HIV/AIDS AND FERTILITY

Jane G. Fortson*
Becker Center on Chicago Price Theory
University of Chicago

This Version: December 2008
First Draft: September 2008

Abstract

This paper studies the response of fertility to the HIV/AIDS epidemic in sub-Saharan Africa. Standard models of fertility have ambiguous predictions in this context. Because childbirth requires unprotected sex (which increases the risk of contracting HIV), we might expect HIV/AIDS to reduce fertility through an infection avoidance motive. Adding to this decline, HIV infection may itself reduce fecundity among infected women. However, because HIV/AIDS affects the expected longevity of children, a quantity-quality model of childbearing would predict an increase in fertility. I use repeated cross-sections of the Demographic and Health Surveys for twelve countries in sub-Saharan Africa to examine this question empirically. Using individual birth histories from these data, I construct estimates of the regional total fertility rate over time. In a difference-in-differences approach, I compare regional HIV prevalence to changes in total fertility rates from the late 1980s to the present. My results suggest that HIV/AIDS had very little impact on fertility, both overall and in a sample of HIV-negative women.

* *Email:* jfortson@uchicago.edu
Address: Becker Center on Chicago Price Theory, 5807 S. Woodlawn Avenue, Chicago, IL 60637

Anne Case, Ken Fortson, Jens Ludwig, and seminar participants at the University of Chicago provided helpful suggestions. I am grateful to Daniel Hirschberg and Kristopher Hult for research assistance, and to Bridgette James for help with the DHS data.

I. INTRODUCTION

In sub-Saharan Africa, HIV has led to numerous changes over the past several decades. Most prominently, AIDS-related mortality has, since the early 1980s, driven declines in life expectancy in a number of countries, particularly in southern Africa. Today, adult HIV prevalence in sub-Saharan Africa is estimated to be about five percent, with significantly higher prevalence in certain areas (UNAIDS and WHO, 2007). In some regions (particularly in southern Africa), more than one fifth of prime-age adults are infected with HIV, and an even greater share will, at some point, become infected. Until very recently, treatment was unavailable for virtually all patients, and some communities have been ravaged by the disease. These dramatic changes in morbidity and mortality have imposed an enormous cost on sub-Saharan Africa; understanding their effects is important in designing policies to mitigate the adverse effects of HIV.

Past research has looked not only at the effects of HIV on sexual behavior (e.g., Oster, 2007) and human capital investment (e.g., Fortson, 2008b), but also economic growth (e.g., Werker, Ahuja, and Wendell, 2006; Young, 2005). One important input to models predicting the long-run economic effects of HIV/AIDS (such as Young, 2005) is an estimate of the change in population size, which is determined not only by mortality but also fertility. Theoretical models of fertility have ambiguous predictions in this setting. Infection avoidance behavior and the biological effects of HIV would predict declines in fertility in response to the epidemic; however, a quantity-quality model of childbearing (e.g., Becker and Lewis, 1973) would suggest increases in fertility in response to the increased mortality risk faced by offspring. Several studies have attempted to estimate the effect of the HIV/AIDS epidemic on fertility in sub-Saharan Africa; however, there is little consensus on the sign or size of the effect. Young (2005) shows that, in South Africa, fertility has fallen in response

to HIV, and, in subsequent work (Young, 2007), confirms this decline in a broader sample of countries. However, Kalemli-Ozcan (2008) finds a positive relationship between HIV and fertility in a cross-sectional, country-level framework. And more recent cross-sectional work by Juhn, Kalemli-Ozcan, and Turan (2008) finds that community level prevalence is not related to fertility among uninfected women.

This paper attempts to resolve this debate by constructing better estimates of fertility over time and linking these estimates to newly-available data on HIV prevalence. The Demographic and Health Surveys (DHS), the source of most available fertility data, are conducted every few years at best. Most researchers have used these data to calculate fertility for the years immediately prior to the survey. As a result, estimates of fertility in the interim years are either imputed or sparse. However, the DHS contains full birth histories for women in the sample. Using a larger share of available birth history information from these surveys, I construct estimates of fertility over time to minimize gaps and imputations. This methodology, which past work has not employed, has the advantage that it allows me to construct a long time-series of fertility rates, eliminating the need to rely solely on cross-sectional variation.[1]

My analysis also takes advantage of newly-available data on HIV prevalence, also from the DHS. These HIV data, the result of population-based testing, have several advantages over past data. First, the estimates are considerably more accurate than past

[1] Because there are strong geographic and economic patterns in both fertility and HIV, cross-sectional studies must adopt an instrumental variables approach. Because many of the factors associated with HIV are also independently related to fertility rates, finding a convincing instrument for HIV has proven to be challenging. However, HIV is arguably exogenous with respect to *trends* in fertility (see Section V), which suggests that a difference-in-differences approach may not be subject to the same concerns.

estimates of prevalence, which were largely based on antenatal clinic testing.[2] Second, these data can be linked to fertility information at the individual, as well as at the regional, level.

Using a difference-in-differences approach, I find that areas with higher levels of HIV saw neither substantial increases nor decreases in fertility over time. I test whether this negligible effect is the result of offsetting effects on HIV-positive and HIV-negative women. While I find some evidence that HIV infection is associated with lower fertility at the individual level, this negative effect is not obscuring a positive behavioral effect among uninfected women. In other words, there is no evidence that living in an area with high levels of HIV has an effect on the fertility of those who are not directly affected (i.e., infected) by the virus. Robustness checks suggest that my results are not driven by sample attrition due to mortality or migration. I find evidence of parallel trends in fertility during the 1980s across areas with eventually-different levels of HIV; the absence of pre-period differences should assuage concerns about the exogeneity of HIV with respect to fertility trends, as well as about the sensitivity of my findings to transformations of the fertility rate. My results rule out effects of the magnitude found by other studies and suggest that the effect of HIV, relative to the overall (negative) trend in fertility over time, is quite small. This evidence casts doubt on the hypothesis that large fertility-driven declines in population will generate future improvements in wellbeing in the countries most affected by HIV. This paper likewise suggests that changes in mortality risk – here, the combined effects of short-run risk related to sexual behavior and long-run risk for offspring – have a relatively small effect on fertility.

[2] Antenatal clinic testing has the limitation that the sample of HIV test participants is neither geographically nor demographically representative.

II. DATA

My analysis uses data from the Demographic and Health Surveys (DHS) for twelve countries in sub-Saharan Africa: Cameroon, Côte d'Ivoire, Ethiopia, Ghana, Kenya, Malawi, Mali, Niger, Rwanda, Tanzania, Zambia, and Zimbabwe.[3] The DHS are nationally-representative household surveys that include information about demographic characteristics and health, as well as detailed birth histories. In a number of countries, a recent cross-section of the survey included the results of HIV testing. My sample includes the twelve countries for which there is both DHS HIV testing data and for which earlier cross-sections of the data can be linked by region. For the most part, my analysis uses these individual-level data to calculate measures of prevalence and fertility at the region level.[4] Additional detail about the surveys, including sampling information and response rates, is available in the data appendix.

Total fertility rates over time are calculated using birth histories reported by women ages 15-49. Rather than calculating an estimate of fertility at the time of the survey (which is the norm), I calculate fertility separately for each of the ten years prior to the survey. However, in order to estimate fertility for women in a consistent age range in each year, this necessitates calculating fertility among women ages 15-39 in each of these years. For example, a woman who is 20 in the survey year is included in the calculation sample for the five years prior to the survey, but not for the five years previous to that (when she would have been under age 15). Likewise, a woman who is 49 in the survey year is excluded from the calculation sample in all but the year ten years prior to the survey, when she was 39. In Section V, I show that fertility among 40- to 49-year-old women is relatively low and should

[3] Source: DHS datasets, www.measuredhs.com, MEASURE DHS, Macro International Inc.
[4] That is, I divide each country into geographic locales that are defined by administrative divisions (e.g., regions). Across the 12 countries in my primary sample, there are 109 regions.

not affect my estimates. In years when the DHS cross-sectional data have overlapping estimates (i.e., when survey years are fewer than ten years apart), I pool estimates from both cross-sections. The total fertility rate is calculated as follows:

$$tfr_{r,t} = 5\sum_{i=1}^{5} \frac{babies_{r,t,5(i+2),5(i+2)+4}}{exposure_{r,t,5(i+2),5(i+2)+4}}, \quad (1)$$

where $babies_{r,t,5(i+2),5(i+2)+4}$ is a measure of babies born in year t to women in the $5(i+2)$ to $5(i+2)+4$ agegroup living in region r. $exposure_{r,t,5(i+2),5(i+2)+4}$ is cumulative years spent in the $5(i+2)$ to $5(i+2)+4$ agegroup during year t among women living in region r. $\frac{babies_{r,t,5(i+2),5(i+2)+4}}{exposure_{r,t,5(i+2),5(i+2)+4}}$ is the age-specific fertility rate for the $5(i+2)$ to $5(i+2)+4$ agegroup.[5] The total fertility rate, which is an aggregation of the age-specific fertility rates, can be thought of as an estimate of the number of children a woman will have over her lifetime based on contemporaneous age-specific fertility rates.

Using the HIV testing data from each of the twelve countries in my sample, I calculate summary statistics for regional HIV prevalence. Table I.A shows that HIV prevalence is highest in Southern Africa, particularly in Malawi, Zambia, and Zimbabwe. Table I.B, which presents summary statistics of fertility rates calculated from multiple waves of DHS data, shows that fertility rates are generally higher in West Africa, particularly in Mali and Niger.

Figure 1 shows trends in fertility rates across countries over time. Whereas in the past these data would have been used to calculate only a handful of estimates for each country, my method of calculating historical fertility rates generates between 15 and 24 years

[5] If women did not straddle agegroups over the year, the age-specific fertility rate would simply be the number of babies born to women in a given agegroup in a given year in a given region, divided by the number of women in that group.

of data per country. These data show that fertility rates fell over the period; consistent with Table I.B, fertility rates were highest in Mali and Niger and lowest in Zimbabwe.

III. RESULTS

My empirical approach combines region-level data on fertility trends over time with cross-sectional, region-level data on HIV prevalence, and aims to test whether trends in fertility differed across areas with different levels of HIV. Figure 2 shows the relationship between regional HIV prevalence (estimated between 2001 and 2006, depending on the country) and the change in fertility over time. Each dot represents a region, where the y-position is determined by the change in fertility between the 1980s and 2000s. The scatterplot shows that, in general, fertility fell between the 1980s and the 2000s. However, the graph shows no apparent systematic relationship between the change in fertility and regional HIV prevalence.

In Table II, I test whether there is a statistically significant relationship between regional HIV prevalence and changes in fertility rates, and whether this is robust to the inclusion of controls. I adopt a difference-in-differences approach, comparing the difference in pre-1990 fertility and post-2000 fertility across areas with different levels of HIV. In the simplest case, I use the following specification:

$$tfr_{r,t} = \beta_0 + \beta_1 HIV_r \times Post_t + \beta_2 HIV_r + \beta_3 Post_t + \varepsilon_{r,t}, \quad (2)$$

where t has just two periods and β_1 is the coefficient of interest. Standard errors are clustered on the region, and each region is weighted by its population.[6,7] This approach

[6] Because there may be serial correlation in fertility, a difference-in-differences model may have a tendency to over-reject the null hypothesis of no effect. Clustering the standard errors on the region is one solution proposed by Bertrand, Duflo, and Mullainathan (2004).

[7] It could be argued that population-weighting is not sensible in this setting, if each region provides equally-good information about the effect of HIV on fertility. However, in practice, weighting seems to have little

implicitly assumes that HIV had a constant effect on fertility after 2000, and no effect on fertility before 1990.[8] Because HIV prevalence in sub-Saharan Africa was very low in the early 1980s, and because survival in the absence of treatment is thought to be about ten years, most of the mortality associated with HIV occurred post-1990. Therefore, there was likely only a minimal behavioral response of fertility to HIV-related mortality risk prior to 1990. Likewise, much of the rise in HIV prevalence occurred in the late 1980s and early 1990s; prevalence and mortality since 2000 has been comparatively flat. Together, these patterns suggest that the assumptions underlying this approach are not unreasonable. The difference-in-differences estimate, β_1, following equation (2) – shown in Table II, column (2) – is small and insignificant, and suggests that changes in fertility are unrelated to HIV.

While this simple difference-in-differences approach is intuitively appealing, it makes sense to control for cross-sectional differences in fertility, as well as flexibly account for trends in fertility. In particular, in my preferred specification, I use the following model:

$$tfr_{r,t} = \beta_0 + \beta_1 HIV_r \times Post_t + \alpha_r + \eta_{n,t} + \varepsilon_{r,t}, \quad (3)$$

where each observation is a region-year; there are between three and twenty regions per country. α_r is a region fixed effect, and $\eta_{n,t}$ is a nation-by-year fixed effect, and β_1 is the coefficient of interest. The region fixed effects allow for level differences across areas in the fertility rate (i.e., these include the main effect of HIV). The nation-by-year fixed effects allow for separate trends in fertility over time for each country, and likewise capture the main effect of time.

---

effect; results which weight each region equally yield estimates that are not appreciably different than the results presented here.

[8] Though there are fertility estimates in the interim years, it is more difficult to pin down estimates of HIV between 1990 and 2000. These results are, however, robust to alternate specifications (discussed in Section V) which use fertility data from 1991-1999 and impute HIV prevalence over the period.

The results, shown in Table II, show that the difference-in-differences estimate is small and insignificant, particularly after accounting for cross-sectional variation in both fertility and HIV (with region fixed effects). My preferred specification (column (6)), which includes country-by-year fixed effects, shows a small and insignificant difference-in-differences estimate. In particular, the point estimate implies that, relative to areas without HIV, areas with prevalence of five percent saw an increase in fertility of 0.01 children between the late 1980s and 2000s – a change which cannot be distinguished from zero.

The 95-percent confidence interval around this estimate is -4.45 to 5.01. This suggests that, for an area with prevalence of five percent (which is estimated 2007 prevalence across sub-Saharan Africa), we can, at the five percent level, rule out effects on fertility smaller than -0.22 and larger than 0.25. This means that, over her lifetime, a woman in an area with prevalence of five percent will have no more than 0.25 more (and no more than 0.22 fewer) children, relative to a woman in an area without HIV. For comparison, Young (2007) finds that prevalence of this magnitude is associated with a ten percent decline in fertility; my estimate and the associated 95-percent confidence interval rule out a decline in fertility of more than four percent. Kalemli-Ozcan (2008) finds that "a country with a high level of HIV/AIDS prevalence, such as Zambia, has one more child per woman on average compared to a country with a low level of HIV/AIDS prevalence, such as Senegal." In my HIV testing data, prevalence in Zambia is 15.6 percent, and prevalence in Senegal is 0.7 percent; therefore, her estimates imply an increase in fertility of 0.34 for an area with prevalence of five percent – an increase that is ruled out by my estimates.

My results likewise suggest that the HIV/AIDS epidemic has had a much smaller effect on fertility than other significant public health events. Bleakley and Lange (2007), using evidence from hookworm eradication in the American South, find that a one standard

deviation decline in hookworm rates is associated with a decline in fertility of roughly 0.15 children. My point estimate of 0.281 suggests a substantially smaller increase in fertility (0.01) for a one standard deviation increase in HIV rates. Finally, relative to the precipitous declines in fertility in sub-Saharan Africa over the past several decades, even the tails of the 95 percent confidence interval are quite small. These results imply that HIV will have a negligible role in influencing fertility trends, neither reversing nor accelerating the demographic transition.

IV. CHANNELS

In Table II, I test, in a difference-in-differences approach, whether areas with higher levels of HIV experienced significantly larger relative increases or decreases in fertility between the 1980s and the present. I find that there is no significant difference in fertility trends between areas with high and low levels of HIV. However, it is possible that this overall null effect obscures significant effects through particular channels. For instance, it is possible that, for physiological reasons, HIV leads to declines in fertility, but that, for behavioral reasons, HIV leads to increases in fertility. This could lead us to find no effect of HIV on fertility overall, but not because women (and their partners) do not respond to the HIV/AIDS epidemic. To disentangle some of the effects through physiological and behavioral channels, I make use of recent DHS data in which individual survey results can be linked to HIV testing results. HIV-negative women should not experience any physiological effects of HIV, but could experience some behavioral effects. In particular, we might expect to see changes in fertility among uninfected women in response to infection risk or in response to changes in the life expectancy of offspring. Estimating these effects may help us understand how fertility responds to mortality risk, which is of general interest. However,

even if I find effects through particular channels, this does not change the fact that the overall effect of HIV on fertility – which is important for understanding the effects of HIV on population size – is negligible.

In Table III, I test whether fertility rates differ between HIV-positive and HIV-negative women. I use the most recent cross-section of data from my main sample along with recent DHS data from Burkina Faso, Guinea, Lesotho, and Senegal.[9] For these thirteen cross-sections, I can calculate total fertility rates separately for HIV-positive and HIV-negative women.

HIV infection is associated with lower fertility (top panel), consistent with existing literature suggesting that there are physiological effects of HIV on fertility (for example, Lewis et al., 2004, Gray et al., 1998).[10] Lower fertility among HIV-positive women may also reflect behavioral responses, if HIV-positive women know their status or perceive themselves to be at higher risk for HIV. For instance, HIV-positive women may be concerned about passing on the virus to their children, or about the care of their children in case of illness or death. Regardless of the channel, fertility is lower among HIV-positive women. Therefore, we might wonder whether the overall null results reflect offsetting effects among HIV-positive and HIV-negative women. In particular, we might think that the effect of HIV among HIV-negative women is positive, but that the inclusion of HIV-positive women masks this. To examine this empirically, I calculate total fertility rates among women known to be HIV-negative. However, because I can only link HIV test results to birth histories for the most recent cross-section from each country, I can only calculate total fertility rates among HIV-negative women for the ten years prior to the most

[9] This analysis excludes Mali, Tanzania, and Zambia, for which individual HIV test results cannot be matched to birth histories.

[10] Contemporaneous work by Juhn, Kalemli-Ozcan, and Turan (2008) using DHS data likewise finds that fertility is lower among HIV-positive women.

recent survey (and not before then).[11] I use these ten years of data on fertility among the HIV-negative and continue to use fertility data from earlier years from a mixed sample (i.e., a sample including both HIV-positive and HIV-negative women) to fill out the time-series of fertility estimates. I then test, on this new set of fertility estimates, whether areas with higher levels of HIV experienced larger relative increases in fertility. The results, shown in Table III (bottom panel), suggest that HIV is associated with neither a rise nor a decline in fertility among HIV-negative women. In particular, the difference-in-differences estimate from the preferred specification (Table III, bottom panel, column (4)) implies that, relative to regions without HIV, regions with prevalence of five percent saw a decline in fertility among HIV-negative women of 0.01 children between the late 1980s and 2000s. This suggests that the effect of HIV among uninfected women is small and insignificant.

Nevertheless, we might be concerned that this robustness check is itself biased. Because early cross-sections of the DHS do not include HIV testing, I cannot distinguish between HIV-negative and HIV-positive women in constructing estimates of fertility using data from earlier waves. Therefore, fertility estimates from the early (pre-1990) sample include both infected and uninfected women. While HIV rates were relatively low in the pre-1990 period, we might nevertheless be concerned that our estimates of overall fertility from this period are not a good measure of fertility among uninfected women. In particular, because the data from earlier years include some HIV-infected women who are likely to have lower fertility, we might think that, in early years, fertility estimates in high HIV areas could understate true levels of fertility among uninfected women. Therefore, the mixed sample composition in early years would lead me to understate the decline in high HIV areas relative to low HIV areas, positively biasing the difference-in-differences estimate. In principle,

[11] Because HIV prevalence was low in the 1980s, fertility estimates for these years based on data from all women should be a reasonably good approximation of fertility among HIV-negative women.

excluding HIV-positive women in early years – if we could do it – would reduce the difference-in-differences estimate. However, since both the overall effect of HIV (Table II) and the estimated effect among uninfected women (Table III, bottom panel) – which should bound the true effect of HIV among uninfected women – are small and insignificant, the size of any bias must be quite small.

V. THREATS TO VALIDITY

Sample Size Concerns.

My primary empirical strategy is to compare changes in fertility (pre-1990 v. post-2000) to current levels of HIV prevalence. This approach implicitly assumes that HIV was negligible in the pre-1990 period, consistent with existing evidence on HIV prevalence (e.g., Timæus and Jasseh, 2004; Oster, 2008). However, my approach has the shortcoming that it uses only a subsample of the available information on fertility rates. In particular, my method of constructing estimates of fertility over time allows me to estimate fertility between 1991 and 1999 as well, though my main results do not use these estimates. With nine additional years of data on fertility rates, I would be able to improve precision. However, doing so requires me to make assumptions about the time path of HIV prevalence – that is, how it changed over the 1990s.

Here, I test whether, by making reasonable assumptions about the time path of HIV over the 1990s, we can more precisely pin down estimates of the effect of HIV on fertility. In particular, I compare results from three approaches. In the first, I again adopt a difference-in-differences approach, but include data from the years 1991-1999. (This is equivalent to the specification in column (6) of Table II, except that I include data between 1990 and 2000, which Table II did not.) I look at the difference-in-differences between pre-

and post-2000, which assumes that fertility between 1991 and 1999 was unaffected by HIV. Next, I test for a relationship between HIV and the total fertility rate assuming that HIV has grown linearly from 1980. In particular, I estimate

$$tfr_{r,t} = \beta_0 + \beta_1 HIV_{r,t} + \alpha_r + \eta_{n,t} + \varepsilon_{r,t}, \quad (4)$$

where $HIV_{r,t}$ is estimated based on HIV_r (HIV prevalence between 2001 and 2006) and a linear time path for HIV. I also estimate a country-level regression, as follows:

$$tfr_{n,t} = \beta_0 + \beta_1 HIV_{n,t} + \alpha_n + \delta_t + \varepsilon_{n,t}, \quad (5)$$

testing for a relationship between HIV and fertility using two different time paths for HIV. First, I assume a linear time path (assuming national HIV prevalence grew linearly from 1980), and next, I use estimates of prevalence from Oster (2008). Oster (2008) uses sibling histories from the DHS to construct estimates of prevalence for a subset of the countries in my sample. Her estimates provide an alternative prevalence path over the 1990s.

The results, presented in Table IV, show that my results are insensitive to the specification chosen. In particular, including data from 1991-1999 (and using different time paths to accommodate this inclusion) does not affect the estimated effect of HIV on fertility – it is still negligible. Though the country-level regressions are admittedly imprecise, they provide corroboratory evidence that the small and insignificant results in the main table (Table II) represent a true zero.

Fertility Among Older Women.

In calculating the total fertility rate in a given year, I aggregate age-specific fertility rates from the birth histories of women who were 15-39 in that year. However, if women outside of this age range (for instance, those 40-49) have children, I may be underestimating

the fertility rate. Therefore, I calculate the fraction of babies born to women who are over 40 years old. I find that less than two percent of children are born to women ages 40-49; this suggests that total fertility rates calculated using births among women ages 15-39 should be a reasonably close estimate of the number of births a woman will have over her lifetime. Even more, I find – in a difference-in-differences approach not unlike my main specification – that the change over time in the fraction of children born to women ages 40-49 is unrelated to the regional HIV rate, which implies that my estimates are unlikely to be biased by the omission of women ages 40-49.

Mortality.

In my main results, I pool data from multiple survey waves when the data overlap. For a given year, I may estimate the fertility rate using births and exposure estimates from more than one year of survey data. However, we might be concerned that, as time passes, the births and exposure measures from a survey may be a poor representation of births and exposure for that year. For instance, 10 years later, it is possible that there has been mortality or migration that has changed the composition of women living in a particular area. To the extent that this is related to HIV (which, especially in the case of mortality, it may be), we might think that this biases the total fertility rate calculations.

In Table V, column (1), I test whether the total fertility rate estimates for a given year vary depending on the year from which the data are drawn. These results suggest that different waves of the survey may yield different estimates of fertility by year – in particular, estimates of fertility from earlier waves are, on average, lower than estimates from more recent waves. If this difference is due to mortality, it is consistent with the idea that lower

fertility women (who may be more likely to be HIV-positive) experience greater mortality between waves.

Because estimates of fertility are systematically lower in earlier waves, pooling data from multiple waves of the survey may lead me to overestimate the total fertility rate. To the extent that I overestimate fertility more in areas with higher levels of HIV, my estimates of the effect of HIV on fertility may be biased. In practice, however, the data overlap between 1986 and 1999. My main analysis does not use data from 1991 through 1999, so any bias from pooling samples is likely to be small. And because HIV prevalence was low between 1986 and 1990, it is unlikely that HIV-related mortality would significantly bias estimates from this period. Nevertheless, I estimate the effect of HIV on fertility using estimates of fertility calculated from the earliest wave available, rather than from pooled data. My results, shown in columns (2)-(5) of Table V, show that this adjustment for mortality (and migration) does not affect the results – that is, the effect of HIV on fertility is small and statistically insignificant, as in Table II.

Endogeneity.

In the analysis thus far, my identifying assumption has been that we can treat HIV as exogenous with respect to fertility trends. For this to be plausible, we have to believe that the factors that made different areas susceptible to HIV are unrelated to trends in fertility over time. Differences in prevalence levels across regions are thought to be related, for instance, to geography, male circumcision rates, and the prevalence of concurrent partnerships – all of which are plausibly unrelated to trends in fertility over time. Nevertheless, we might be concerned that the difference-in-differences estimate is biased by the omission of sexual behavior controls. For instance, because pregnancy and HIV

infection are both direct results of unprotected sexual intercourse, changes in sexual behavior could drive increases in both HIV infection and fertility rates. However, in Table II I find no significant relationship between regional HIV prevalence and trends in fertility rates. This suggests that, if anything, these estimates are positively biased by the omission of sexual behavior estimates, implying that the true effect of HIV prevalence on fertility may actually be negative.

To assess whether the omission of sexual behavior biases my results, I can test whether there are differences in sexual behavior in the early 1990s across areas with different levels of HIV (in the 2000s). In particular, using data from DHS cross-sections from 1996 and earlier, I can test whether the eventual regional HIV rate (estimated using more recent data) is related to age at first intercourse in the early 1990s.[12] I find that, controlling for country, there is no relationship between regional HIV prevalence and age at first intercourse. Therefore, any bias due to the omission of sexual behavior controls is likely to be small.

Another way to evaluate the plausibility of this assumption is to test whether there are differences in fertility rate trends across areas when HIV rates were low relative to current levels. If, over the 1990s, areas with different levels of HIV were subject to different shocks than areas with lower levels of HIV, then we might expect these areas to be subject to different shocks in the 1980s (pre-"treatment"), as well. In particular, using fertility data through 1990, I test whether the change in fertility between the early 1980s and the late 1980s is related to the eventual HIV rate – it is not (Table VI, column (1)). I also interact

[12] This analysis excludes data from Ethiopia and Rwanda, for which there were no DHS surveys in the early 1990s. I also exclude Malawi, for which there is not information about age at first intercourse. In Zambia, there are two surveys (1992 and 1996) in the early 1990s. I use data from the 1996 survey. Because not all women in the sample have had intercourse, I cannot reliably calculate the average age at first intercourse. Instead, I calculate the fraction of women who have had intercourse by age 20 using a sample of women ages 20-49.

the HIV rate with a linear year term and with year indicators. The results, shown in columns (2) and (3) of Table VI, suggest that there was no difference in pre-HIV fertility rate trends across regions that ultimately had different HIV prevalence. Because trends in fertility across these areas did not differ before HIV, this lends support to my identifying assumption, and suggests that it may be reasonable to interpret the difference-in-differences estimate as the causal effect.

HIV Test Non-Response.

My analysis links information about regional HIV prevalence to data on fertility rates over time. If estimates of regional HIV prevalence are biased (for instance, because non-response is related to HIV infection), and if that bias is related to trends in fertility, then it is possible that there is an effect of HIV on fertility, but that HIV test non-response prevents me from detecting that effect. However, response rates to the HIV test are reasonably high – about 80 percent of eligible adults agreed to be tested. Furthermore, previous work by Mishra et al. (2006) looking at differences in observable characteristics between respondents and non-respondents finds that national prevalence estimates based on DHS data are not biased by non-response.

Though adjustments based on observable demographic characteristics suggest that HIV prevalence estimates are not biased, we might still be concerned that respondents and non-respondents differ in their probabilities of HIV infection. Though I cannot test this directly, one measure of HIV infection risk may be sexual behavior. Using data from Burkina Faso, Cameroon, Ghana, Kenya, and Tanzania, Fortson (2008a) finds no differences in reported sexual behavior between HIV test respondents and non-respondents, which

supports the idea that infection risk may not differ across response groups. Therefore, non-response is unlikely to bias my estimates of regional HIV prevalence.[13]

Measurement Error.

My analysis uses information from birth histories for the ten years prior to the survey to calculate fertility rates. If respondents do not accurately recall birth histories – and if their recall is poorer for earlier years – then my estimates of fertility in early years may be measured with substantially greater error. Even if measurement error is unrelated to HIV rates, it could attenuate my estimates, contributing to the estimated null effect. One way to test whether this is a concern would be to use estimates of fertility from the earliest survey wave available, rather than pooling estimates from multiple waves. Estimates of fertility from earlier waves – because less time has elapsed between the survey and reporting period – should be less susceptible to measurement error from poor recall. Therefore, the results in Table V, columns (2)-(5), in addition to testing for bias due to mortality, should provide some evidence of the effect of measurement error. Allowing for country-level trends, the difference-in-differences estimate is small and insignificant (column (5)); therefore, the evidence suggests that measurement error in fertility rates is not driving the estimated null effect.

[13] Recent work by Janssens, van der Gaag, and Rinke-deWit (2008) calls into question the conclusion that non-response bias in population-based testing samples is small. Using non-DHS data from Namibia, the authors use interviewer assignment as an instrument for HIV test response. They find substantial differences in the probability of infection between respondents and non-respondents, even after adjusting for socioeconomic characteristics, and argue that non-response bias may be larger than previously recognized. Despite this potential shortcoming, these data have many advantages for this application, particularly the ability to match HIV prevalence to fertility rate trends at the regional level.

Transformations of the Dependent Variable.

In a difference-in-differences approach – such as the one adopted here – the results may be sensitive to transformations of the dependent variable (as discussed by Meyer, 1995). In particular, if baseline fertility rates differ across areas with high and low levels of HIV, then the difference-in-differences estimate may depend on how the outcome variable is specified. Intuitively, we may be concerned that fertility may not change over time at the same rate across the fertility distribution. If the rate of change over time (in the absence of HIV) depends on the baseline level of fertility, it is possible that my approach fails to detect a difference-in-differences effect because of how it is specified, rather than because there is no effect. For instance, because fertility rates are, on average, lower in high HIV areas (e.g., Table II, column (2)), there may be less scope for declines there than elsewhere. Therefore, we might be worried that the decline in high HIV areas (though commensurate with the absolute decline in low HIV areas) signals a true decline in fertility in response to HIV. If this were the case, we might expect to see differences in the rate of change of fertility across high and low HIV areas, even before HIV was very prevalent. However, as Table VI shows, there were not significant differences in fertility rate trends across these areas during the 1980s. Furthermore, if I include in my main specifications (in Table II) an interaction between the 1990 fertility rate in a region and a post-2000 indicator – which should control for different trends across regions with different baseline levels of fertility – the difference-in-differences estimate continues to be small and insignificant. As a final check, I look at the effect of HIV on the log fertility rate. Again, the coefficient of interest is small and insignificant. These results suggest that the negligible estimated difference-in-differences effect of HIV on fertility is valid, and is not driven by the difference in baseline levels of fertility across areas with different levels of HIV.

VI. CONCLUSION

My results show evidence of a robust, null effect of the HIV/AIDS epidemic on fertility rates in sub-Saharan Africa.[14] These estimates imply that, through its effect on fertility, HIV will have a minimal impact on population size. This result calls into question the conclusions of past research forecasting improvements in wellbeing for future generations of Africans.

Furthermore, my findings suggest that the behavioral response of fertility to mortality risk is negligible. This fertility response reflects (non-)effects through two channels: the effect of short-run mortality risk for adults (through unprotected sex) and the effect of long-run mortality risk for children (i.e., effects on the desired number of children). HIV could affect fertility through both of these channels. While it would indeed be interesting to distinguish between these effects, the composite effect is nonetheless of great interest. In addition to HIV/AIDS, a leading cause of adult mortality in the developing world is maternal mortality (WHO, 2004); like HIV, the effect of maternal mortality risk on fertility may arise through short-run (mortality avoidance) and long-run (life expectancy) channels. My results show that these combined effects are negligible in the case of HIV; because maternal mortality represents a smaller risk, my results imply that reducing maternal mortality would have a relatively small, if not null, impact on fertility.

This paper provides evidence that the AIDS epidemic has had only a minimal impact, if any, on fertility behavior and the demographic transition in sub-Saharan Africa. In addition to these substantive findings, the methodology developed here – which allows me

[14] To the extent that fertility is a useful biomarker for sexual behavior, this paper likewise provides new time-series data on the response of sexual behavior to HIV risk. Because contraceptive use is widespread in sub-Saharan Africa, pregnancy may be a flawed biomarker for unprotected sex. However, it is arguably a better measure of unprotected sex than self-reported sexual behavior data, which have been shown to be biased by misreporting (e.g., de Walque, 2006). Nevertheless, my results are consistent with past work showing little response of sexual behavior to HIV/AIDS (e.g., Oster, 2007).

to construct estimates of historical fertility rates – has the potential to be applied to other research questions in other settings. This methodology is perhaps as much a contribution as the empirical findings.

REFERENCES

Becker, Gary S., and H. Gregg Lewis (1973), "On the Interaction Between the Quantity and Quality of Children," *Journal of Political Economy*, 81(2), Part 2, pp. S279-S288.

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

Bleakley, Hoyt, and Fabian Lange (2007), "Chronic Disease Burden and the Interaction of Education, Fertility, and Growth," *Review of Economics and Statistics*, forthcoming.

de Walque, Damien (2006), "Discordant Couples: HIV Infection Among Couples in Burkina Faso, Cameroon, Ghana, Kenya, and Tanzania," World Bank Policy Research Working Paper 3956.

Fortson, Jane (2008a), "The Gradient in Sub-Saharan Africa: Socioeconomic Status and HIV/AIDS," *Demography*, 45(2), pp. 303-322.

Fortson, Jane (2008b), "Morality Risk and Human Capital Investment: The Impact of HIV/AIDS in Sub-Saharan Africa," working paper, University of Chicago.

Gray, Ronald H., Maria J. Wawer, David Serwadda, Nelson Sewankambo, Chuanjun Li, Frederick Wabwire-Mangen, Lynn Paxton, Noah Kiwanuka, Godrey Kigozi, Joseph Konde-Lule, Thomas C. Quinn, Charlotte A. Gaydos, and Denise McNairn (1998), "Population-Based Study of Fertility in Women with HIV-1 Infection in Uganda," *The Lancet*, 351, pp. 98-103.

Janssens, Wendy, Jacques van der Gaag, and Tobias Rinke-deWit (2008), "Some Pitfalls in the Estimation of HIV Prevalence," Amsterdam Institute for International Development Research Series RS 08-03.

Juhn, Chinhui, Sebnem Kalemli-Ozcan, and Belgi Turan (2008), "HIV and Fertility in Africa: First Evidence form Population Based Surveys," NBER Working Paper 14248.

Lewis, James J.C., Carine Ronsmans, Alex Ezeh, and Simon Gregson (2004), "The Population Impact of HIV on Fertility in Sub-Saharan Africa," *AIDS*, 18(Supplement 2), pp. S35-S43.

Kalemli-Ozcan, Sebnem (2008), "AIDS, 'Reversal' of the Demographic Transition and Economic Development: Evidence from Africa," NBER Working Paper 12181.

Mishra, Vinod, Martin Vaessen, J. Ties Boerma, Fred Arnold, Ann Way, Bernard Barrere, Anne Cross, Rathavuth Hong, and Jasbir Sangha (2006), "HIV Testing in National Population-Based Surveys: Experiences from the Demographic and Health Surveys," *Bulletin of the World Health Organization*, 84(7), pp. 537-545.

Meyer, Bruce D. (1995), "Natural and Quasi-Experiments in Economics," *Journal of Business & Economic Statistics*, 13(2), pp. 151-161.

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

Oster, Emily (2008), "Routes of Infection: Exports and HIV Incidence in Sub-Saharan Africa," working paper, University of Chicago.

Timæus, Ian M. and Momodou Jasseh (2004), "Adult Mortality in Sub-Saharan Africa: Evidence from Demographic and Health Surveys," *Demography*, 41(4), pp. 757-772.

UNAIDS and WHO (2007), *AIDS Epidemic Update: December 2007*. (Joint United Nations Programme on HIV/AIDS and World Health Organization), Geneva, Switzerland: UNAIDS and WHO.

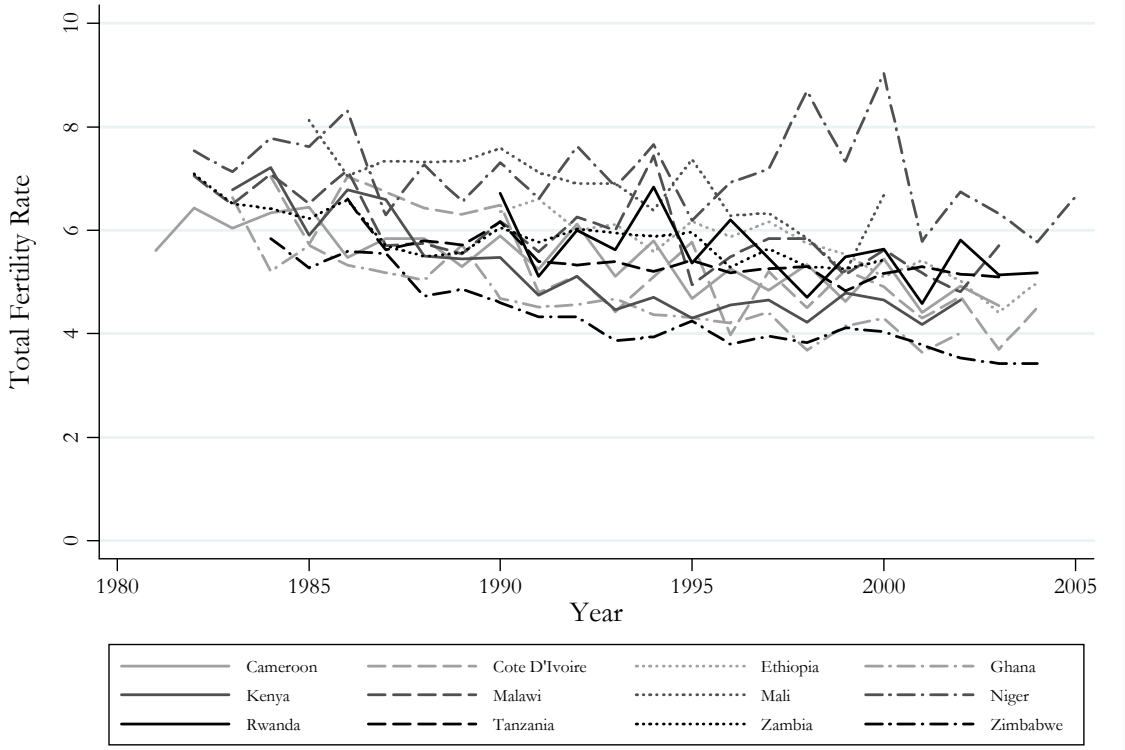
Werker, Eric, Amrita Ahuja, and Brian Wendell (2006), "Male Circumcision and AIDS: The Macroeconomic Impact of a Health Crisis," Harvard Business School Working Paper 07-025.

World Health Organization (2004), *Maternal Mortality in 2000: Estimates Developed by WHO, UNICEF, and UNFPA*. Geneva, Switzerland: WHO.

Young, Alwyn (2005), "The Gift of the Dying: The Tragedy of AIDS and the Welfare of Future African Generations," *Quarterly Journal of Economics*, 120(2), pp. 423-466.

Young, Alwyn (2007), "In Sorrow to Bring Forth Children: Fertility Amidst the Plague of HIV," *Journal of Economic Growth*, 12(4), pp. 283-327.

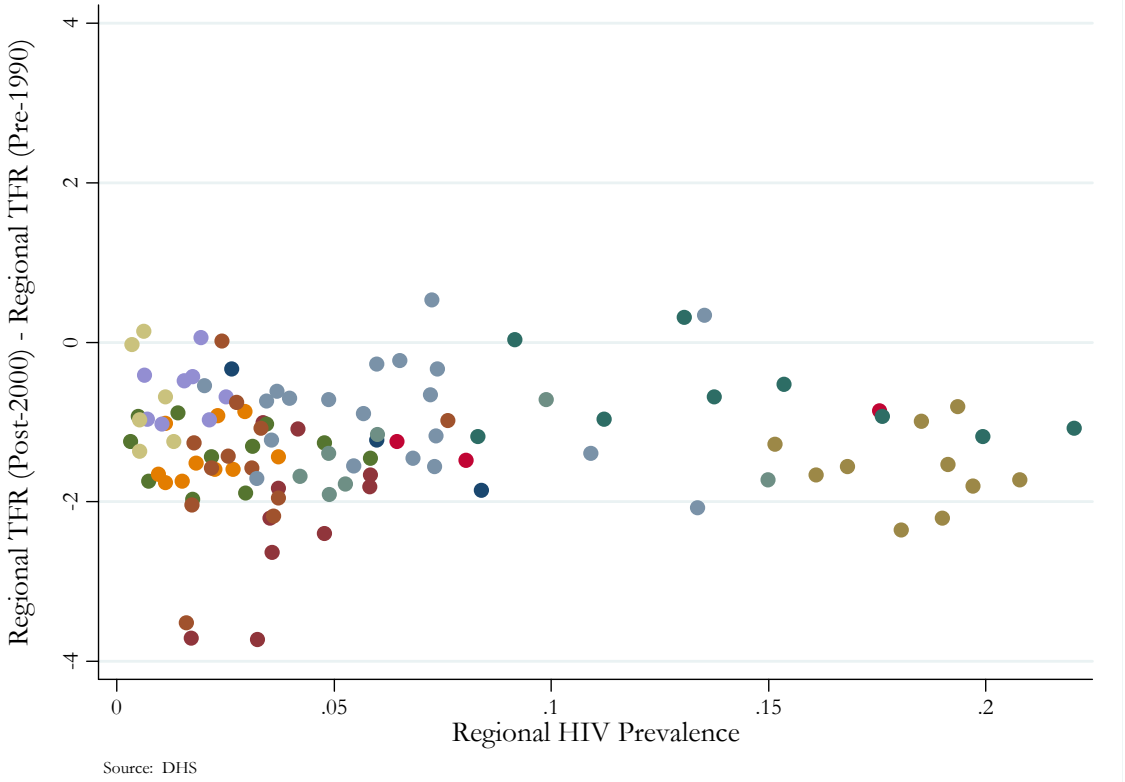
FIGURE 1
NATIONAL TOTAL FERTILITY RATES OVER TIME, BY COUNTRY



Source: DHS

Notes. Results are from the DHS for Cameroon (1991, 1998, 2004), Côte d'Ivoire (1994, 2005), Ethiopia (2000, 2005), Ghana (1993, 1998, 2003), Kenya (1993, 1998, 2003), Malawi (1992, 2000, 2004), Mali (1995/1996, 2001), Niger (1992, 1998, 2006), Rwanda (2000, 2005), Tanzania (1996, 1999, 2003, 2004), Zambia (1992, 1996, 2001/2002), and Zimbabwe (1994, 1999, 2005/2006). These total fertility rates are calculated from the DHS individual recode data using a sample which includes women ages 15-49 in the survey year, or 15-39 in the year of calculation, weighted using appropriate individual sample weights.

FIGURE 2
CHANGE IN REGIONAL TOTAL FERTILITY RATE (POST-2000 – PRE-1990),
BY REGIONAL HIV PREVALENCE (POST-2000)



Notes. Fertility data are from the DHS for Cameroon (1991, 1998, 2004), Côte d'Ivoire (1994, 2005), Ethiopia (2000, 2005), Ghana (1993, 1998, 2003), Kenya (1993, 1998, 2003), Malawi (1992, 2000, 2004), Mali (1995/1996, 2001), Niger (1992, 1998, 2006), Rwanda (2000, 2005), Tanzania (1996, 1999, 2003, 2004), Zambia (1992, 1996, 2001/2002), and Zimbabwe (1994, 1999, 2005/2006). These total fertility rates are calculated from the DHS individual recode data using a sample which includes women ages 15-49 in the survey year, or 15-39 in the year of calculation, weighted using appropriate individual sample weights. The y-axis plots the difference in the average regional total fertility rate for the years after (and including) 2000 and the years before (and including) 1990. An observation is a region (within a country). HIV is estimated (based on DHS HIV data) HIV prevalence among men and women ages 15-49 in the region in 2001 (Mali), 2001-2002 (Zambia), 2003 (Ghana, Kenya, and Tanzania), 2004 (Cameroon and Malawi), 2005 (Côte d'Ivoire, Ethiopia, and Rwanda), 2005-2006 (Zimbabwe), or 2006 (Niger).

TABLE I.A
REGIONAL HIV PREVALENCE IN SURVEY YEAR, DETAILED SUMMARY STATISTICS

Regional HIV Prevalence	CM	CI	ET	GH	KE	MW	ML
Mean	5.67	3.97	2.46	2.04	7.16	10.68	1.53
Standard Deviation	2.89	1.25	1.78	0.91	3.93	6.01	0.68
25th Percentile	2.64	3.37	0.73	1.12	4.87	6.44	0.87
Median	5.97	3.65	2.18	2.04	5.26	8.04	1.64
75th Percentile	8.39	4.77	3.44	2.67	9.89	17.56	2.03
Observations (Regions)	3	10	11	10	7	3	8

Regional HIV Prevalence	NI	RW	TZ	ZM	ZW	Total
Mean	0.74	3.03	6.48	14.49	18.26	5.36
Standard Deviation	0.38	1.61	3.14	4.70	1.75	4.93
25th Percentile	0.52	1.97	3.83	11.21	16.82	1.93
Median	0.57	2.66	6.25	13.75	18.75	3.72
75th Percentile	1.11	3.46	7.33	17.63	19.35	6.44
Observations (Regions)	6	12	20	9	10	109

Notes. Results are from the DHS for Cameroon (2004, CM), Côte d'Ivoire (2005, CI), Ethiopia (2005, ET), Ghana (2003, GH), Kenya (2003, KE), Malawi (2004, MW), Mali (2001, ML), Niger (2006, NI), Rwanda (2005, RW), Tanzania (2003, TZ), Zambia (2001/2002, ZM), and Zimbabwe (2005/2006, ZW). Table shows detailed summary statistics for the regional HIV rate in the survey year, which is estimated for each region of residence (within each country). The unit of observation is a region. These HIV rates are calculated from the DHS HIV data using a sample which includes men and women ages 15-49, weighted using appropriate HIV sample weights. In calculating summary statistics overall, country observations are weighted by population.

TABLE I.B
REGIONAL TOTAL FERTILITY RATES, SUMMARY STATISTICS

Regional TFR	CM	CI	ET	GH	KE	MW	ML
Mean	5.46	5.35	5.67	4.68	5.25	5.98	6.82
Standard Deviation	1.00	1.33	1.09	1.16	1.35	0.84	1.14
Observations	69	200	165	200	140	66	128
Regions	3	10	11	10	7	3	8
Years	23	20	15	20	20	22	16

Regional TFR	NI	RW	TZ	ZM	ZW	Total
Mean	7.14	5.57	5.43	5.91	4.32	5.54
Standard Deviation	1.19	1.07	1.23	0.90	1.07	1.31
Observations	144	180	360	171	210	2033
Regions	6	12	20	9	10	109
Years	24	15	18	19	21	--

Notes. Results are from the DHS for Cameroon (1991, 1998, 2004, CM), Côte d'Ivoire (1994, 2005, CI), Ethiopia (2000, 2005, ET), Ghana (1993, 1998, 2003, GH), Kenya (1993, 1998, 2003, KE), Malawi (1992, 2000, 2004, MW), Mali (1995/1996, 2001, ML), Niger (1992, 1998, 2006, NI), Rwanda (2000, 2005, RW), Tanzania (1996, 1999, 2003, 2004, TZ), Zambia (1992, 1996, 2001/2002, ZM), and Zimbabwe (1994, 1999, 2005/2006, ZW). Table shows summary statistics for the regional total fertility rate, which is estimated for each year for each region of residence (within each country). The unit of observation is a region/year. These total fertility rates are calculated from the DHS individual recode data using a sample which includes women ages 15-49 in the survey year, or 15-39 in the year of calculation, weighted using appropriate individual sample weights. In calculating summary statistics for a country, region observations are weighted by the sum of individual weights in a region; in calculating summary statistics for all countries combined, these weights are adjusted by population.

TABLE II
DIFFERENCE-IN-DIFFERENCES RESULTS

FERTILITY RATE	(1) NATION	(2) REGION	(3) REGION	(4) REGION	(5) REGION	(6) REGION
HIV Prevalence × Post-2000	0.560	0.021	-1.997	0.388	-0.440	0.281
	(2.010)	(1.249)	(1.643)	(1.457)	(1.551)	(2.387)
HIV Prevalence	-4.841	-5.838*	-5.471*			
	(2.603)	(1.777)	(1.905)			
Post-2000	-1.314*	-1.229*				
	(0.146)	(0.125)				
Year			-0.082*	-0.089*		
			(0.011)	(0.010)		
Constant	6.499*	6.516*	7.090*	6.816*	5.340*	5.367*
	(0.221)	(0.126)	(0.237)	(0.103)	(0.181)	(0.108)
Year FEs	No	No	No	No	Yes	No
Country × Year FEs	No	No	No	No	No	Yes
Region FEs	No	No	No	Yes	Yes	Yes
Unit of Observation	Nation × Pre/Post	Region × Pre/Post	Region × Year	Region × Year	Region × Year	Region × Year
Observations	24	218	1062	1062	1062	1062

Notes. Results are from the DHS for Cameroon (1991, 1998, 2004), Côte d'Ivoire (1994, 2005), Ethiopia (2000, 2005), Ghana (1993, 1998, 2003), Kenya (1993, 1998, 2003), Malawi (1992, 2000, 2004), Mali (1995/1996, 2001), Niger (1992, 1998, 2006), Rwanda (2000, 2005), Tanzania (1996, 1999, 2003, 2004), Zambia (1992, 1996, 2001/2002), and Zimbabwe (1994, 1999, 2005/2006). "Pre" is defined to be all years (for which we have observations) prior to and including 1990; "Post" is defined to be all years after and including 2000. Specifications using data at the region/year level likewise exclude data from years 1991-1999. All regressions are weighted least squares regressions, weighted using the population (1) or the sum of provided individual sample weights adjusted by population size (2 through 6). The dependent variable is the national total fertility rate (1) or regional total fertility rate (2-6). HIV is estimated (based on DHS HIV data) HIV prevalence among men and women ages 15-49 in the country (1) or region (2-6) in 2001 (Mali), 2001-2002 (Zambia), 2003 (Ghana, Kenya, and Tanzania), 2004 (Cameroon and Malawi), 2005 (Côte d'Ivoire, Ethiopia, and Rwanda), 2005-2006 (Zimbabwe), or 2006 (Niger). Total fertility rates are calculated using data from the individual recode on women ages 15-39, weighted using provided individual sample weights. Post-2000 indicator indicates whether the year is 2000 or later. Year is rescaled to be years since 1980. Huber-White standard errors, clustered on the nation (1) or on the region (2-6), are in parentheses. * = p-value < .05

TABLE III
CHANNEL ANALYSIS

A. SUMMARY STATISTICS

	(1) N (Countries)	(2) N	(3) Mean	(4) Std Dev.
Fertility Rate (Full Sample)	13	1150	5.318	1.466
Fertility Rate (HIV-Negative Sample)	13	1150	5.460	1.535
Difference (HIV-Negative – Full)	13	1150	0.146	0.500

B. REGRESSION ANALYSIS

Fertility Rate (HIV-Negative Sample)	(1) REGION	(2) REGION	(3) REGION	(4) REGION
HIV Prevalence × Post	-2.092	1.071	0.279	-0.248
	(2.196)	(2.044)	(2.199)	(2.419)
HIV Prevalence	-5.228			
	(2.651)			
Year	-0.067*	-0.085*		
	(0.013)	(0.012)		
Constant	6.900*	6.791*	5.769*	5.590*
	(0.272)	(0.126)	(0.092)	(0.071)
Year FEs	No	No	Yes	No
Country × Year FEs	No	No	No	Yes
Region FEs	No	Yes	Yes	Yes
Unit of Observation	Region × Year	Region × Year	Region × Year	Region × Year
Observations	736	736	736	736

Notes. The top panel uses data from Burkina Faso (2003), Cameroon (2004), Côte d'Ivoire (2005), Ethiopia (2005), Ghana (2003), Guinea (2005), Kenya (2003), Lesotho (2004), Malawi (2004), Niger (2006), Rwanda (2005), Senegal (2005), and Zimbabwe (2005/2006). The bottom panel uses data from Cameroon (1991, 1998, 2004), Côte d'Ivoire (1994, 2005), Ethiopia (2000, 2005), Ghana (1993, 1998, 2003), Kenya (1993, 1998, 2003), Malawi (1992, 2000, 2004), Niger (1992, 1998, 2006), Rwanda (2000, 2005), and Zimbabwe (1994, 1999, 2005/2006). "Pre" is defined to be all years (for which we have observations) prior to and including 1990; "Post" is defined to be all years after and including 2000. Specifications using data at the region/year level likewise exclude data from years 1991-1999. All regressions are weighted least squares regressions, weighted using the sum of provided individual sample weights adjusted by population size. The dependent variable is the regional total fertility rate. HIV is estimated (based on DHS HIV data) HIV prevalence among men and women ages 15-49 in the region in 2003 (Ghana and Kenya), 2004 (Cameroon and Malawi), 2005 (Côte d'Ivoire, Ethiopia, and Rwanda), 2005-2006 (Zimbabwe), or 2006 (Niger). Total fertility rates are calculated using data from the individual recode on women ages 15-39, weighted using provided individual sample weights. For the ten years prior to the most recent cross-section, this is calculated using the sample of HIV-negative women. In earlier years, the fertility rate is calculated using the full sample of women. Post-indicator indicates whether the year is 2000 or later. Year is rescaled to be years since 1980. Huber-White standard errors, clustered on the region, are in parentheses. * = p-value < .05

TABLE IV
ALTERNATE TIME PATHS

FERTILITY RATE	(1) REGION	(2) REGION	(3) NATION	(4) NATION
HIV Prevalence (Survey Year) × Post-2000	0.579			
	(1.453)			
HIV Prevalence (Contemporaneous, Linear)		-2.075	2.161	
		(3.534)	(5.535)	
HIV Prevalence (Contemporaneous, Oster)				1.735
				(3.306)
Constant	5.166*	5.275*	5.071*	5.355*
	(0.074)	(0.158)	(0.161)	(0.247)
Year FEs	No	No	Yes	Yes
Country FEs	No	No	Yes	Yes
Country × Year FEs	Yes	Yes	No	No
Region FEs	Yes	Yes	No	No
Unit of Observation	Region × Year	Region × Year	Nation × Year	Nation × Year
Observations	2033	2033	233	46

Notes. Results are from the DHS for Cameroon (1991, 1998, 2004), Côte d'Ivoire (1994, 2005), Ethiopia (2000, 2005), Ghana (1993, 1998, 2003), Kenya (1993, 1998, 2003), Malawi (1992, 2000, 2004), Mali (1995/1996, 2001), Niger (1992, 1998, 2006), Rwanda (2000, 2005), Tanzania (1996, 1999, 2003, 2004), Zambia (1992, 1996, 2001/2002), and Zimbabwe (1994, 1999, 2005/2006). Results in column (4) use data from Cameroon, Kenya, Malawi, Mali, Zambia, and Zimbabwe. Columns (1)-(3) use fertility rate data from all available years (i.e., including 1991-1999). Column (4) uses fertility rate data from alternating years (those that match with Oster's HIV data), generally from the mid-1980s to the mid-1990s. All regressions are weighted least squares regressions, weighted using the sum of provided individual sample weights adjusted by population size (1 and 2) or the population (3 and 4). The dependent variable is the regional total fertility rate (1 and 2) or the national total fertility rate (3 and 4). HIV Prevalence (Survey Year) is estimated (based on DHS HIV data) HIV prevalence among men and women ages 15-49 in the region in 2001 (Mali), 2001-2002 (Zambia), 2003 (Ghana, Kenya, and Tanzania), 2004 (Cameroon and Malawi), 2005 (Côte d'Ivoire, Ethiopia, and Rwanda), 2005-2006 (Zimbabwe), or 2006 (Niger). HIV Prevalence (Contemporaneous, Linear) is HIV prevalence in the same year for which the total fertility rate is estimated, calculated assuming that HIV grew linearly from zero percent in 1980 to the survey year estimates. HIV Prevalence (Contemporaneous, Oster) uses HIV prevalence data (matched to the year for which the total fertility rate is estimated) from Oster (2008). Total fertility rates are calculated using data from the individual recode on women ages 15-39, weighted using provided individual sample weights. Post-2000 indicator indicates whether the year is 2000 or later. Huber-White standard errors, clustered on the region (1-2) or on the nation (3-4), are in parentheses. * = p-value < .05

TABLE V
SAMPLE MORTALITY

FERTILITY RATE	(1) WAVES	(2) REGION	(3) REGION	(4) REGION	(5) REGION
HIV Prevalence × Post-2000		1.783	4.068*	1.372	0.194
		(1.015)	(0.922)	(1.071)	(1.972)
HIV Prevalence		-7.602*			
		(1.730)			
Year		-0.099*	-0.105*		
		(0.009)	(0.008)		
Earlier Wave	-0.655*				
	(0.120)				
Constant	5.813*	7.254*	6.936*	5.801*	5.270*
	(0.197)	(0.172)	(0.113)	(0.224)	(0.090)
Year FEs	No	No	No	Yes	No
Country × Year FEs	No	No	No	No	Yes
Region FEs	No	No	Yes	Yes	Yes
Unit of Observation	Region × Year × Wave	Region × Year	Region × Year	Region × Year	Region × Year
Observations	1556	2033	2033	2033	2033

Notes. Results are from the DHS for Cameroon (1991, 1998, 2004), Côte d'Ivoire (1994, 2005), Ethiopia (2000, 2005), Ghana (1993, 1998, 2003), Kenya (1993, 1998, 2003), Malawi (1992, 2000, 2004), Mali (1995/1996, 2001), Niger (1992, 1998, 2006), Rwanda (2000, 2005), Tanzania (1996, 1999, 2003, 2004), Zambia (1992, 1996, 2001/2002), and Zimbabwe (1994, 1999, 2005/2006). Earlier Wave indicates whether the fertility estimate comes from an earlier (of two) wave; in column (1), the sample is restricted to region/years for which there are two waves of data. Columns (2)-(5) use estimates of fertility from the earliest wave available, rather than pooling data from multiple waves. "Pre" is defined to be all years (for which we have observations) prior to and including 1999; "Post" is defined to be all years after and including 2000. All specifications use data from years 1991-1999. All regressions are weighted least squares regressions, weighted using the sum of provided individual sample weights adjusted by population size. The dependent variable is the regional total fertility rate. HIV is estimated (based on DHS HIV data) HIV prevalence among men and women ages 15-49 in the region in 2001 (Mali), 2001-2002 (Zambia), 2003 (Ghana, Kenya, and Tanzania), 2004 (Cameroon and Malawi), 2005 (Côte d'Ivoire, Ethiopia, and Rwanda), 2005-2006 (Zimbabwe), or 2006 (Niger). Total fertility rates are calculated using data from the individual recode on women ages 15-39, weighted using provided individual sample weights. Year is rescaled to be years since 1980. Huber-White standard errors, clustered on the region, are in parentheses. * = p-value < .05

TABLE VI
FALSIFICATION TESTS

FERTILITY RATE	(1) REGION	(2) REGION	(3) REGION
HIV Prevalence (Survey Year) × Post-1985	-1.913		
	(2.534)		
HIV Prevalence (Survey Year) × Year		-0.521	
		(0.284)	
HIV Prevalence (Survey Year) × (Year = 1982)			8.704
			(4.571)
HIV Prevalence (Survey Year) × (Year = 1983)			2.367
			(2.716)
HIV Prevalence (Survey Year) × (Year = 1984)			1.465
			(3.923)
HIV Prevalence (Survey Year) × (Year = 1985)			5.848*
			(2.913)
HIV Prevalence (Survey Year) × (Year = 1986)			1.354
			(3.471)
HIV Prevalence (Survey Year) × (Year = 1987)			-3.076
			(2.924)
HIV Prevalence (Survey Year) × (Year = 1988)			1.729
			(2.818)
HIV Prevalence (Survey Year) × (Year = 1989)			1.490
			(2.602)
HIV Prevalence (Survey Year) × (Year = 1990)			0.484
			(2.561)
Constant	6.106*	6.309*	5.890*
	(0.137)	(0.193)	(0.182)
Year FEs	No	No	No
Country FEs	No	No	No
Country × Year FEs	Yes	Yes	Yes
Region FEs	Yes	Yes	Yes
Unit of Observation	Region × Year	Region × Year	Region × Year
Observations	639	639	639

Notes. Results are from the DHS for Cameroon (1991, 1998), Côte d'Ivoire (1994), Ethiopia (2000), Ghana (1993, 1998), Kenya (1993, 1998), Malawi (1992, 2000), Mali (1995/1996), Niger (1992, 1998), Rwanda (2000), Tanzania (1996, 1999), Zambia (1992, 1996), and Zimbabwe (1994, 1999). The sample includes data from years before (and including) 1990. All regressions are weighted least squares regressions, weighted using the sum of provided individual sample weights adjusted by population size. The dependent variable is the region total fertility rate. HIV Prevalence (Survey Year) is estimated (based on DHS HIV data) HIV prevalence among men and women ages 15-49 in the region in 2001 (Mali), 2001-2002 (Zambia), 2003 (Ghana, Kenya, and Tanzania), 2004 (Cameroon and Malawi), 2005 (Côte d'Ivoire, Ethiopia, and Rwanda), 2005-2006 (Zimbabwe), or 2006 (Niger). Total fertility rates are calculated using data from the individual recode on women ages 15-39, weighted using provided individual sample weights. Column (1) interacts HIV prevalence with a post-1985 indicator; column (2) interaction HIV prevalence with a linear year term (rescaled to be years since 1980); and column (3) interacts HIV prevalence with year indicators – 1981 is the omitted year. Huber-White standard errors, clustered on the region, are in parentheses. * = p-value < .05

DATA APPENDIX

Data for the analysis come from Demographic and Health Surveys (DHS), which are available from ORC Macro (http://www.measuredhs.com). The primary sample includes twelve countries: Cameroon, Côte d'Ivoire, Ethiopia, Ghana, Kenya, Malawi, Mali, Niger, Rwanda, Tanzania, Zambia, and Zimbabwe.[15] The calculation of HIV rates uses the following cross-sections: Cameroon (2004), Côte d'Ivoire (2005), Ethiopia (2005), Ghana (2003), Kenya (2003), Malawi (2004), Mali (2001), Niger (2006), Rwanda (2005), Tanzania (2004), Zambia (2001/2002), and Zimbabwe (2005/2006).[16] Fertility information come from these and other cross-sections, and are matched on the region to HIV prevalence information.[17] In particular, I use the following regional groupings and the following cross-sections:

Cameroon (1991, 1998, 2004): north/extreme north/adamaoua, central/south/east/west/ littoral, northwest/southwest

Côte d'Ivoire (1994, 2005): center, center-east, center-north, center-west, north, northeast, northwest, west, south (includes abidjan), southwest

Ethiopia (2000, 2005): tigray, afar, amhara, oromiya, somali, ben-gumz, snnp, gambela, harari, addis abeba, dire dawa

Ghana (1993, 1998, 2003): western, central, greater accra, volta, eastern, ashanti, brong ahafo, northern, upper west, upper east

Kenya (1993, 1998, 2003): nairobi, central, coast, eastern, nyanza, rift valley, western

Malawi (1992, 2000, 2004): northern, central, southern

[15] In some robustness checks, we use data from four additional countries: Burkina Faso (2003), Guinea (2005), Lesotho (2004), and Senegal (2005). Regions in Burkina Faso and Guinea are not sufficiently consistent over time to permit linking multiple cross-sections; the 1999 data for Senegal are also difficult to compare to the most recent cross-section. Lesotho is excluded from the main analysis because there is only one cross-section from which to draw data.

[16] The 2003 DHS for Tanzania is also referred to as the HIV/AIDS Indicator Survey (AIS), and covers only mainland Tanzania. The 2005 DHS for Côte d'Ivoire is also referred to as the HIV/AIDS Indicator Survey (AIS). The data used in this analysis from those surveys as well as from Ethiopia (2005), Niger (2006), Senegal (2005) and Zimbabwe (2005/2006) are from preliminary releases of the data. The Ethiopia survey lists 1997 as the survey year because the Ethiopian calendar differs from the Gregorian calendar; however, when calculating year of birth, this analysis refers to the Gregorian calendar. To convert, 92 months are added to the Ethiopian calendar dates.

[17] The Tanzania (2003) data, from which Tanzania's HIV data are drawn, do not include fertility data. While the 1998/1999 DHS survey for Côte d'Ivoire has fertility information, it lacks region information, making it difficult to compare to data from other survey years; therefore, this analysis uses only 1994 and 2005 data from Côte d'Ivoire.

Mali (1995/1996, 2001): kayes, koulikoro, sikasso, segou, mopti, tombouctou, gao, bamako

Niger (1992, 1998, 2006): niamey, dosso, maradi, tahoua/agadez, tillaberi, zinda/diffa

Rwanda (2000, 2005): ville de kigali, kigali ngali, gitarama, butare, gikongoro, cyangugu, kibuye, gisenyi, ruhengeri, byumba, umutara, kibungo

Tanzania (1996, 1999, 2003, 2004): dodoma, arusha/manyara, kilimanjaro, tanga, morogoro, coast, dar es salam, lindi, mtwara, ruvuma, iringa, mbeya, singida, tabora, rukwa, kigoma, shinyanga, kagera, mwanza, mara

Zambia (1992, 1996, 2001/2002): central, copperbelt, eastern, luapula, lusaka, northern, northwestern, southern, western

Zimbabwe (1994, 1999, 2005/2006): manicaland, mashonaland central, mashonaland east, mashonaland west, matabeleland north, matabeleland south, midlands, masvingo, harare, bulawayo

The standard DHS has several survey components, including a household questionnaire, women's questionnaire, and men's questionnaire. HIV testing was also conducted in a recent wave in the twelve countries in my sample. In Cameroon, Côte d'Ivoire, Ethiopia, Ghana, Kenya, Malawi, Niger, Rwanda, Tanzania, and Zimbabwe, these HIV test results can be linked to other respondent characteristics. In Mali and Zambia, HIV test results are unlinked to the survey data; however, test results can be used to calculate HIV prevalence for various subgroups, including regional prevalence. Regional HIV prevalence is estimated at the region level, and is calculated as HIV prevalence among adults ages 15-49 (for comparability across countries).

In this analysis, I link regional HIV prevalence and (in all but Mali, Tanzania, and Zambia) individual HIV status to trends in fertility over time. Information from the household questionnaire is used to calculate prevalence and link results (for instance, to determine region) but otherwise all information about respondents is drawn from the individual (women's) questionnaire.[18,19]

[18] The individual file for Côte d'Ivoire (2005) includes both men and women. I exclude male respondents from the analysis.

[19] Fertility rates are calculated for the following years: Cameroon (1981-2003), Côte d'Ivoire (1984-1993, 1995-2004), Ethiopia (1990-2004), Ghana (1983-2002), Kenya (1983-2002), Malawi (1982-2003), Mali (1985-2000), Niger (1982-2005), Rwanda (1990-2004), Tanzania (1986-2003), Zambia (1982-2000), and Zimbabwe (1984-2004).

In Malawi, Mali, and Zambia, HIV tests were conducted in a one-in-three subsample of households. In Cameroon, Ethiopia, Kenya, Niger, and Rwanda, HIV tests were conducted in a one-in-two subsample of households. In Côte d'Ivoire, Ghana, Tanzania, and Zimbabwe, HIV tests were conducted in all households.

Each respondent eligible for HIV testing was asked to provide a blood sample for testing. In Cameroon, Côte d'Ivoire, Ethiopia, Ghana, Kenya, Malawi, Mali, Niger, Rwanda, Tanzania, and Zimbabwe, HIV testing was conducted on dried blood spot specimens collected by finger prick. In Zambia, the dried blood spot specimen came from a venous blood specimen.

Survey and HIV test non-response rates for the most recent wave of the survey are shown in Appendix Table A.I. Response rates for the individual questionnaire are quite high. Response rates for the HIV test components are somewhat lower.

This analysis weights each region by a population-adjusted regional weight. The regional weight is the sum of the individual weights of women in that region. The individual weights adjust for the sampling probability and response rate (Rutstein and Rojas, 2003). I adjust these regional weights by population estimates from the 2007 CIA World Factbook so that the sum of the regional weights in a given country equals that country's population. (In country-level regressions, observations are weighted by population alone.) Regional HIV prevalence is calculated using DHS-provided HIV weights, which adjust for individual sampling probabilities and test non-response (separately by sex).

TABLE A.I
SURVEY RESPONSE RATES

Country	Survey Year	Individual Questionnaire	HIV Test		
			Men	Women	Total
Burkina Faso	2003	96.3	85.8	92.3	89.3
Cameroon	2004	94.3	89.8	92.1	91.0
Cameroon	1998	95.5	-	-	-
Cameroon	1991	94.3	-	-	-
Côte d'Ivoire	2005	89.8	76.3	79.1	77.8
Côte d'Ivoire	1994	97.9	-	-	-
Ethiopia	2005	95.6	75.6	83.2	79.4
Ethiopia	2000	97.8	-	-	-
Ghana	2003	95.7	80.0	89.3	84.9
Ghana	1998	97.4	-	-	-
Ghana	1993	97.1	-	-	-
Guinea	2005	97.2	88.2	92.5	90.6
Kenya	2003	94	70.3	76.3	73.4
Kenya	1998	95.7	-	-	-
Kenya	1993	94.8	-	-	-
Lesotho	2004	94.3	68.0	80.7	74.7
Malawi	2004	95.7	63.3	70.4	67.0
Malawi	2000	97.7	-	-	-
Malawi	1992	96.6	-	-	-
Mali	2001	94.9	75.6	85.2	80.7
Mali	1995/1996	96.1	-	-	-
Niger	2006	95.6	84.2	90.7	87.8
Niger	1998	96.4	-	-	-
Niger	1992	96.3	-	-	-
Rwanda	2005	98.1	95.6	97.3	96.5
Rwanda	2000	98.1	-	-	-
Senegal	2005	93.7	75.5	84.5	80.4
Tanzania	2004	97.3	-	-	-
Tanzania	2003	-	77.0	83.5	80.5
Tanzania	1999	97.8	-	-	-
Tanzania	1996	95.5	-	-	-
Zambia	2001/2002	96.4	73.3	79.4	76.5
Zambia	1996	96.7	-	-	-
Zambia	1992	97.4	-	-	-
Zimbabwe	2005/2006	90.2	63.4	75.9	70.0
Zimbabwe	1999	95.2	-	-	-
Zimbabwe	1994	95.6	-	-	-

Notes. Percent surveyed or tested among eligible respondents. In all countries, the individual response rate refers to the rate among women.