max	Min
Total Fertility Rate, WB	44	5.71	1.22	8.06	2.08
Total Fertility Rate, DHS	34	6.07	0.87	7.40	3.90
Desired Fertility Rate, DHS	34	4.71	0.99	6.90	2.30
AIDS (per 100,000)	42	22.38	32.84	162.16	0.02
HIV	44	0.08	0.07	0.25	0.001
HIV-EPP	38	0.06	0.05	0.21	0.004
HIV-Oster	9	0.04	0.03	0.09	0.009
Know Someone Died of AIDS (%)	23	40.81	21.03	88.45	6.90
Primary School for Female (%)	38	77.17	29.11	133.65	25.06
Secondary School for Female (%)	38	24.55	23.62	112.82	4.24
Primary School for Male (%)	38	88.73	24.96	136.49	38.47
Secondary School for Male (%)	38	29.79	15.18	80.97	7.33
Female Literacy (%)	42	45.29	23.32	92.26	10.09
Male Literacy (%)	42	62.77	17.89	91.41	26.73
GDP per capita (PPP 1996 $s)	44	798.24	1217.78	6168.33	95.92
Urban Population (%)	44	31.68	13.37	73.76	7.53
Infant Mortality (per 1000)	44	100.30	36.06	176.75	14.99
Mortality Under 5 (per 1000)	44	159.11	63.80	295.76	17.60
Contraceptive Use (%)	27	20.16	14.38	63.75	4.00

Notes: All variables are averaged over 1985−2000 depending on the availability. Total Fertility Rate is the sum of age-specific fertility rates (number of children that a woman would have if she lived through all of her child-bearing years and experienced the current age-specific fertility rates at each age); from World Bank (WB), World Development Indicators (WDI) and from DHS, www.measuredhs.com, MEASURE DHS, Macro International Inc., respectively. The survey years for the data from DHS fall between 1986–2004. Data Appendix reports the survey years for each country. Desired Fertility Rate represents wanted fertility per woman ages 15–49; from DHS. AIDS represents the number of officially reported AIDS cases per 100,000 per country per year, calculated as multiplying the officially reported AIDS cases by 100,000 and dividing by population; from WHO/UNAIDS, Epidemiological Fact Sheets. HIV represents percent HIV-1 sero-prevalence infection rate among pregnant women attending antenatal clinics; from U.S. Census Bureau, HIV Surveillance Database. HIV-EPP represents *estimated* national HIV prevalence among 15 to 49 year olds calculated by fitting an epidemiological model to data (Estimation and Projection Package-EPP) from urban and rural surveillance sites; from U.S. Census Bureau, International Programs Center. HIV-Oster represents Oster (2006) estimates that are based on mortality data from sibling histories in the DHS. Know Someone Died of AIDS, represents the percent female who know someone personally who has the virus that causes AIDS or has died of AIDS; from DHS. Schooling variables are gross enrollment rates for primary and secondary school respectively; from WDI. Literacy rates are in percent for male and female in total population respectively; from WDI. GDP per capita is the Gross Domestic Product (PPP 1996 $) divided by population; from WDI. Urban Population is the percent of urban population in total population; from WDI. Infant Mortality is the infant mortality rate per 1000 births; from WDI. Mortality under 5 is the age 5 and under mortality per 1000 births; from WDI. Contraceptive Use represents the percent women aged 15-49 who are using, or whose partners are using, any form of contraception; from WDI. See Appendix for more information on the variables.

Table 3: AIDS, HIV and Fertility: Between Regressions

Dependent variable: Total Fertility Rate (TFR)

Source for TFR:	WB	WB	WB	DHS	DHS	DHS
	(1)	(2)	(3)	(4)	(5)	(6)
Log AIDS	0.14	–	–	0.20	–	–
	(0.05)	–	–	(0.08)	–	–
Log HIV	–	0.13	–	–	0.20	–
	–	(0.07)	–	–	(0.09)	–
Log HIV-EPP	–	–	–0.04	–	–	–
	–	–	(0.11)	–	–	–
Know Someone Died	–	–	–	–	–	0.02
of AIDS	–	–	–	–	–	(0.07)
Female Schooling	–0.02	–0.02	–0.02	–0.02	–0.03	–0.02
	(0.006)	(0.006)	(0.006)	(0.009)	(0.01)	(0.009)
Log GDP per capita	–0.11	–0.10	–0.11	–0.01	–0.12	–0.13
	(0.06)	(0.08)	(0.07)	(0.17)	(0.20)	(0.19)
Infant Mortality	0.02	0.02	0.01	0.01	0.01	0.01
	(0.003)	(0.003	(0.003))	(0.004)	(0.01)	(0.007)
R^2	0.84	0.88	0.80	0.63	0.61	0.67
Observations	33	35	30	26	26	22

Notes: Robust standard errors (White correction) are in parentheses. The Between Regressions report the results using country averages depending on availability, and including a constant. See table 2 for the detailed explanation of the variables.

Table 4: AIDS, HIV and Fertility: Falsification on Between Regressions

Falsification Exercise—Dependent variable: Total Fertility Rate in 1980s		
	(1)	(2)
Log AIDS	0.01	–
	(0.07)	–
Log HIV	–	0.08
	–	(0.10)
Controls	Yes	Yes
R^2	0.70	0.72
Observations	32	33

Notes: Robust standard errors (White correction) are in parentheses. The Between Regressions report the results using country averages depending on availability, and including a constant. See table 3 for the set of controls. See table 2 for the detailed explanation of the variables.

Table 5: AIDS, HIV and Fertility: Pooled and Within Regressions

Dependent variable: Total Fertility Rate (TFR)

	Pooled (1)	Pooled (2)	Pooled (3)	Within (4)	Within (5)	Within (6)	Within (8)	Within (9)	Within (10)
Log AIDS	0.09	–	–	0.09	–	–	–0.06	–	–
	(0.04)	–	–	(0.04)	–	–	(0.05)	–	–
Log HIV	–	0.12	–	–	–0.01	–	–	–0.08	–
	–	(0.06)	–	–	(0.05)	–	–	(0.05)	–
Log	–	–	–0.01	–	–	0.10	–	–	–0.15
HIV-EPP	–	–	(0.07)	–	–	(0.10)	–	–	(0.05)
Female	–0.02	–0.03	–0.02	–0.004	–0.001	–0.002	–0.03	–0.003	–0.03
Schooling	(0.005)	(0.005)	(0.005)	(0.01)	(0.01)	(0.01)	(0.02)	(0.008)	(0.01)
Log GDP	–0.10	–0.12	–0.12	–0.27	–0.29	–0.43	–0.27	–0.29	–0.13
per capita	(0.06)	(0.08)	(0.07)	(0.31)	(0.27)	(0.29)	(0.27)	(0.27)	(0.24)
Infant	0.01	0.01	0.01	0.003	0.006	0.004	0.007	0.01	0.005
Mortality	(0.002)	(0.032)	(0.002)	(0.01)	(0.004)	(0.01)	(0.007)	(0.005)	(0.004)
Common Trend	–0.05	–0.04	–0.03	–	–	–	–	–	–
	(0.02)	(0.01)	(0.01)	–	–	–	–	–	–
Country Effects	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Year Effects	No	No	No	Yes	Yes	Yes	No	No	No
Country Trends	No	No	No	No	No	No	Yes	Yes	Yes
R^2	0.76	0.82	0.76	0.87	0.87	0.87	0.95	0.95	0.95
Observations	111	153	139	111	153	139	111	153	139
Countries	33	35	30	33	35	30	33	35	30

Notes: Robust standard errors (White correction; clustered on countries) are in parentheses. The Within Regressions report results using country fixed effects. See table 2 for the detailed explanation of the variables.

Table 6: Instrumental Variables Regression: First Stage

Dependent variable:	Log AIDS (1)	Log AIDS (2)	Log HIV (3)	Log HIV (4)
Distance (1000km)	–0.6	–0.7	–0.5	–0.6
	(0.2)	(0.2)	(0.2)	(0.2)
Female Schooling	–	–0.01	–	0.05
	–	(0.01)	–	(0.01)
Log GDP per capita	–	–0.13	–	–0.67
	–	(0.26)	–	(0.27)
Infant Mortality	–	0.01	–	0.01
	–	(0.01)	–	(0.01)
Regional Dummies	No	Yes	No	Yes
R^2	0.28	0.34	0.20	0.36
Observations	33	33	35	35

Notes: Robust standard errors (White correction) are in parentheses. The Between Regressions report the results using country averages of the variables, and including a constant. Distance to Dem. Congo is in 1000km. See table 2 for the detailed explanation of the variables.

Table 7: Instrumental Variable Regressions: Second Stage

Dependent variable: Total Fertility Rate (TFR)

	(1)	(2)
Log AIDS	0.28	–
	(0.09)	–
Log HIV	–	0.37
	–	(0.14)
Female Schooling	–0.02	–0.03
	(0.006)	(0.007)
Log GDP per capita	–0.15	–0.02
	(0.08)	(0.11)
Infant Mortality	0.02	0.01
	(0.003)	(0.004)
R^2	0.80	0.82
Observations	33	35

Notes: Robust standard errors (White correction) are in parentheses. The 2SLS Regressions report the results using country averages, and including a constant. Distance to Dem. Congo is in 1000km. See table 2 for the detailed explanation of the variables.

Table 8: Falsification on Exclusion Restriction

Dependent var.:	TFR in 1980s (1)	Log GDP (2)	Infant Mort. (3)	Child Mort. (4)
Distance (1000km)	0.03	0.01	5.83	–12.1
	(0.1)	(0.1)	(5.98)	(11.4)
R^2	0.00	0.00	0.03	0.04
Observations	33	35	35	35

Notes: Robust standard errors (White correction) are in parentheses. The Between Regressions report the results using country averages of the variables, and including a constant. Distance to Dem. Congo is in 1000km. See table 2 for the detailed explanation of the variables.

Table 9: HIV and Fertility: Between Regressions at the Regional Level

Dependent variable: TFR in 1998–2004		
	(1)	(2)
Log HIV in 1990–1995	0.29	0.29
	(0.05)	(0.08)
Country Dummies	Yes	Yes
Cluster	Region	Country
R^2	0.79	0.79
Observations	32	32
Countries	12	12

Notes: Robust standard errors (column 1: clustered on regions, column 2: clustered on countries) are in parentheses. All regressions report results using country fixed effects. Regional TFRs are from DHS, various survey years (mean: 5.07, std dev.: 1.60, max: 8.7, min: 1.9). Each country's survey year falls between 1987–2004. The data are averaged over the survey years. Regional HIV rates (percent HIV-1 sero-prevalence among pregnant women) are from the U.S. Census Bureau, HIV Surveillance Database (2003) (mean: 0.047, std dev.: 0.079, max: 0.3094, min: 0). HIV prevalence rates are averaged over 1990–1995 or used as a single year depending on the availability.

Table 10: HIV and Individual Fertility in South Africa

Dependent variable is Retrorespective Fertility

Estimation Time Period	Poisson (1961–1998) (1)	Poisson (1990–1998) (2)	OLS (1961–1998) (3)	OLS (1990–1998) (4)
HIV-US Census by age group	–1.36 (0.189)	1.64 (0.539)	–0.164 (0.025)	0.227 (0.066)
Age	0.503 (0.008)	0.181 (0.023)	0.043 (0.000)	0.004 (0.002)
Age^2	–0.010 (0.000)	–0.005 (0.000)	–0.001 (0.000)	–0.000 (0.000)
Wage Index	–0.335 (0.013)	–0.339 (0.023)	–0.041 (0.002)	–0.040 (0.003)
Year of Birth	–0.003 (0.001)	–0.065 (0.014)	–0.001 (0.000)	–0.007 (0.002)
Observations	171206	58208	171206	58208

Notes: Only women who are 25 or older are used and there are 7276 of them in the data set. The panel is constructed using each womans birth histories since age 12, and it includes the period between 1961–1998. Retrorespective fertility is the number of pregnancies of each woman in each year, including that were lost before term or resulted in stillbirths. HIV prevalence rates for South Africa are available since 1990, therefore, following Young's methodology, in columns (1) and (3) HIV prevalence is taken as 0 before 1990. In columns (2) and (4) only part of the panel that is after 1990 is used in regressions. Wage index is the estimated wage for each individual depending on own education following Young's methodology. Each regression has a constant. Robust standard errors (White correction; clustered on individuals) are in parentheses.

Table 11: HIV and Individual Fertility in Africa

Dependent variable is Past Year Births

Estimation Cluster:	Poisson Individual (1)	Poisson Individual (2)	Poisson Country (3)	Poisson Country (4)	OLS Individual (5)	OLS Individual (6)	OLS Country (7)	OLS Country (8)
HIV-EPP	–1.260	–0.963	–1.270	–0.963	–0.218	–0.234	–0.218	–0.234
	(0.380)	(0.385)	(1.17)	(0.908)	(0.057)	(0.054)	(0.201)	(0.190)
Primary	–0.294	0.071	–0.294	0.071	0.060	0.010	0.060	0.010
Education	(0.054)	(0.057)	(0.005)	(0.001)	(0.009)	(0.010)	(0.001)	(0.000)
Secondary	–0.798	0.252	–0.798	0.252	–0.132	0.029	–0.132	0.029
Education	(0.091)	(0.102)	(0.005)	(0.012)	(0.010)	(0.012)	(0.001)	(0.002)
Tertiary	–1.190	0.759	–1.190	0.759	–0.155	0.116	–0.155	0.116
Education	(0.445)	(0.461)	(0.007)	(0.036)	(0.030)	(0.033)	(0.001)	(0.007)
Age	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age^2	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	No	Yes	No	Yes	No	Yes
Country Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country Trends	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	403100	350586	403100	350586	403100	350586	403100	350586

Notes: Countries and survey years are Benin (1996, 2001), Burkina Faso (1992/1993, 1998/1999, 2003), Burundi (1987), Cameroon (1991, 1998), Central Republic of Africa(1994/1995), Chad (1996/1997), Cote D'Ivoire (1998/1999), Ethiopia (2000), Gabon (2000), Ghana (1988, 1993, 1998, 2003), Guinea (1999), Kenya (1989, 1993, 1998, 2003), Liberia (1986), Malawi (1992, 2000), Mali (1987, 1995/1996, 2001), Mozambique (1997), Namibia (1992, 2000), Niger (1992, 1998), Nigeria (1990, 1999, 2003), Rwanda (1992, 2000), Senegal (1986, 1992/1993, 1997), South Africa (1998), Tanzania (1992, 1996, 1999), Togo (1988, 1998), Uganda (1988, 1995, 2000/2001), Zambia (1992, 1996, 2001/2002), Zimbabwe (1988, 1994, 1999). Other controls in the regressions are marital status (never, currently or formerly married), urban/rural location, the number and square of born and living children, and the presence of a radio, television, refrigerator or bicycle (each entered separately) in the household. Each regression has a constant. Robust standard errors (White correction; clustered as indicated) are in parentheses.