

HIV and Fertility: Long-term Evidence from Sub-Saharan Africa

Günther Fink
Harvard School of Public Health

Sebastian Linnemayr¹
Harvard School of Public Health

Preliminary version - Please do not cite

This version: September 2008

Abstract

HIV/AIDS constitutes the single largest negative health shock to the international health system over the last fifty years. Reducing life expectancy by up to 20 years in some countries, the arrival of HIV has resulted in a unique disruption of the global upward trend in life expectancy, and thus provides a natural study ground for the interactions between health and socioeconomic change. In this paper, we focus on one particular aspect of behavioral change: the changes in fertility triggered by the HIV epidemic. Combining historical micro data from the World Fertility Surveys (WFS) with the most recent rounds of the Demography and Health Surveys (DHS) we investigate the HIV fertility response exploring variations both across regions and time. Our results suggest a weak relation between HIV and fertility on the regional level. On the individual level, however, the fertility response is large, and critically depends on the human capital of the exposed populations: while women with primary school or less increase their fertility in the presence of HIV, the opposite is true for women with secondary or higher education. The differential fertility response to HIV across educational groups is consistent with different discount rates across groups as well as with models of differential disease specific knowledge absorption across educational categories.

¹ Corresponding author: Center for Population and Development Studies, 9 Bow Street, Cambridge, MA 02138. E-mail: slinnema@hsph.harvard.edu.

The authors are grateful to David Canning, Jocelyn Finlay, Michael Moore, Younghwan Song and the participants of the Population Economics Program (PEP) Lunch Seminar Series for the comments and suggestions.

1. Introduction

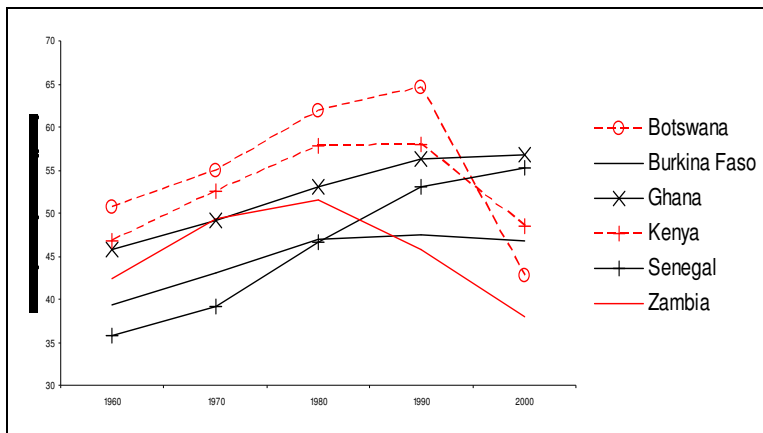
Worldwide, about 40 million people are HIV-positive; the disease has cost at least 12 million lives to date. About two thirds of the people infected with HIV are concentrated in Sub-Saharan Africa, where large parts of the adult population are infected with the virus (UNAIDS, 2008). In this region, the HIV epidemic represents the largest health shock experienced in the last fifty years. Life expectancy has been reduced by over 10 years in countries like Botswana, Kenya, and Zambia, and child mortality has crept back up to levels last observed in or before 1980 since the surge of the epidemic in the early 1990s (1 and 1b).

Given the large degrees of variation in the exposure to HIV within and across countries, the arrival of HIV/AIDS constitutes a very interesting natural experiment to investigate the effects of health on socio-economic behavior. In this paper, we look at one specific aspect of the interplay between health and demography: the change in family size triggered by changes in health. While there is an extensive literature highlighting the negative correlation between health (as measured by child and adult mortality) in the cross-section of countries and regions² (Schultz, 1997), evidence regarding a direct fertility response to changing health at the micro-level is scarce.

In this study, we combine recent data from the Demography and Health Surveys (DHS) with previous waves of the World Fertility Surveys (WFS) to investigate the effects of deteriorating health on reproductive behavior. By linking the original WFS data to the later DHS waves, we are able to compare populations prior to and after the arrival of the epidemic, and thus to identify the effects of health in a standard difference-in-difference framework. This allows us to control for region specific factors that are constant over time as well as for overall and country specific time trends in fertility.

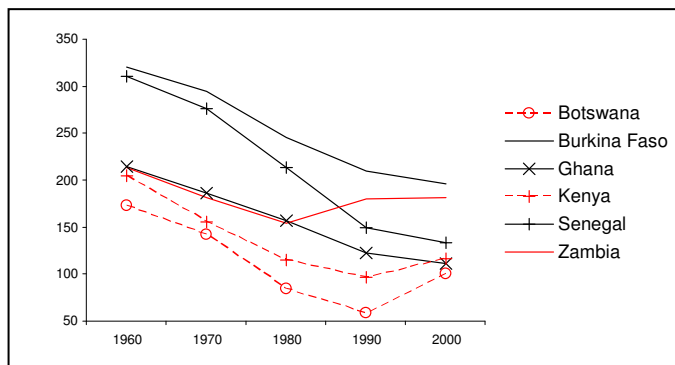
² The correlation between life expectancy and fertility across countries was -0.81 in 2005 (World Bank, 2007).

Figure 1a: Life expectancy



Source: World Development Indicators 2007 (World Bank, 2007)

Figure 1b: Under-5 mortality rate (deaths per 1,000 live births)



Source: World Development Indicators 2007 (World Bank, 2007)

On the regional level, the correlation between changes in fertility and HIV on the regional level is small and negative, but not significant in multivariate analysis. A quite different and more nuanced picture emerges when we allow for interactions between education and HIV at the individual level. While populations with no education or merely primary education increase their fertility in response to HIV, the opposite is true for populations with three or more years of schooling, who reduce their fertility in the presence of HIV. These differences may be interpreted as further evidence in the differential capacity to absorb HIV related information as argued by de Walque (2007)

and Glick and Sahn (2007), but are also consistent with differential discounting rates and planning horizons across income groups as argued by Oster (2007). The emerging differences on the regional level are large; while the HIV epidemic has triggered a significant and positive fertility response in populations with very little education, the arrival of the epidemic has actually accelerated the trend towards smaller family sizes in regions with high average educational attainment.

The empirical results presented in this paper contribute to, and partially reconcile, an emerging literature on the effects of HIV on reproductive behavior. While Young's seminal work on the socioeconomic effects of the HIV epidemic finds a negative correlation between HIV prevalence and fertility (Young, 2005), two recent studies by Kalemli-Ozcan and coauthors argue that the fertility response to HIV is zero or positive as predicted by traditional hoarding models (Kalemli-Ozcan, 2006; Juhn, Kalemli-Ozcan and Turan, 2008). All existing studies either rely on country level variation over time or cross-sectional variation at a given point of time. The approach chosen in this paper allows us to explore both variation within countries and across time, and is to our knowledge the first paper to carefully investigate the long-term interactions between health and human capital in a Sub-Saharan African setting.

The analysis presented in this paper also contributes to an extensive literature on the socioeconomic effects of health in general, and HIV/AIDS in particular. Early economic studies such as Cuddington (1993) have argued for a relatively mitigated, or even positive, impact of the epidemic on economic growth, as a population that is declining at a faster rate than the output can share more output per capita than before the epidemic.³ HIV can, however, have other, more indirect effects on demographics. The large shifts in mortality can trigger behavioral changes in fertility that have the potential to significantly alter the reproductive schedule of women in areas heavily hit by the epidemic. This effect on fertility is of particular importance for sub-Saharan Africa, the region least advanced in its demographic transition from high mortality/higher fertility to

³ A second mechanical effect is biological: HIV-positive women tend to be less fertile than non-infected ones. The magnitude of this effect is of the order of a decrease in the probability of conceiving a child during the past year of around 25% (Gray et al., 1998).

a low mortality/low fertility state. Several papers have stated the case for significant economic growth effects of entering the transition given the right institutional framework (see, for example Bloom and Canning (2008)). Schooling is one of the key factors highlighted in this literature, a result which appears to be strengthened further by the evidence presented in this paper.

The rest of the paper is organized as follows: we introduce a basic theoretical framework in Section 2 of the paper, and then present the data in Section 3. We present the main regional trends in section 4, and our main multivariate results in Section 5; Section 6 concludes.

2. A Simple Framework of HIV and Fertility

A large set of biological, behavior and evolutionary models have been used to analyze the linkage between health and mortality environment and reproductive behavior. From an economic perspective, the easiest way to think about the linkages between HIV and fertility is the traditional static fertility framework introduced by Becker (1960; 1965) and Becker and Lewis (1973). In this framework, the decision making household or woman makes a lifetime decision over the quality and quantity of children given a set of prices and wages. In the simplest form of the model, incomes and prices are given, and higher incomes are associated with increased spending on children. More resources allocated to children do, however, not necessarily imply larger family sizes. If the income elasticity of child quality is sufficiently large, higher incomes will lead to a smaller number of children and higher expenditure on child quality (Becker, 1960). A similar quality-quantity tradeoff emerges in a framework where the decision maker's human capital has positive income effects, but also raises the relative cost of child rearing which requires time spent out of the labor force (Becker, 1965; Willis, 1973).

HIV affects the fertility outcomes determined in a life time optimization framework through several channels: first, HIV has a direct and negative effect on fecundity, and thus imposes additional constraints on family planning. Second, HIV

increases child mortality, making child bearing more risky and costly from the parent's perspective. Last, and potentially most importantly, HIV strongly affects the degree of lifetime uncertainty faced by young adults. This uncertainty entails the uncertainty of treatment possibilities for individuals knowing that they are infected, as well as uncertainties regarding the decision maker's current and future health status.

From an empirical perspective, the direct medical effects in terms of reduced fecundity are likely to be minor. Even with an estimated reduction in the unconditional pregnancy rates around 25 percent among infected women (Gray, Wawer, Serwadda, Sewankambo, Li, Wabwire-Mangen, Paxton, Kiwanuka, Kigozi, Konde-Lule, Quinn, Gaydos and McNairn, 1998), the total effect on fertility rates is rather small in our sample, since the fraction of infected women is relatively small and even those women infected women likely still can get pregnant.

Higher child mortality affects parental choices by making child bearing more risky. In a stochastic setting where parents are sufficiently risk-averse, a higher risk of child death will make parents desire more kids⁴ (Schultz, 1997). This result, however, does not necessarily hold if the degree of risk aversion is small, or HIV more broadly affects the costs associated with child bearing. Mahy (1999) argues that HIV does not only impose an additional burden to women by forcing them to undergo blood tests at antenatal clinics in many cases, but also by exposing their own HIV status to their communities in the case of child death. If the stigma associated with HIV is large enough, the optimal number of children may decrease rather than increase even with very high degrees of risk-aversion.

The effect of parents' lifetime uncertainty on fertility is likely negative; higher lifetime uncertainty generally implies higher discount rates, and, in general, a shift towards short-term consumption. More importantly, in highly risky environments, altruistic parents have to account for the possibility of not being able to support their

⁴ Parents may anticipate child death and therefore plan to give birth to more children ex-ante (hoarding), but may also want to replace children lost ex-post.

children during their childhood; the risk of exposing children to orphanhood constitutes an additional cost of child bearing, lowering the optimal number of children. This effect will be particularly pronounced for parents with a high propensity to invest into the quality of their children; the longer parents want children to stay in school, the higher a cost the life-time uncertainty imposed by HIV constitutes.

The theoretical prior regarding the effect of HIV on fertility is thus highly ambiguous. Even though parents may respond to increases in child mortality by higher fertility, concerns about revealing their own HIV status as well as concerns regarding their own ability to support their children during infancy imply a more restrictive family planning, especially for parents characterized by high human capital and/or inclined to invest substantially into their offsprings' human capital.

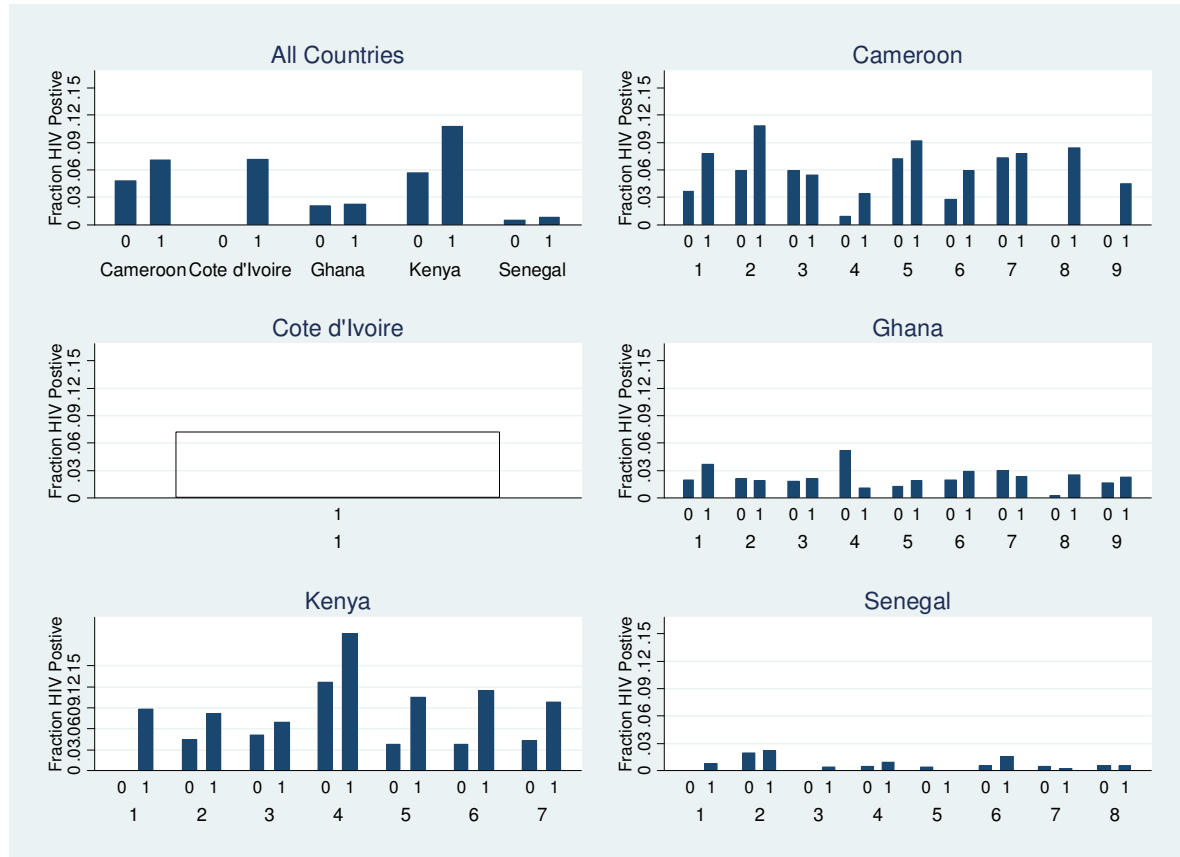
3. The Data

To evaluate the long term relation between HIV and fertility we combine data from the last rounds of the Demographic and Health Surveys (DHS) with data from the World Fertility Surveys (WFS). The World Fertility Surveys were conducted in 41 countries between 1975 and 1982, and were essentially a shorter and less comprehensive version of the later DHS surveys. The WFS was chosen because they offer many of the variables needed in the analysis that were collected when HIV prevalence was most probably close to zero in the countries considered (UNAIDS 2004). While the focus of the World Fertility Surveys was family planning, most of the original survey questions are very similar to the questions used in the later rounds of the DHS surveys.

Currently, 38 WFS surveys are publicly available, 8 of which can be matched to DHS data sets with HIV measurements: Cameroon, Cote d'Ivoire, Dominican Republic, Ghana, Haiti, Kenya, Lesotho and Senegal. Unfortunately, there is a complete mismatch between the sampling framework used in the WFS and DHS surveys for the Dominican Republic, Haiti and Lesotho, which makes a dynamic regional analysis in these three countries impossible. We include only those areas which were targeted in both the DHS

and WFS surveys, which leaves us with 32 geographical sampling areas in the 5 remaining countries. Figure 1 shows the HIV prevalence for all regions in our sample; since prevalence rates vary differ largely between rural and urban areas, we show prevalence rates separately for the rural and urban populations in each region.

Figure 1: HIV Prevalence in Urban and Rural Areas



Notes: 0 is rural, 1 is urban. Bottom line indexes the region number. For Cote d'Ivoire, we have only one (urban) region, the metropolitan area of Abidjan.

As Figure 1 shows, the regions with the highest HIV prevalence rates in our sample are located in Cameroon and Kenya, while Senegal and Ghana have mostly low prevalence rates (1-2% prevalence). All countries show significant regional and urban/rural variation, with urban prevalence rates 5-10 times the rural rates in the most hard-hit areas in each country. The region with the highest prevalence rate in our sample is Nyanza (Kenya), with HIV prevalence rates of 19.6% (urban) and 12.7% (rural), respectively.

Table 1 shows descriptive statistics for the full sample as well as for the respective surveys in each country.

Table 1: Descriptive Statistics

Full Sample

Variable	WFS ⁵		DHS	
	Mean	Std. Dev.	Mean	Std. Dev.
Age	28.54	9.81	27.95	9.50
Years of education	2.92	3.83	4.91	4.53
Partner	0.75	0.43	0.65	0.48
Urban residence	0.30	0.46	0.45	0.50
HIV: female prevalence	0.00	0.00	0.05	0.04
HIV: male prevalence	0.00	0.00	0.03	0.03
Number of observations	27,319		37,815	

Cameroon (9 regions)

Variable	WFS 1978		DHS 2004	
	Mean	Std. Dev.	Mean	Std. Dev.
Age	29.84	10.42	27.50	9.52
Years of education	2.48	3.27	6.45	3.58
Partner	0.78	.42	0.65	.48
Urban residence	0.27	.44	0.53	0.50
HIV: female prevalence	0.00	0.00	0.07	.04
HIV: male prevalence	0.00	0.00	0.04	.02
Number of observations	8,112		8,822	

Cote d'Ivoire (Abidjan only)

Variable	WFS 1980		DHS 2005	
	Mean	Std. Dev.	Mean	Std. Dev.
Age	25.83	8.36	27.58	8.62
Years of education	2.47	3.85	6.38	5.29
Partner	0.76	0.43	0.41	0.49
Urban residence	1.00	0.00	1.00	0.00
HIV: female prevalence	0.00	0.00	0.10	0.00
HIV: male prevalence	0.00	0.00	0.04	0.00
Number of observations	1,090		947	

⁵ The HIV prevalence rates are assumed to be zero in the WFS. Even though there likely were some cases of HIV in the late 1970s, HIV AIDS was formerly recognized only in the early 1980s.

Ghana (9 regions)

Variable	WFS 1979		DHS 2003	
	Mean	Std. Dev.	Mean	Std. Dev.
Age	28.20	9.48	29.25	9.63
Years of education	4.14	4.67	5.39	4.68
Partner	0.72	0.45	.65	.48
Urban residence	0.34	0.47	.42	.49
HIV: female prevalence	0.00	0.00	.03	.01
HIV: male prevalence	0.00	0.00	.02	.01
Number of observations	6,107		5,691	

Kenya (7 regions)

Variable	WFS 1977		DHS 2003	
	Mean	Std. Dev.	Mean	Std. Dev.
Age	27.98	9.59	28.05	9.31
Years of education	3.56	3.75	7.47	4.09
Partner	0.70	0.46	.59	.49
Urban residence	0.20	0.40	.34	.47
HIV: female prevalence	0.00	0.00	.09	.04
HIV: male prevalence	0.00	0.00	.05	.03
Number of observations	8,038		7,753	

Senegal (8 regions)

Variable	WFS 1978		DHS 2005	
	Mean	Std. Dev.	Mean	Std. Dev.
Age	28.30	9.50	27.70	9.54
Years of education	0.80	2.22	2.34	3.79
Partner	0.83	0.38	.70	.46
Urban residence	.35	0.48	.43	.50
HIV: female prevalence	0.00	0.00	.01	.01
HIV: male prevalence	0.00	0.00	.00	.01
Number of observations	3,972		14,602	

While the average age is fairly similar across the WFS and DHS surveys, the average years of education have gone up significantly over time, from an average of 2.92 in the WFS to 4.91 in the DHS surveys. This increase largely represents the general upward trend in educational attainment across countries, but also some sampling differences across surveys. To provide a clearer picture of the differences in the samples

used, we compare women of an age to be included both in the WFS and the DHS surveys in Table 2. For example, women who were 20 years of age in 1978 in Cameroon were 46 when the DHS 2004 was administered, making them eligible for inclusion in both the WFS and the DHS. Given a similar sampling approach in both surveys, we would therefore expect fixed variables such as education not to differ between the two waves for this subgroup of women.

Table 2: Comparability of DHS and WFS Samples

Full Sample

Variable	WFS		DHS	
	Ages 20-24		Ages 45-49	
	Mean	Std. Dev.	Mean	Std. Dev.
Age	21.836	1.452	46.713	1.428
Years of education	4.300	4.159	3.480	4.361
Partner	0.779	0.415	0.817	0.387
Urban residence	0.368	0.482	0.390	0.488
T-test of same average education across samples (p-value): 0.000				
Number of observations	5,324		2,672	

Cameroon (9 regions)

Variable	WFS 1978		DHS 2004	
	Mean	Std. Dev.	Mean	Std. Dev.
Age	21.813	1.485	46.795	1.411
Years of education	4.052	3.416	4.573	3.610
Partner	0.830	0.376	0.756	0.430
Urban residence	0.322	0.467	0.435	0.496
T-test of same average education across samples (p-value): 0.000				
Number of observations	1,577		620	

Cote d'Ivoire (Abidjan only)

Variable	WFS 1980		DHS 2005	
	Mean	Std. Dev.	Mean	Std. Dev.
Age	21.847	1.394	46.741	1.507
Years of education	3.256	4.322	7.389	5.963
Partner	.808	.395	0.741	0.442
Urban residence	1.000	0.000	1.000	0.000

T-test of same average education across samples (p-value): 0.000

Number of observations	281	54
------------------------	-----	----

Ghana (9 regions)

Variable	WFS 1979		DHS 2003	
	Mean	Std. Dev.	Mean	Std. Dev.
Age	21.868	1.452	46.649	1.392
Years of education	5.822	4.623	4.491	5.265
Partner	0.759	0.428	.814	.389
Urban residence	0.365	0.482	.355	.479

T-test of same average education across samples (p-value): 0.000

Number of observations	1,216	501
------------------------	-------	-----

Kenya (7 regions)

Variable	WFS 1977		DHS 2003	
	Mean	Std. Dev.	Mean	Std. Dev.
Age	21.847	1.439	46.704	1.434
Years of education	5.053	4.160	4.959	4.520
Partner	0.721	0.449	0.726	0.446
Urban residence	0.288	0.453	0.255	0.436

T-test of same average education across samples (p-value): 0.401

Number of observations	1,492	486
------------------------	-------	-----

Senegal (8 regions)

Variable	WFS 1977		DHS 2003	
	Mean	Std. Dev.	Mean	Std. Dev.
Age	21.807	1.432	46.698	1.449
Years of education	1.280	2.745	1.389	3.071
Partner	0.809	0.394	0.902	0.297
Urban residence	0.389	0.488	0.410	0.492

T-test of same average education across samples (p-value): 0.288

Number of observations	758	1,011
------------------------	-----	-------

On average, the women in the DHS samples are slightly more likely to live in an urban area; there are some differences in the average years of education, which appear to be particularly pronounced in Abidjan (Ivory Coast), where the sample of women in the age group 45-49 is very small (54 women in total).

Overall, the women sampled in the two surveys look fairly similar; given that neither the DHS nor the WFS samples were stratified by age groups, some differences in the sub-samples are unavoidable⁶ and will be further discussed when presenting the main result.

4. Regional Trends in HIV and Fertility

Average total fertility has fallen substantially between the early 1980 and the early 2000s. As Figure 2 shows, total fertility was between 4 and 8 in the WFS surveys, and between 2 and 6 in the latest DHS surveys; the average reduction in the total fertility rates in our sample is very close to 2, with a significant variation both within and across countries.

Quite interestingly, the picture looks quite different when comparing completed (cohort) fertility across time and regions. In both the WFS and the DHS, the average woman aged 45 or older has given birth to around 6 children. This implies that the (unconditional) reproductive behavior of the 1930 birth cohorts (aged 45+ in 1980) was on average very similar to the reproductive behavior of women born in the late 1950s; the changes in total fertility rates observed over the time period thus mostly represent changes in behavior of the younger cohorts.

For the purpose of investigating the effects of HIV on fertility, total fertility is clearly the more interesting measure, since women whose completed fertility we can observe in the last rounds of the DHS had already passed age 30 when the HIV epidemic truly started. The drawback of total fertility rate is that they may in some cases pick up “tempo effects”; total fertility rates may fall or increase in the short run if women change the timing of their birth without changing their desired fertility. In the case of HIV, this

⁶ Both the WFS and DHS are nationally representative for the populations aged 15-49. With some mortality between ages 20-49 and a larger fraction of the sample drawn from the younger cohorts due to their larger relative size, a perfect match between the WFS sub-sample of the 20-24 year old and the DHS sample of the 45-49 old is virtually impossible.

is a non trivial issue, since young mothers may opt for earlier child birth to limit the likelihood of vertical HIV transmission; if this is true, total fertility numbers in HIV regions would be biased upwards relative to the historical numbers as well as final fertility outcomes.

Figure 2: Total Fertility Rate and Completed Fertility: WFS vs. DHS

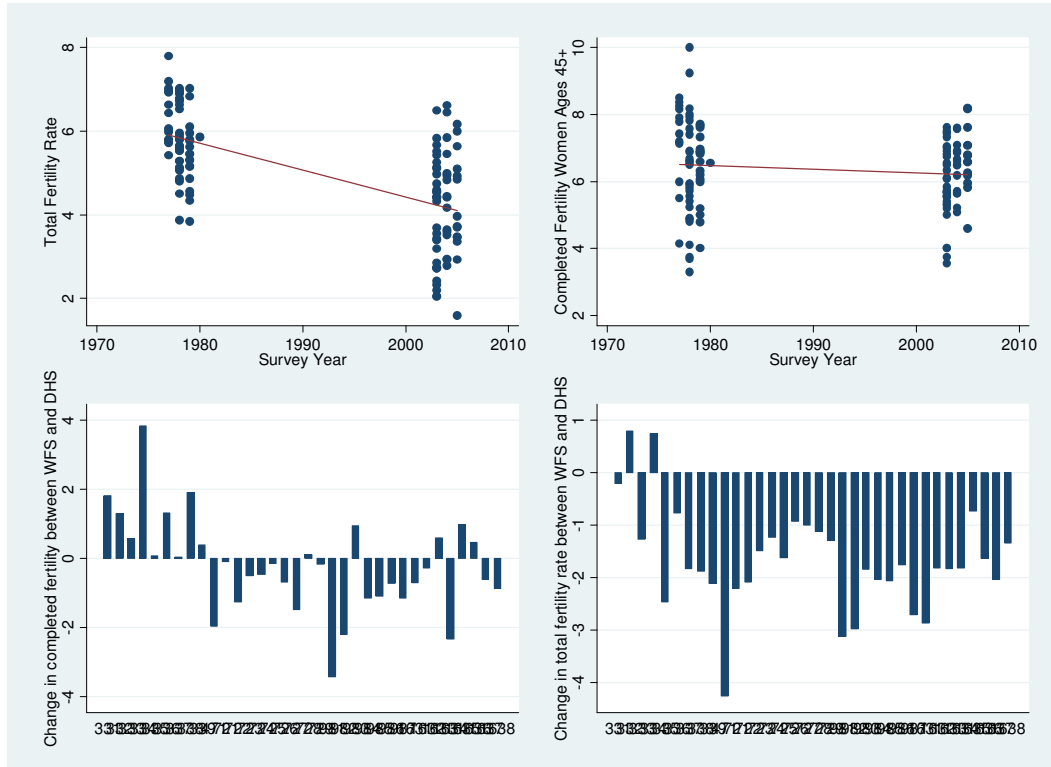


Figure 3 shows the basic correlation between HIV prevalence and changes in fertility rates on the regional level. While the correlation is slightly negative (the correlation coefficient -0.24), HIV prevalence does not appear to have any significant effect on changes in fertility on the regional level once we control for country specific time trends as shown in Table 3 below.

Figure 3: Changes in Fertility and Regional HIV Prevalence Rates

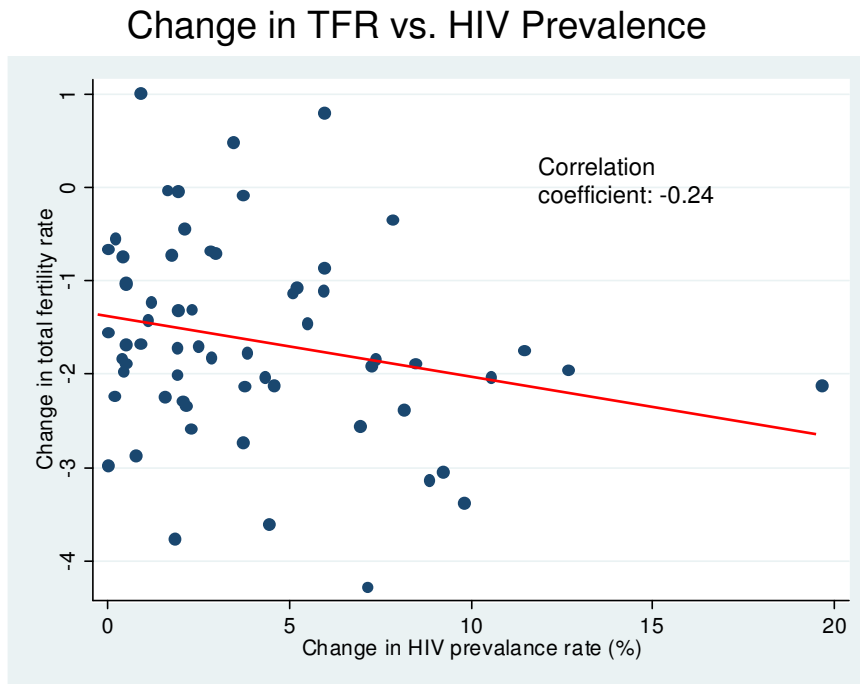


Table 3: Regional Difference-in-Differences: HIV and Total Fertility

Dependent Variable	Total Fertility Rate		
HIV prevalence	-0.0658** (0.028)	-0.0798 (0.057)	-0.0676 (0.045)
Regional fixed effects	YES	YES	YES
Country time trend	NO	YES	YES
Regional time trend	NO	NO	YES
Observations	124	124	124
R-squared	0.85	0.90	0.96

*** p<0.01, ** p<0.05, * p<0.1

5. Empirical Specification and Results

We start our empirical work with a standard model for estimating fertility, and augment it with both the woman's own HIV status (in order to get an idea of the biological effect once the woman has contracted the virus), and with the regional prevalence rate⁷ (in order to measure the fertility response in reaction to the local community HIV prevalence rate). In addition, we include educational status of the woman as an explanatory variable, as it has been found to be a major determinant of fertility. In the context of HIV, education has also been found to influence the behavioral reaction to HIV. For example, de Walque (2007) finds that in the context of an HIV information campaign in rural Uganda, educated women were more responsive to the messages of the campaign, and used condoms more frequently than their less-educated peers. Similarly, Glick and Sahn (2006) find that the education gradient for prevention knowledge is substantial and seems to have increased over time in nine African countries.

In light of this evidence, we also include an interaction term between HIV status and the woman's education. Apart from these variables of interest, we also include variables included in standard fertility regressions, such as the age of the woman and its square, her relationship status, and whether she is living in an urban or rural area.

The main model we would like to estimate is given by

$$fert_{ijkt} = \alpha + \beta hiv_{ijkt} + \gamma educ_{ijkt} + \chi hiv_{ijkt} * educ_{ijkt} + \kappa hiv_{jkt} + \lambda educ_{ijkt} * hiv_{jkt} + \phi X_{ijkt} + post + \delta j + \delta j * post + \varepsilon_{ijkt} \quad (1)$$

where $fert_{ijkt}$ is the fertility measure for woman i in region j in country k period t . For most of the specifications, we look at births of the woman in the last five years⁸. Hiv_{ijkt} is the woman's own HIV status, whereas hiv_{jkt} is the unweighted regional HIV

⁷ Juhn, Kalemli-Ozcan, and Turan (2007) use the cluster HIV rate for this purpose. We prefer the regional HIV rate as on average there are only about 10 women per cluster, resulting in highly variable cluster HIV rates, and feel that such fluctuations are better smoothed out at the regional level.

⁸ As a robustness check, we also looked at the more short-term measure of whether the woman is currently pregnant and found essentially the same results, which are available from the authors.

prevalence rate, which is set to zero in all WFS surveys by default as discussed above. $Educ$ is years of education⁹, and X is a matrix of additional controls. δ_j are regional fixed effects and the $post$ is an indicator equal to 1 if the data come from the DHS surveys, and zero otherwise. To control for the highly heterogeneous political and economic experiences over the last decades, we also use a regional or country specific time trend ($\delta_j * post$) in all our specifications.

In Table 4, we present our main specification. We focus on the impact of the regional HIV prevalence rate on the individual fertility decision of woman in order to investigate the question whether women (conditional on their own HIV status, as in Column 2) react to the HIV prevalence rate in their geographical vicinity. A priori, it is not clear whether they would increase their fertility (for example, in order to reach a certain target number of children in the presence of increased uncertainty about child survival as suggested by Kalemli-Ozcan (2003)) or reduce it, and to what extent this decision would be influenced by the woman's education status. Given that the two papers by de Walque (2007) and Glick and Sahn (2007) posit that more educated women understand the messages from HIV prevention campaigns better and seem to change their sexual behavior more than less-educated women, it seems a reasonable hypothesis that fertility may be reduced more by more educated women, which we test with the interaction term between regional HIV prevalence and the woman's education. In all specifications in Table 4, we control for regional fixed effects that control for regional unobservables that are time-invariant, such as geography, cultural or religious factors, or climate. We also include a time-trend to control for changes in fertility between the observations from the WFS and those in the DHS.

In the first column, the results are presented when including all women irrespective of their own HIV status. The first main finding is that the regional HIV prevalence has a positive and statistically significant effect on average fertility. A 10% HIV prevalence in a region would therefore lead to a reduction in the number of children born in the last five years by about 0.1. Given that the mean value for the fertility variable

⁹ We also try specifications with schooling categories; see Table 5 for further discussion.

is about 1, this corresponds to a proportional reduction (in this case, of 10%) in fertility. Not surprisingly, women with a higher number of years of education have fewer children, although the magnitude of this decrease is not large: for women in our sample, one year more of education leads to a reduction in the children born in the last five years of only .01, or about 1% of the mean value. The second main finding is that the interaction of education with regional HIV prevalence is negative (as hypothesized before) and statistically significant, meaning that women with higher education increase their fertility less in response to HIV than less-educated women. Although the overall response to HIV prevalence found in our sample is an increase in fertility, this response turns negative for women with about 5 years of education. In our sample, about half of the women in the DHS survey have more than five years of education. This result then indicates a heterogeneous response to the HIV epidemic, with a quite substantial education gradient that previous studies on the impact of HIV on fertility have neglected. The other control variables included confirm the findings in previous fertility studies: older women have higher fertility but at a decreasing rate, married women have more children, and women living in urban areas have somewhat fewer children than people living in the countryside.

In the second column, we restrict our sample to women who are HIV-negative in order to arrive at a purely anticipatory behavioral effect (the sample used to arrive at the results in column 1 include also HIV-positive women who may see their fertility change for biological reasons or due to reactions about learning her own HIV-status). For this restricted sample, we lose about 800 observations (as about 5% of the roughly 18,000 women in the DHS sample have tested positive for HIV), whereas the results remain virtually of the same magnitude and statistical significance. Given that the DHS does not provide the tested individuals with the results of the HIV test, this result is not too surprising (if we assume, for example, that private testing is not available, or if we assume that the same fraction of HIV-positive and HIV-negative women have learned about their HIV status from sources other than the DHS).

In the third column, we further restrict the sample to women aged 40 or younger. The reason for this restriction is that women over this age were not exposed to the HIV

epidemic for a large part of their fertile period. Women older than 40 in the DHS would have been experiencing the effects of HIV at a time when most of her fertility is already completed, and the number of children born to her in the last five years is probably close to zero, irrespective of the regional HIV prevalence. For this reason, younger women who are in the prime of their fertile period and are currently facing the full effect of the HIV epidemic provide more information. This sample restriction leads to a significant increase in the magnitude of the coefficient for regional HIV prevalence, which now almost doubles in size. This means that for this sample, a 10% regional HIV prevalence translates into an increase in the number of children born in the last five years by .25 children, or roughly a more than 25% increase compared to the mean value found in the sample. Also the interaction term of mother's education with regional HIV prevalence increases in magnitude. The turning point at which the fertility response turns negative from positive is at about 7 years, in contrast to the 4-5 years found previously. In column 4 we further restrict the sample to women age 18 and over, as for women younger than that the fertility measure could be rather noisy. However, the results change only marginally as a result of this restriction.

In the last column, we restrict the sample to a cross-section of DHS surveys only, in order to get an idea of the potential bias when not controlling for regional fixed effects, and in order to facilitate comparison with the literature such as Juhn, Kalemli-Ozcan, and Turan (2007) who use the DHS data without controlling for regional fixed effects. The main finding for this cross-sectional sample is that the results are very similar to the results when including regional fixed effects. The coefficients for age are virtually identical, whereas education has increased in significance as a determinant of fertility. A similar result holds for the urban dummy: women living in a city have now significantly lower fertility than their rural counterparts. For our main variable of interest, the regional HIV prevalence rate, the coefficient in this cross-section is of the same magnitude as when including the WFS in the sample. The coefficient on the interaction term between mother's education and regional HIV prevalence is about 30% smaller than before. It is remarkable, however, how close the results for the cross-section sample and the sample combining WFS and DHS are.

Table 4: Multivariate Analysis: Children born and HIV prevalence

Dependent Variable: Number of children born in the last 5 years					
	(1)	(2)	(3)	(4)	(5)
Age	0.220*** (0.00808)	0.221*** (0.00816)	0.317*** (0.0107)	0.313*** (0.0102)	0.285*** (0.0122)
Age squared	-0.00369*** (0.000129)	-0.00372*** (0.000130)	-0.00558*** (0.000189)	-0.00552*** (0.000169)	-0.00506*** (0.000224)
Years of education	-0.0130*** (0.00191)	-0.0129*** (0.00193)	-0.0128*** (0.00202)	-0.0128*** (0.00223)	-0.0189*** (0.00323)
Married	0.583*** (0.0268)	0.588*** (0.0274)	0.608*** (0.0290)	0.636*** (0.0327)	0.572*** (0.0409)
Urban	-0.104*** (0.0166)	-0.101*** (0.0163)	-0.107*** (0.0183)	-0.133*** (0.0217)	-0.224*** (0.0314)
Regional HIV prevalence	1.113* (0.616)	1.297* (0.658)	2.517*** (0.800)	2.798*** (0.945)	2.377*** (0.816)
HIV * education	-0.225*** (0.0488)	-0.242*** (0.0531)	-0.349*** (0.0562)	-0.375*** (0.0606)	-0.254*** (0.0556)
DHS sample	-0.0984*** (0.0359)	-0.0963** (0.0356)	-0.106** (0.0405)	-0.134*** (0.0485)	-
Constant	-2.222*** (0.109)	-2.253*** (0.110)	-3.401*** (0.145)	-3.322*** (0.147)	-2.901*** (0.161)
Sample restrictions	None	HIV neg.	HIV neg. age < 41	HIV neg. 17 < age < 41	HIV neg. 17 < age < 41 DHS only
Observations	45180	44367	38124	32133	12209
R-squared	0.304	0.306	0.323	0.224	0.269

*** p<0.01, ** p<0.05, * p<0.1

Robust standard errors in parentheses

In Table 5, we present some robustness checks and also include the woman's own HIV status in the regression, as indicated in equation (1). In columns 1 and 2, we include a country specific time trend in the regression in order to control for heterogeneous country experiences with family planning programs and other country-wide effects. While the results remain very similar to those in the previous Table, the inclusion of a country time trend leads to the coefficient on the regional HIV prevalence rate to turn insignificant. As expected, HIV-positive women experience a decline in fertility either for biological reasons or due to an ex-post response to learning about their HIV status. The results change only marginally when own HIV status is omitted in column 2.

Columns 3 and 4 of Table 5 repeat the specification in columns 1 and 2 but exclude the regions with the lowest HIV prevalence rates (under 1%) and with the highest HIV prevalence rates (over 7%) to test whether the results are robust to the exclusion of these extreme values. When focusing on regions with moderate HIV prevalence rates, the magnitude of the coefficient on regional HIV prevalence more than doubles in magnitude, indicating that the results are by no means driven by regions with very low or very high HIV prevalence rates, whereas the coefficient estimate for the interaction between education and HIV remains about the same size. As before, the results remain the same when excluding own HIV status.

One of the questions unanswered in Table 4 is the effect of different education levels on fertility in the context of HIV, as education is entered linearly into the regression above. For this reason, we split education by level of attainment in columns 5 and 6 of Table 5. The categories used are “no schooling”, which serves as the omitted group, “some primary”, “primary completed”, “some secondary”, “secondary completed”, and “tertiary”. This stratification allows us to investigate in more detail which level of schooling is important in determining fertility in general, and when interacted with regional HIV status. In addition, this stratification also allows best-possible comparability of this variable between the WFS and DHS. When entering the education variable this way, we find that up to primary education there is little evidence on a negative impact on fertility, the sign of the coefficient begins to turn for women having some secondary education, and becomes significant for completed secondary education and tertiary education. A similar result holds when interacting the education categories with regional HIV prevalence. The results show that only women with completed secondary or tertiary education respond differently to the HIV prevalence than non-educated women.

Table 5: Robustness Checks

Dependent Variable: Number of children born in the last 5 years						
<i>Sample</i>	All (1)	HIV neg. (2)	All (3)	HIV neg. (4)	All (5)	HIV neg. (6)
Age	0.220*** (0.00786)	0.221*** (0.00789)	0.219*** (0.0119)	0.220*** (0.0124)	0.223*** (0.00819)	0.224*** (0.00825)
Age squared	-0.00370*** (0.000127)	-0.00371*** (0.000128)	-0.00371*** (0.000207)	-0.00374*** (0.000214)	-0.00372*** (0.000133)	-0.00373*** (0.000134)
Years of education	-0.0128*** (0.00174)	-0.0126*** (0.00175)	-0.0171*** (0.00304)	-0.0165*** (0.00312)		
Married	0.580*** (0.0258)	0.586*** (0.0264)	0.491*** (0.0413)	0.504*** (0.0439)	0.576*** (0.0262)	0.582*** (0.0268)
Urban	-0.114*** (0.0159)	-0.112*** (0.0157)	-0.207*** (0.0268)	-0.201*** (0.0274)	-0.114*** (0.0173)	-0.112*** (0.0170)
HIV status	-0.121*** (0.0264)		-0.139*** (0.0356)		-0.128*** (0.0257)	
HIV regional preval.	1.351 (1.095)	1.384 (1.070)	3.186*** (0.585)	3.521*** (0.581)	0.317 (0.555)	0.343 (0.591)
HIV * education	-0.220*** (0.0404)	-0.235*** (0.0437)	-0.198** (0.0692)	-0.226*** (0.0714)		
DHS wave	0.0574 (0.0928)	0.0612 (0.0903)			-0.116*** (0.0260)	-0.115*** (0.0255)
<i>Schooling categories</i>						
Some primary					0.0655 (0.0394)	0.0654 (0.0393)
Primary completed					0.0251 (0.0224)	0.0265 (0.0225)
Some secondary					-0.0288 (0.0245)	-0.0271 (0.0244)
Second. completed					-0.116*** (0.0281)	-0.112*** (0.0278)
Tertiary					-0.323*** (0.0372)	-0.324*** (0.0366)
<i>HIV interactions</i>						
Some primary					0.0572 (0.964)	0.292 (1.033)
Primary completed					0.463 (0.595)	0.475 (0.631)
Some secondary					-0.297 (0.590)	-0.398 (0.622)
Second. completed					-1.738** (0.801)	-1.869** (0.835)
Tertiary					-1.540** (0.688)	-1.661** (0.739)
Constant	-2.302*** (0.104)	-2.319*** (0.105)	-2.199*** (0.129)	-2.238*** (0.136)	-2.413*** (0.0994)	-2.433*** (0.0997)
Observations	45180	44367	10822	10405	45180	44367
R-squared	0.308	0.310	0.327	0.335	0.302	0.304

Column 1 and 2 include a country specific time trend. In columns 3 and 4, the sample is restricted to regions with prevalence rates above 1 and below 7% (DHS only). In columns 5 and 6, educational categories rather than years of schooling are used.

*** p<0.01, ** p<0.05, * p<0.1, Robust standard errors in parentheses

One cautionary note must be made concerning the interpretation of these results. The inclusion of the WFS prevents us from creating a wealth index (there is no income or expenditure information in the DHS, therefore the common approach in papers using the DHS is to create a wealth index using principal component analysis), therefore we cannot separate the direct impact of education on fertility (for example through changes in preferences or because the content of fertility or HIV information campaigns is more easily processed) from an income effect as educated women have often been found to also have higher socioeconomic status. For example, women with relatively high education may be married to a man with high education and high income (or they themselves may have a higher-paying job), confounding the knowledge-aspect of education with a wealth effect; education may thus proxy for both knowledge and wealth.

6. Conclusion

In this paper we evaluate the effect of HIV prevalence on fertility in Sub-Saharan Africa using a new data set which combines existing DHS data with historical data from the World Fertility Surveys. The main result emerging from this paper is that a weak and statistically insignificant correlation between HIV prevalence and fertility on the regional level hides important heterogeneity across socioeconomic groups within and across regions. We find a positive fertility response to the regional HIV prevalence for non-educated mothers and mothers with primary schooling, but a negative fertility response for mothers with completed secondary schooling and higher. An intuitive explanation for these results in the light of the theoretical discussion presented in Section 2 of this paper as well as empirical evidence discussed in the introduction, is that more highly educated people better understand the risks and costs associated with HIV (De Walque, 2007), and thus adjust their fertility more than people with lower human capital. An alternative explanation is that people with different educational attainment have different discount rates resulting in differing fertility decisions.

The results in this paper indicate that HIV - as a large negative health shock for some of regions in our sample - induces significant fertility changes in the affected

populations, leading to significantly larger family sizes in the poorly educated subgroups, thus potentially delaying the demographic transition and the economic benefits associated with it.

References

- Becker, G. S. (1960). An Economic Analysis of Fertility. Demographic and Economic Change in Developed Countries. G. S. Becker. Princeton, N.J., Princeton University Press.
- Becker, G. S. (1965). "A Theory of the Allocation of Time." Economic Journal **75**: 493-517.
- Becker, G. S. and H. G. Lewis (1973). "On the Interaction between Quantity and Quality of Children." Journal of Political Economy **81**: S279-S288.
- Cuddington, J. T. (1993). "Modeling the Macroeconomic Impact of the AIDS in Tanzania." World Bank Economic Review **7**(2): 173-189.
- De Walque, D. (2007). "How Does the Impact of an HIV/AIDS Information Campaign Vary with Educational Attainment? Evidence from Rural Uganda." Journal of Development Economics, **84**(2): 686-714.
- Glick, P. and D. E. Sahn (2007). "Changes in HIV/AIDS Knowledge and Testing Behavior in Africa: how much and for whom?" Journal of Population Economics **20**: 383-422.
- Gray, R. H., M. J. Wawer, D. Serwadda, N. Sewankambo, C. Li, F. Wabwire-Mangen, L. Paxton, N. Kiwanuka, G. Kigozi, J. Konde-Lule, T. C. Quinn, C. A. Gaydos and D. McNairn (1998). "Population Based Study of Fertility in Women with HIV-1 Infection in Uganda." Lancet **351**(9096): 98-104.
- Juhn, C., S. Kalemli-Ozcan and B. Turan (2008). "HIV and Fertility in Africa: First Evidence from Population Based Surveys." NBER Working Paper **14248**.
- Kalemli-Ozcan, S. (2006). "AIDS, Reversal of the Demographic Transition and Economic Development: Evidence from Africa." NBER Working Papers **12181**.
- Mahy, M. (1999). Perceptions of Child Mortality and their Effects on Fertility Intentions in Zimbabwe. Baltimore, Maryland, Johns Hopkins University.
- Oster, E. (2007). "HIV and Sexual Behavior Change: Why Not Africa?" NBER Working Paper **13049**
- Schultz, T. P. (1997). Demand for Children in Low Income Countries. Handbook of Population and Family Economics. M. R. Rosenzweig and O. Stark. Amsterdam, North-Holland. **1A**: 643.
- UNAIDS (2008). 2008 Report on the Global AIDS Epidemic, UNAIDS.
- Willis, R. (1973). "A new Approach to the Economic Theory of Fertility Behavior." Journal of Political Economy **81**: S14-S64.
- World Bank (2007). World Bank Development Indicators CD-ROM.
- Young, A. (2005). "The gift of the dying: The tragedy of AIDS and the welfare of future African generations." Quarterly Journal of Economics **120**(2): 423-466.

