Disease Risk and Fertility: Evidence from the HIV/AIDS Pandemic

Yoo-Mi Chin and Nicholas Wilson[[1]](#footnote-1)\*

First draft: October 1, 2015

Current version: December 29, 2016

*Abstract:* A fundamental question about human behavior is whether fertility responds to disease risk. The standard economic theory of household fertility decision-making generates ambiguous predictions and the response has large implications for human welfare. We examine the fertility response to the HIV/AIDS pandemic using national household survey data from fourteen sub-Saharan African countries. Instrumental variable (IV) estimates using distance to the origin of the pandemic suggest that HIV/AIDS has increased the total fertility rate (TFR) and the number of surviving children. These results rekindle the debate about the fertility response to disease risk and highlight the question of whether the HIV/AIDS pandemic has reduced GDP per capita.

*Keywords:* disease, fertility, HIV/AIDS, instrumental variable regression

*JEL codes:* I15, J13, O12

**1. Introduction**

A fundamental question about human behavior is whether fertility responds to disease risk. The standard economic theory of household fertility decision-making (Becker and Lewis 1973) suggests that disease risk may affect demand for children through changes in household resources and the shadow prices of child quantity and quality, yet this theory does not yield clear predictions about the net effect of disease risk. Not only does the main economic theory of fertility yield an ambiguous prediction about the effect of disease risk on fertility, but also the nature of the response may have large implications for human welfare. For example, reductions in child mortality risk that are not associated with fertility reductions are likely to reduce GDP per capita by increasing population growth (Malthus 1798, Acemoglu and Johnson 2007). We examine the fertility response to the leading cause of adult mortality (WHO 2011) and one of the leading causes of child mortality (WHO 2011) in sub-Saharan Africa: HIV/AIDS.

A major barrier to answering this question is the endogeneity of disease risk to the process determining fertility. Existing economic literature on the fertility response to HIV/AIDS has addressed this issue using geographic and time fixed effects (Durevall and Lindskog 2011, Kalemli-Ozcan 2012, Juhn et al 2013), difference-in-differences approaches exploiting heterogeneity in the rise of the pandemic (Fortson 2009, Fink and Linnemayr 2009, Kalemli-Ozcan 2012), a synthetic control group method for comparative case studies (Karlsson and Pichler 2015), and system Generalized Method of Moments (GMM) (Boucekkine et al 2009). These studies present mixed evidence on the fertility response to HIV/AIDS, with some finding a negative association (Young 2005, Young 2007, Boucekkine et al 2009), others finding no clear association on average (Fortson 2009, Kalemli-Ozcan and Turan 2011, Kalemli-Ozcan 2012, Juhn et al 2013, Karlsson and Pichler 2015), and at least three studies finding heterogeneous responses (Fink and Linnemayr 2009, Durevall and Lindskog 2011, Karlsson and Pichler 2015).

Our study applies an established instrument for HIV prevalence in health behavior regressions to provide further evidence on this question. As demonstrated in Oster (2012) and Chin (2013) and confirmed in our analysis, distance to the origin of the HIV/AIDS pandemic is a strong predictor of the early spread of HIV. Distance to the origin of the HIV/AIDS pandemic, particularly conditional on latitude and longitude, is also plausibly exogenous to the processes determining health behaviors and health outcomes.[[2]](#footnote-2) We appear to be the first to use this instrument to analyze the effect of HIV on fertility.[[3]](#footnote-3)

We implement our instrumental variable (IV) regression analysis using household survey data from a sample of 14 countries in sub-Saharan Africa. These countries are those that have the first HIV testing modules in standardized national household surveys (i.e. the Demographic and Health Surveys (DHS)), information on lifetime individual birth histories, and cluster GPS coordinates. As in previous studies using distance to the origin of the HIV/AIDS pandemic as an instrument, we limit our main analysis to the set of surveys conducted in the early-to-mid-2000s, a period over which pandemic countries were not in a steady-state and hence distance to the origin was strongly associated with HIV prevalence (Oster 2012, Chin 2013), and omit two countries with large-scale civil conflicts during this period (Oster 2012).[[4]](#footnote-4) We use these data to construct cluster-level measures of HIV prevalence and distance to the origin of the HIV/AIDS pandemic. In addition, we follow the method proposed and implemented in Fortson (2009) and use these data to construct cluster-level measures of total fertility rates (TFR), as well as complementing this measure with a cluster-level measure of the total number of surviving children. In total, we examine the fertility response to HIV risk using data from a sample of approximately 130,000 women in nearly 5,500 distinct geographic locations.

We find that fertility has increased in response to HIV/AIDS risk in our sample of countries. Our un-instrumented results suggest that there exists a negative correlation between cluster-level HIV prevalence and fertility, consistent with omitted factors such as respondent-level schooling being simultaneously associated with low fertility and high HIV prevalence (Fortson 2008). In contrast, the IV regression results suggest cluster-level HIV prevalence has increased TFR and has increased the number of surviving children. In the full sample of countries, we find that a doubling of HIV prevalence increased TFR by approximately two births and increased the number of surviving children by approximately 0.5 children. A variety of robustness and falsification checks support our main results. An analysis of the heterogeneous effects by age suggest that older women of child bearing age have responded more than women of younger child bearing age.

By using an empirical strategy that heretofore has not been applied to this research question, our analysis provides new causal evidence on an unresolved question in the existing economic literature. Although our results do not resolve the mixed findings in the existing literature (Young 2005, Young 2007, Boucekkine et al 2009, Fink and Linnemayr 2009, Fortson 2009, Durevall and Lindskog 2011, Kalemli-Ozcan and Turan 2011, Kalemli-Ozcan 2012, Juhn et al 2013, Karlsson and Pichler 2015), they do highlight the fact that the evidence, including ours, on the fertility response to the HIV/AIDS pandemic is not definitive.

In addition, we expand the very small body of economic literature on the effect of HIV/AIDS on the number of surviving children. Among the existing economic analyses of the fertility response to HIV/AIDS, only Young and (2005) and Boucekkine et al (2009) provide evidence on the effect on the number of surviving children. Evidence on whether HIV/AIDS increases the number of surviving children and not just births is a major factor in determining the effect of the pandemic on gross domestic product (GDP) per capita or other measures of material standard of living. Young (2005) suggests that that HIV/AIDS has increased GDP per capita in South Africa. Our finding using data from 14 pandemic countries in sub-Saharan Africa that HIV/AIDS has increased the number of surviving children suggests that HIV/AIDS may have reduced GDP per capita by increasing dependency ratios.

Our study also contributes to a broader economic literature on the fertility response to other diseases (or their local eradication) such as hookworm (Bleakley and Lange 2009), malaria (Aksan and Chakborty 2013, Lucas 2013, McCord et al 2017), and diarrhea (Aksan and Chakborty 2013), or disease-risk in general (Bhalotra and Soest 2008, Angeles 2010). Our finding that HIV/AIDS risk has increased fertility is consistent with the fertility responses in Bhalotra and Soest (2008), Bleakley and Lange (2009), Angeles (2010), Aksan and Chakraborty (2013), and McCord et al (2017) for other diseases, suggesting that fertility typically increases in response to disease risk. Although our finding on the fertility response to disease risk differs from that presented in Lucas (2013) for malaria eradication, malaria may provide a unique biomedical mechanism in that malaria has a larger effect on the survival probability of the first birth compared to later pregnancies (Lucas 2013). Becoming HIV positive is an absorptive state until death, with the cumulative probability of being HIV positive strictly increasing in age for any given woman, meaning that HIV/AIDS may have a larger effect on the survival of later births.[[5]](#footnote-5)

The rest of the paper is organized as follows. Section 2 describes economic mechanisms linking HIV/AIDS to fertility as well as existing evidence on the fertility response to HIV/AIDS in sub-Saharan Africa. Section 3 describes our data sources and variables construction. Section 4 presents our empirical strategy for estimating the causal effect of HIV/AIDS on fertility. Section 5 presents our results and we discuss these in more detail in Section 6. Section 7 provides concluding remarks.

**2. Fertility in the HIV/AIDS Pandemic**

**2.1 Conceptual framework**

The standard economic theory of household fertility decision-making (Becker and Lewis 1973) predicts that disease risk will affect demand for children through several mechanisms. Adult health risk reduces time horizons for decision-making and, via reduced labor productivity and increased demand for care labor in the home, reduces household income. Child health risk increases the shadow prices of child quantity and quality. Because HIV/AIDS simultaneously affects adult and child health risk, the sign of the net effect through all of these channels is not clear.

 Economic research on the fertility response to HIV/AIDS has extended the Becker model in several key ways. Young (2005) embeds the Becker model in a Solow growth model and finds that HIV may reduce fertility by raising labor scarcity and the value of women’s time. Boucekkine et al (2009) embeds the Becker model in an overlapping generations (OLG) model where fertility and labor supply respond to adult and child mortality risks, finding that adult mortality risk has an ambiguous effect on fertility and child mortality increases total fertility. Fink and Linnemayr (2009) distinguishes between mortality risks of infant and adult children and between incentives faced by more educated and less educated parents, generating the prediction that more educated women reduce fertility more than less educated women. The empirical analysis in Juhn et al (2012) distinguishes between the effects of own disease status and community-level disease risk to allow for direct physiological effects of HIV/AIDS on fecundity (Gray et al 1998).

**2.2 Preliminary and existing evidence**

 Figure 1 presents (unweighted) time series evidence on the association between the rise of the HIV/AIDS pandemic and fertility in sub-Saharan Africa using country-level data from the United Nations. Our HIV time series begins in 1990, the start of comprehensive national statistics on HIV prevalence in this region of the world. We include TFR going back to 1960, a period when HIV prevalence was virtually zero in each of these countries. The figure reveals a striking similarity in the timing of the fertility decline in this sample and the onset of the HIV/AIDS pandemic. Although the exact trajectory of HIV prevalence prior to 1990 is not available, HIV prevalence increased from zero during the decades leading up to 1990 to around 4% by 1990. In other words, the fall in fertility coincided with the rise of the HIV/AIDS pandemic. However, concurrent expansions in much of this region in schooling (Ainsworth et al 1996, Lloyd et al 2000, Osili and Long 2008, Behrman 2015) and access to modern contraception technologies (Caldwell et al 1992, Ainsworth et al 1996, Lam 2011) make it impossible to assign causality to HIV/AIDS in explaining fertility decline this figure.

 Cross-sectional and panel economic analyses have yielded mixed evidence on the association between HIV prevalence and fertility in sub-Saharan African settings. Young (2005) and Young (2007) use data from South Africa and find evidence suggesting that fertility has fallen in response to HIV/AIDS. Boucekkine et al (2009) and Kalemli-Ozcan (2012) each use data from multiple sub-Saharan African countries and find evidence similarly suggesting that fertility has fallen in response to HIV/AIDS. In contrast, Fortson (2009), Kalemli-Ozcan (2012), and Juhn et al (2013) use data from multiple sub-Saharan African countries and find evidence suggesting that fertility has not responded to HIV/AIDS.[[6]](#footnote-6) Kalemli-Ozcan and Turan (2011) revisit the analysis of data from South Africa in Young (2005) and find evidence that fertility has not responded to HIV/AIDS. Fink and Linnemayr (2009) uses data from Cameroon, Cote d’Ivoire, Ghana, Kenya, Mali, and Senegal to show that heterogeneous responses by education may underlie an estimated zero average response. Using data from Malawi, Durevall and Lindskog (2011) shows that heterogeneous responses by age and parity may underlie an estimated zero average response. In a synthetic control method comparative case study of Mozambique, South Africa and Zimbabwe, Karlsson and Pichler (2015) find some evidence of heterogeneous responses by age and that fertility on average has not responded to HIV/AIDS.

**3. Data**

Data for our analysis come from cross-sectional geo-referenced national household surveys (i.e. the Demographic and Health Surveys (DHS)). Our empirical strategy relies on individual birth history files, anonymous HIV testing files, and the cluster GPS coordinates. We use data from the 14 sub-Saharan African countries included in the first round of DHS HIV testing and for which GPS data are available.[[7]](#footnote-7)

**3.1 HIV prevalence**

Existing economic research on the fertility response to HIV almost exclusively relies on HIV data from the DHS HIV testing modules. We also use these data, which provide anonymous blood sample results from a sub-sample of individuals who are asked and who consent to blood tests as part of the DHS survey. Response rates in the HIV testing modules are approximately 80 percent. In the 14 countries in our sample, the DHS links the results of these tests to the individual’s survey cluster of residence. We use these individual-level data to calculate cluster-level HIV prevalence. Following previous economic literature (e.g. Fortson 2009, Oster 2012, and Chin 2013), we calculate HIV prevalence using adult females and males.

Table 1 presents descriptive statistics for HIV prevalence in our sample. In total, we calculate HIV prevalence for 5,483 DHS clusters using data from 124,007 adult respondents. Mean cluster-level HIV prevalence ranges from a low of 0.9% in Senegal to a high of 23.75% in Lesotho.

**3.2 Total fertility rates**

We use the birth history files to calculate total fertility rates (TFR) at the cluster level. Following Fortson (2009), we calculate TFR as:

$$TFR\_{c,t}=5 \sum\_{i=3}^{9}\frac{babies\_{c,t,5i,5i+4}}{exposure\_{c,t,5i,5i+4}}\_{} \left(1\right)$$

where $TFR\_{c,t}$ is an aggregation of the age-specific fertility rates (provided by the righthand-side of this equation), $babies\_{c,t,5i,5i+4}$ measures the number of children born in year *t* to women in the age group 5i to 5i+4, and $exposure\_{c,t,5i,5i+4}$ measures the total years represented by women in cluster *c* during year *t* in the age group 5i to 5i+4. One key difference in our study from Fortson (2009) is that we use average fertility across the three years preceding the survey year for women age 15-49.

Table 1 presents descriptive statistics for TFR in our sample. In total, we calculate TFR for 5,483 DHS clusters using data from 135,435 adult female respondents. Mean cluster-level TFR ranges from a low of 3.49 births in Lesotho to a high of 6.53 births in Malawi.

**3.3 Distance to origin**

For our sample countries, the DHS includes anonymized cluster GPS coordinates of the centroids of the survey clusters. Following the existing literature (Oster 2012, Chin 2013), we assume that the center of the Democratic Republic of the Congo (i.e. latitude=-6.31, longitude=23.59) is the origin of the HIV/AIDS pandemic.

We calculate the great-circle distance between each cluster centroid and the origin of the HIV/AIDS pandemic. Great-circle distance is the shortest distance between two points on a sphere, measure along the surface of the sphere. An analysis of the approximation error associated with the sphericality assumption suggests that it is less than 0.5% latitude and 0.2% longitude (Ministry of Defence (Navy) 1987). A leading alternative, Euclidean distance, is the shortest direct distance between two points on a plane. In Section 5, we explore the sensitivity of our analysis to choice of distance measure.

**4. Empirical Strategy**

**4.1 Instrumental variable (IV) regression**

Our identification strategy relies on an established instrument for HIV prevalence in heath behavior regressions (Oster 2012, Chin 2013). As previously demonstrated (Oster 2012, Chin 2013), distance to the origin of the HIV/AIDS pandemic is negatively correlated with local HIV prevalence and plausibly satisfies the exclusion restriction in many health behavior regressions. The mechanism that likely underlies this relationship is that HIV/AIDS arrived earlier in locations closer to the origin (Oster 2012). As the HIV/AIDS pandemic has progressed, the relationship between distance to the origin and HIV prevalence is likely to have weakened (Oster 2012). We address this concern by focusing on the same time period (i.e. a likely pre-steady state period) as in existing literature using this instrument.

The primary regression equation, which closely follows the specification in Oster (2012), is:

$$TFR\_{jc}=β\_{0}+β\_{1}\hat{ln(HIV)\_{jc}}+X\_{jc}^{'}Π+ε\_{jc} \left(2\right)$$

where *TFRjc* is the total fertility rate (TFR) in cluster *j* in country *c*, $\hat{ln(HIV)\_{jc}}$is the fitted cluster-level adult HIV prevalence, $X\_{jc}^{'}$ is a vector of covariates (including indicator variables for Eastern and for Southern Africa, cluster latitude and longitude, indicator variables for decile of cluster latitude and for decile of cluster longitude, and cluster means for age, years of schooling, work outside of the home, urban, Muslim and asset ownership), and $ε\_{jc}$ is an idiosyncratic error term.[[8]](#footnote-8) We aggregate TFR at the cluster-level and do the same for HIV prevalence because the cluster is the level at which our instrument, distance to the origin, varies in our data. In all our regression analyses, we estimate heteroskedasticity-robust standard errors.

We calculate $\hat{ln(HIV)\_{jc}}$ for Equation (2) using fitted values from the following regression equation:

$$ln(HIV)\_{jc}=α\_{0}+α\_{1}distance\_{jc}+X\_{jc}^{'}Φ+η\_{jc} \left(3\right)$$

where $ln(HIV)\_{jc}$ is the natural log of cluster-level adult HIV prevalence[[9]](#footnote-9), $distance\_{jc}$is the distance from cluster *j* in country *c* to the origin of the HIV/AIDS pandemic, $X\_{jc}^{'}$ is the vector of covariates (as in Equation (2)), and $η\_{jc}$ is an idiosyncratic error term.

**4.2 Instrument validity**

**4.2.1 Correlation between distance and HIV**

One of the criteria our instrument must satisfy is that it is correlated with HIV prevalence. Columns (1)-(3) of Table 2 present first-stage estimates of the association between distance to the origin and cluster-level natural log of HIV prevalence. In these three columns, we estimate this first- stage equation using ordinary least squares (OLS) regression. The simple correlation is in Column (1). In Column (2), we control for whether the location is in Eastern Africa, whether it is in Southern Africa, and we control for latitude and longitude. In Column (3), we add the socioeconomic and demographic controls. Eastern and Southern Africa are the sub-regions with the greatest HIV prevalence and countries in these sub-regions may otherwise have had large fertility differences from countries in other sub-regions. Controls for latitude and longitude also help address concerns about unobserved heterogeneity in factors determining fertility possibly associated with distance to the origin of the HIV/AIDS pandemic. Socioeconomic and demographic factors affecting fertility may be correlated with distance to the origin. Throughout, there is a large and statistically significant negative association between distance and the natural log of HIV prevalence. For example, the point estimate from the regression specification with the full set of controls (i.e. Column (3)) suggests that the natural log of HIV prevalence falls by approximately 0.98 for a 1,000 kilometer increase in distance from the origin (p-value<0.01).[[10]](#footnote-10)

**4.2.2 Exclusion restriction**

The second criteria our instrument must satisfy is the exclusion restriction that distance does not affect fertility except through HIV prevalence. Although exclusion restrictions are not testable hypotheses, we can conduct suggestive checks of the identifying assumption. In particular, if the identifying assumption is true, then it is reasonable to presume that distance to the origin should not be correlated with fertility prior to the rise of the pandemic.

We investigate this falsification hypothesis by estimating a reduced-form version of Equation (2) using fertility histories that should not have been affected by HIV because the fertility decisions occurred prior to the rise of the pandemic. In particular, we examine the relationship between distance and TFR in the three-year period 1987-1989 using data from women who were age 15-39 in 1987, 1988, and 1989, using nine DHS countries that have birth histories and cluster GPS coordinates.[[11]](#footnote-11) Table 3 presents the results of this test. The TFR regression appears in Column (1) and the number of surviving children appears in Column (2). The association between distance and TFR prior to the rise of the pandemic is not statistically significant (p-value=0.578) and the point estimate is positive (i.e. 0.4569), consistent with the identifying assumption underlying our main empirical strategy. The association between distance and number of surviving children prior to the rise of the pandemic is not statistically significant (p-value=0.350) and small. If prior to the HIV/AIDS pandemic there was an underlying relationship between distance to the origin of the pandemic and fertility, the point estimates in Table 3 suggests that this relationship was one in which fertility was higher (not lower) further from the origin and the number of surviving children did not vary substantially with respect to the origin.

**5. Results**

**5.1 Baseline results**

Table 4 presents the main regression results. Columns (1) and (2) display the log-linear regressions OLS regressions TFR and number of surviving children. The point estimate in Column (1), -0.0319 (p-value<0.10), indicates that a doubling of HIV prevalence from the sample mean of 6.9% is associated with an approximately 0.03 reduction in TFR, a very small, but statistically significant association. The point estimate in Column (2), -0.0096 (p-value<0.01), indicates that a doubling of HIV prevalence from the sample mean is associated with an approximately 0.01 reduction in number of surviving children, a very small, but statistically significant association.

Instrumenting for HIV prevalence with distance to the origin of the pandemic causes the estimate for ln(HIV) to reverse sign and greatly increase in absolute value. The point estimate in Column (3) of Table 4, 1.9691 (p-value<0.01), indicates that a doubling of HIV prevalence from the sample mean of 6.9% is associated with an approximately 1.97 increase in TFR. Although large and statistically significant, our standard errors are relatively large and we cannot reject the hypothesis of slightly more than a 1.0 increase in TFR in response to a doubling of HIV prevalence. The point estimate in Column (4) of Table 4, 0.5461 (p-value<0.01), indicates that a doubling of HIV prevalence from the sample mean is associated with an approximately 0.55 increase in number of surviving children. Again, although large and statistically significant, our standard errors are relatively large and we cannot reject the hypothesis of around a 0.3 increase in number of surviving children in response to a doubling of HIV prevalence.

Columns (5) and (6) of Table 4 present reduced-form regression results where we regress fertility outcomes on distance to the origin of the HIV/AIDS pandemic. Consistent with the IV results, we find that greater distance from the origin is associated with lower fertility.

**5.2 Robustness check using Euclidean distance**

Table 5 presents the results of a main robustness check. Instead of using great-circle distance, we use Euclidean distance. Column (1) presents the TFR results and Column (2) presents the number of surviving children results. For both outcomes, we find results very similar to those presented in the main analysis.[[12]](#footnote-12)

**5.2 Heterogeneity by age**

Previous economic research has suggested that heterogeneous responses by age may underlie the average fertility response (Durevall and Lindskog 2011). Table 6 displays the results of regressions where the outcome of interest is the age-specific fertility rate, calculated by five-year age group. Panel A presents OLS results and Panel B presents IV results. The OLS results suggest a heterogeneous relationship between HIV prevalence and fertility. The IV results reveal some evidence of an inverted U-shape underlying the average positive effect, with women in the middle of the age distribution increasing their fertility more in response to HIV than women in the early or later childbearing years. This finding is roughly consistent with the inverted U-shape in the effect sizes by age estimated using panel data and individual-level fixed effects models from Malawi in Durevall and Lindskog (2011). Table 7 repeats this exercise for number of surviving children. The results suggest that number of surviving children increased more for women in the middle of the age distribution than women in the early childbearing years.

**6. Discussion**

Our regression analysis yielded three main results. First, the cross-sectional cluster-level OLS estimate of the associations between HIV prevalence and fertility and HIV prevalence and number of surviving children are negative and small. We estimate that doubling local HIV prevalence is associated with a 0.03 reduction in TFR. This result contrasts with the positive cross-sectional national-level association in a similar set of countries in Kalemli-Ozcan (2012). Thus, although fertility may be higher in high HIV prevalence countries, fertility is lower in high HIV prevalence locations when those locations are measured using a more granular method. One explanation for this sub-national relationship is that within countries schooling is simultaneously associated with lower fertility and higher HIV prevalence (Fortson 2008). Consistent with the OLS TFR results, our OLS results for number of surviving children suggest that doubling of HIV prevalence is associated with an approximately 0.01 reduction in the number of surviving children. The negative OLS estimates are not consistent with a typical econometric concern in this setting about reverse causality in which increased fertility (i.e. unprotected sex) causes increased HIV.

Second, the cluster-level association reverses sign when instrumenting for HIV prevalence with distance to the origin of the pandemic. Our point estimate suggests that doubling local HIV prevalence increases TFR by 1.9 births. The fact that instrumenting for HIV prevalence reverses the OLS estimates is consistent with key determinants of fertility such as schooling being associated with HIV prevalence in the cluster-level cross-section. Although we control for mean years of schooling at the cluster level, this is unlikely to fully capture possible omitted variables bias in the OLS estimates driven by heterogeneity in educational attainment that is systematically associated with HIV prevalence. An identification assumption of our instrumentation strategy is that distance to the origin, conditional on other regression covariates such as cluster latitude and longitude, is not associated with observed heterogeneity. Consistent with this assumption, adding controls for socioeconomic and demographic variables in the first-stage regression does not substantially affect our estimate of the relationship between distance and HIV prevalence (i.e. as demonstrated in Columns (3) and (4) of Table 2).

Third, HIV prevalence has increased the number of surviving children and not just TFR. As in the TFR analysis, the cross-sectional cluster-level OLS estimate of the association between HIV prevalence and number of surviving children is negative (and small) and cluster-level association reverses sign when instrumenting for HIV prevalence with distance to the origin of the pandemic. Our point estimate suggests that doubling local HIV prevalence increases number of surviving children by approximately 0.5.

Our main results provide several key policy implications. First, the rise of the HIV/AIDS pandemic appears to have increase total fertility, so HIV prevention efforts may reduce total fertility through reductions in HIV prevalence. Second, population growth appears to have risen in response to the HIV/AIDS pandemic because total fertility increased more than mortality increased, meaning that HIV prevention efforts may reduce population growth.

**7. Conclusion**

We examine the fertility response to the HIV/AIDS pandemic in a sample of sub-Saharan African countries. To identify the effect of local HIV prevalence on fertility, we instrument for HIV prevalence using distance to the origin of the HIV/AIDS pandemic. Our instrumental variables results suggest that the HIV/AIDS pandemic increased total fertility rates (TFR).

 Our main finding that the HIV/AIDS pandemic has increased TFR raises the question as to whether HIV/AIDS has increased the number of surviving children. Although HIV/AIDS likely has increased child mortality, our findings suggest that births have risen disproportionately in response and the pandemic has caused an increase in the number of surviving children. An increase in the number of surviving children likely would have decreased consumption in afflicted sub-Saharan African countries by decreasing GDP per capita through an increase in dependency ratios. The net effect of the HIV/AIDS pandemic on GDP per capita is unknown, as it depends partly on the age distribution of AIDS-related mortality across workers (i.e. non-elderly adults) and non-workers (i.e. children and elderly adults). Future quasi-experimental research should examine the effect of HIV/AIDS on GDP per capita and consumption.

**References**

Acemoglu, Daron and Simon Johnson. 2007. Disease and development: The effect of life expectancy on economic growth. Journal of Political Economy, 115(6): 925-985.

Ainsworth, Martha, Kathleen Beegle, and Andrew Nyamete. 1996. The impact of women's schooling on fertility and contraceptive use: A study of fourteen sub-Saharan African countries. TheWorld Bank Economic Review, 10(1): 85-122.

Aksan, Anna-Maria and Shankha Chakraborty. 2013. Childhood disease and the precautionary demand for children. Journal of Population Economics, 26: 855-885.

Anderson, T. W., and H. Rubin. 1949. Estimation of the parameters of a single equation in a complete system of stochastic equations. Annals of Mathematical Statistics, 20: 46-63.

Angeles, Luis, 2010. Demographic transitions: Analyzing the effects of mortality on fertility. Journal of Population Economics, 23: 99-120.

Becker, Gary and H. Gregg Lewis. 1973. On the interaction between the quantity and quality of children. Journal of Political Economy, 81(2, Part 2): S279-S288.

Behrman, Julia Andrea. 2015. Does schooling affect women’s desired fertility? Evidence from Malawi, Uganda, and Ethiopia. Demography 52(3): 787-809.

Bleakley, Hoyt and Fabian Lange. 2009. Chronic disease burden and the interaction of education, fertility, and growth. Review of Economics and Statistics, 91(1): 52-65.

Bhalotra, Sonia, and Arthur Van Soest. 2008. Birth-spacing, fertility and neonatal mortality in India: Dynamics, frailty, and fecundity. Journal of Econometrics, 143(2): 274-290.

Boucekkine, Rauof, Rodolphe Desbordes, and Helen Latzer. 2009. How do epidemics induce behavioral changes? Journal of Economic Growth, 14: 233-264.

Caldwell, John C., Israel O. Orubuloye, and Pat Caldwell. 1992. Fertility decline in Africa: A new type of transition? Population and Development Review, 18(2): 211-242.

Chin, Yoo-Mi. 2013. Does HIV increase the risk of spousal violence in sub-Saharan Africa? Journal of Health Economics, 32: 997-1006.

Durevall, Dick and Annik Lindskog. 2011. Uncovering the impact of the HIV epidemic on fertility in sub-Saharan Africa: The case of Malawi. Journal of Population Economics, 24: 629-655.

Fink, Gunther and Sebastian Linnemayr. 2009. HIV does matter for fertility: Human capital, mortality and family size. Harvard University working paper.

Fortson, Jane. 2009. HIV/AIDS and fertility. American Economic Journal: Applied Economics, 1(3): 170-194.

Gray RH,Wawer MJ, Serwadda D. 1998. Population-based study of fertility of women with HIV-1 infection in Uganda. Lancet, 351(9096): 98-103.

Juhn, Chinhui, Sebnem Kalemli-Ozcan, and Belgi Turan. 2013. HIV and fertility in Africa: First evidence from population-based surveys. Journal of Population Economics, 26(3): 835-853.

Kalemli-Ozcan, Sebnem. 2012. AIDS, “reversal” of the demographic transition and economic development: Evidence from Africa. Journal of Population Economics, 25(3): 871-897.

Kalemli-Ozcan, Sebnem and Belgi Turan. 2011. HIV and fertility revisited. Journal of Development Economics, 96(1): 61-65.

Karlsson, Martin, and Stefan Pichler. 2015. Demographic consequences of HIV. Journal of Population Economics, 28(4): 1097-1135.

Lam, David. 2011. How the world survived the population bomb: Lessons from 50 years of extraordinary demographic history. Demography 48(4):1231-1262.

Lloyd CB, Kaufman CE, Hewett P. 2000. Implications for fertility change of the spread of primary schooling in sub-Saharan Africa. Population and Development Review. 26(3):483-515.

Lucas, Adrienne. 2013. The impact of malaria eradication on fertility. Economic Development and Cultural Change, 61(3): 607-631.

Malthus, Thomas. 1798. An Essay on the Principle of Population. London: J. Johnson, St. Paul's Churchyard.

McCord, Gordon, Dalton Conley, and Jeffrey D. Sachs. 2017. Malaria ecology, child mortality & fertility. Economics & Human Biology, 24: 1-17.

Ministry of Defence (Navy). 1987. Admiralty of Manual of Navigation, Volume 1: General Navigation, Coastal Navigation and Pilotage. Volume 45 of BR Series. The Stationary Office. Norwich, England.

Osili, Una Okonkwo and Bridget Terry Long. 2008. Does female schooling reduce fertility? Evidence from Nigeria. Journal of Development Economics 87(1): 57-75.

Oster, Emily. 2012. HIV and sexual behavior change: Why not Africa? Journal of Health Economics, 31: 35-49.

WHO. 2011. Causes of Death Summary Tables. Health Statistics and Informatics Department, World Health Organization. Geneva, Switzerland.

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): 423-466.

Young, Alwyn. 2007. In sorrow to bring forth children. Journal of Economic Growth, 12: 283-327.

Table 1. Summary Statics of HIV, TFR, and Distance by Country

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | BF | CM | ET | GH | GN | KE | LS |  |
| HIV Mean | 1.89 | 5.58 | 1.95 | 2.32 | 1.64 | 6.83 | 23.75 |  |
| HIV SD | 3.78 | 6.79 | 4.06 | 3.84 | 3.32 | 9.72 | 16.29 |  |
| TFR Mean | 5.94 | 4.73 | 4.82 | 4.45 | 5.23 | 4.93 | 3.49 |  |
| TFR SD | 2.04 | 2.33 | 2.70 | 3.06 | 1.83 | 4.03 | 2.06 |  |
| Distance Mean (in 1000km) | 3.50 | 1.90 | 2.45 | 3.11 | 4.33 | 1.60 | 2.63 |  |
| Distance SD | 0.13 | 0.16 | 0.25 | 0.14 | 0.20 | 0.18 | 0.05 |  |
| Observations | 397 | 464 | 527 | 410 | 288 | 399 | 380 |  |
|  | ML | MW | SL | SN | SZ | ZM | ZW | Total |
| HIV Mean | 1.25 | 12.93 | 1.48 | 0.90 | 20.90 | 14.16 | 18.13 | 7.82 |
| HIV SD | 2.93 | 16.77 | 3.37 | 2.52 | 10.85 | 10.07 | 9.30 | 11.72 |
| TFR Mean | 6.53 | 5.63 | 4.83 | 5.46 | 3.93 | 6.05 | 3.75 | 5.00 |
| TFR SD | 1.89 | 2.14 | 2.51 | 2.07 | 3.05 | 2.52 | 2.04 | 2.65 |
| Distance Mean (in 1000km) | 3.94 | 1.51 | 4.27 | 4.91 | 2.40 | 1.01 | 1.58 | 2.73 |
| Distance SD | 0.22 | 0.15 | 0.09 | 0.14 | 0.04 | 0.19 | 0.14 | 1.17 |
| Observations | 405 | 512 | 350 | 366 | 270 | 319 | 396 | 5,483 |

Notes: This table shows summary statistics of HIV rate, total fertility rate (TFR), and the distance of a survey cluster from the origin of HIV virus. An observation is a survey cluster. Results are from the DHS for Burkina Faso (BF) 2003, Cameroon (CM) 2004, Ethiopia (ET) 2005, Ghana (GH) 2003, Guinea (GN) 2005, Kenya (KE) 2003, Lesotho (LS) 2004, Mali (ML) 2006, Malawi (MW) 2004, Sierra Leone (SL) 2005, Senegal (SN) 2005, Swaziland (SZ) 2006, Zambia (ZM) 2007, Zimbabwe (ZW) 2005. HIV rates are the percentage of men and women ages 15-49 who are infected with HIV in a cluster, weighted using HIV sample weights. TFR is an aggregation of age-specific fertility rates of 7 age groups (15-19, 20-24, 25-29, 30-34, 35-39, 40-44, 45-49) using data for 3 years before the survey year. Distance is a great-circle distance of the centroid of a survey cluster from the origin of HIV virus.

Table 2. First Stage Estimates of the Effect of Distance on HIV Prevalence

|  |  |  |  |
| --- | --- | --- | --- |
| Dependent Variable: | ln(HIV) | ln(HIV) | ln(HIV) |
|  | (1) | (2) | (3) |
| Specification:  | OLS | OLS | OLS |
|  |  |  |  |
| GC Distance | -0.9392\*\* | -0.9320\*\* | -0.9788\*\* |
|  | (0.022) | (0.189) | (0.190) |
| Latitude |  | -0.0164 | -0.0392\* |
|  |  | (0.018) | (0.018) |
| Longitude |  | -0.0513\* | -0.0408\* |
|  |  | (0.021) | (0.020) |
| Eastern |  | -0.8329+ | -1.0081\* |
|  |  | (0.425) | (0.408) |
| Southern |  | 0.2335 | -0.1967 |
|  |  | (0.520) | (0.504) |
| Age |  |  | -0.0139 |
|  |  |  | (0.012) |
| Education |  |  | 0.1468\*\* |
|  |  |  | (0.018) |
| Work |  |  | -0.0046 |
|  |  |  | (0.121) |
| Urban |  |  | 0.4491\*\* |
|  |  |  | (0.076) |
| Muslim |  |  | -0.0394 |
|  |  |  | (0.106) |
| Asset |  |  | -0.1049\*\* |
|  |  |  | (0.036) |
|  |  |  |  |
| Observations | 5,483 | 5,483 | 5,483 |
| First-stage F | 1764.81 | 24.38 | 26.59 |
| Latitude Decile Dummies | No | Yes | Yes |
| Longitude Decile Dummies | No | Yes | Yes |

Notes: This table shows the first-stage regressions. An observation is a survey cluster. GC Distance is a great-circle distance of the centroid of a cluster to the origin of the HIV virus (in 1000km), Latitude and Longitude are the latitude and longitude of the centroid of a cluster. Eastern is a dummy variable that indicates countries in eastern Africa (Ethiopia, Kenya, Malawi, Zambia, and Zimbabwe), Southern is a dummy variable that indicates countries in southern Africa (Lesotho, Swaziland). Age is the mean age of women, Education is the mean years of schooling of women, Work is a percentage of women who work, Muslim is a percentage of a Muslim household, and Asset is the mean number of durable goods (electricity, radio, TV, refrigerator, bicycle, motorcycle, and car) of a household in a cluster. Urban is a dummy that indicates whether a cluster is an urban area or not. Latitude and Longitude Decide Dummies are dummies variables indicating in which decile of latitude and longitude a cluster is included. Robust standard errors in parentheses. \*\* p<0.01, \* p<0.05, + p<0.10

Table 3. Falsification Tests

|  |  |  |
| --- | --- | --- |
| Dependent Variable:  | Total Fertility Rate | Number of Surviving Children |
|  | (1) | (2) |
| Specification | OLS | OLS |
|  |  |  |
| GC Distance | 0.4569 | -0.1190 |
|  | (0.820) | (0.127) |
| Latitude | -0.0623 | -0.0076 |
|  | (0.081) | (0.015) |
| Longitude | 0.0669 | 0.0256\* |
|  | (0.075) | (0.012) |
| Age | 0.0511 | 0.1557\*\* |
|  | (0.031) | (0.006) |
| Education | -0.2163\*\* | -0.0779\*\* |
|  | (0.038) | (0.009) |
| Work | -0.9196+ | -0.1505\* |
|  | (0.532) | (0.060) |
| Urban | -0.6209+ | -0.1660\*\* |
|  | (0.335) | (0.036) |
| Muslim | 0.4453 | -0.1353\*\* |
|  | (0.305) | (0.042) |
| Asset | -0.0677 | -0.0180 |
|  | (0.114) | (0.016) |
|  |  |  |
| Latitude Decile Dummy | Yes | Yes |
| Longitude Decile Dummy | Yes | Yes |

Notes: This table shows falsification tests using 9 countries of DHS that have both birth records and GIS of the survey cluster (Burkina Faso 1992, Cameroon 1991, Ethiopia 2000, Ghana 1993, Guinea 1999, Mali 1995, Malawi 2000, Senegal 1992-3, Zimbabwe 1999). Column 1 shows the relationship between distance and the total fertility rate of women ages 15-39 pre-1990 (1987-89). Column 2 shows the relationship between distance and the mean number of surviving children per woman in a cluster. Robust standard errors in parentheses. \*\* p<0.01, \* p<0.05, + p<0.10

Table 4. Effect of HIV on Fertility and Surviving Children

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Dependent Variable:  | Total Fertility Rate | Number of Surviving Children | Total Fertility Rate | Number of Surviving Children | Total Fertility Rate | Number of Surviving Children |
|  | (1) | (2) | (3) | (4) | (5) | (6) |
| Specification:  | OLS | OLS | IV | IV | OLS | OLS |
|  |  |  |  |  |  |  |
| ln(HIV) | -0.0319+ | -0.0096\*\* | 1.9691\*\* | 0.5461\*\* |  |  |
|  | (0.017) | (0.003) | (0.464) | (0.121) |  |  |
| GC Distance |  |  |  |  | -1.9274\*\* | -0.5345\*\* |
|  |  |  |  |  | (0.234) | (0.053) |
|  |  |  |  |  |  |  |
| Observations | 5,483 | 5,483 | 5,483 | 5,483 | 5,483 | 5,483 |
| First-stage F |  |  | 26.59 | 26.59 |  |  |

Notes: This table shows the results of the effect of HIV on the total fertility rate and the number of surviving children. An observation is a survey cluster. Columns 1 & 2 show OLS regressions. Column 3 & 4 show IV regressions. Column 5 & 6 show reduced-form regression results. All specifications include the latitude and longitude of the centroid of a cluster, a dummy variable that indicates countries in eastern Africa (Ethiopia, Kenya, Malawi, Zambia, and Zimbabwe), a dummy variable that indicates countries in southern Africa (Lesotho, Swaziland), the mean age of women, the mean years of schooling of women, a percentage of women who work, a percentage of a Muslim household, the mean number of durable goods (electricity, radio, TV, refrigerator, bicycle, motorcycle, and car) of a household in a cluster, dummies for deciles of longitude and latitude, and a dummy that indicates whether a cluster is an urban area or not. Robust standard errors in parentheses. \*\* p<0.01, \* p<0.05, + p<0.10

Table 5. Robustness Checks Using Euclidean Distance

|  |  |  |
| --- | --- | --- |
| Dependent Variable: | Total Fertility Rate | Number of Surviving Children |
|  | (1) | (2) |
| Specification:  | IV | IV |
|  |  |  |
| ln(HIV) | 1.9758\*\* | 0.5484\*\* |
|  | (0.465) | (0.122) |
|  |  |  |
| Observations | 5,483 | 5,483 |
| First-stage F | 26.61 | 26.61 |

Notes: This table shows the effect of HIV on the total fertility rate using Euclidean distance as an instrument. An observation is a survey cluster. All specifications include the latitude and longitude of the centroid of a cluster, a dummy variable that indicates countries in eastern Africa (Ethiopia, Kenya, Malawi, Zambia, and Zimbabwe), a dummy variable that indicates countries in southern Africa (Lesotho, Swaziland), the mean age of women, the mean years of schooling of women, a percentage of women who work, a percentage of a Muslim household, the mean number of durable goods (electricity, radio, TV, refrigerator, bicycle, motorcycle, and car) of a household in a cluster, dummies for deciles of longitude and latitude, and a dummy that indicates whether a cluster is an urban area or not. Robust standard errors in parentheses. \*\* p<0.01, \* p<0.05, + p<0.10

Table 6. Heterogeneous Effects on Age-specific Fertility Rates by Age

|  |
| --- |
| Panel A: OLS regressions |
| Dependent Variable:  | Fertility Rate | Fertility Rate | Fertility Rate | Fertility Rate | Fertility Rate | Fertility Rate | Fertility Rate |
|  |  |  |  |  |  | (6) | (7)  |
| Group: | Age15-19 | Age20-24 | Age25-29 | Age30-34 | Age35-39 | Age40-44 | Age45-49 |
|  |  |  |  |  |  |  |  |
| ln(HIV) | 0.0022\*\* | -0.0014 | -0.0008 | -0.0035\* | -0.0026\* | 0.0015 | -0.0018\* |
|  | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.002) | (0.001) |
|  |  |  |  |  |  |  |  |
| Mean Fertility Rate  | 0.1303 | 0.2356 | 0.2252 | 0.1916 | 0.1318 | 0.0664 | 0.0189 |
| SD Fertility Rate | 0.1299 | 0.1666 | 0.1772 | 0.1865 | 0.1739 | 0.2166 | 0.0996 |
| Observations | 5,483 | 5,483 | 5,483 | 5,483 | 5,483 | 5,483 | 5,483 |
| Panel B: IV regressions |
| Dependent Variable:  | Fertility Rate | Fertility Rate | Fertility Rate | Fertility Rate | Fertility Rate | Fertility Rate | Fertility Rate |
|  |  |  |  |  |  | (6) | (7)  |
| Group: | Age15-19 | Age20-24 | Age25-29 | Age30-34 | Age35-39 | Age40-44 | Age45-49 |
|  |  |  |  |  |  |  |  |
| ln(HIV) | 0.0533\*\* | 0.0415\* | 0.0816\*\* | 0.1073\*\* | 0.0541\* | 0.0516\*\* | 0.0045 |
|  | (0.017) | (0.019) | (0.027) | (0.029) | (0.023) | (0.019) | (0.010) |
|  |  |  |  |  |  |  |  |
| Mean Fertility Rate  | 0.1303 | 0.2356 | 0.2252 | 0.1916 | 0.1318 | 0.0664 | 0.0189 |
| SD Fertility Rate  | 0.1299 | 0.1666 | 0.1772 | 0.1865 | 0.1739 | 0.2166 | 0.0996 |
| Observations | 5,483 | 5,483 | 5,483 | 5,483 | 5,483 | 5,483 | 5,483 |

Notes: This table shows age-specific fertility rate response to HIV. An observation is a survey cluster. Age-specific Fertility rate is the number of babies born during 3 years before the survey year to women in each age group divided by cumulative years spent by women in each age group during 3 years before the survey year. All specifications include the latitude and longitude of the centroid of a cluster, a dummy variable that indicates countries in eastern Africa (Ethiopia, Kenya, Malawi, Zambia, and Zimbabwe), a dummy variable that indicates countries in southern Africa (Lesotho, Swaziland), the mean age of women, the mean years of schooling of women, a percentage of women who work, a percentage of a Muslim household, the mean number of durable goods (electricity, radio, TV, refrigerator, bicycle, motorcycle, and car) of a household in a cluster, dummies for deciles of longitude and latitude, and a dummy that indicates whether a cluster is an urban area or not. Robust standard errors in parentheses. \*\* p<0.01, \* p<0.05, + p<0.10

Table 7. Heterogeneous Effects on Number of Living Children by Age

|  |
| --- |
| Panel A: OLS regressions |
| Dependent Variable:  | Number of Surviving Children | Number of Surviving Children | Number of Surviving Children | Number of Surviving Children | Number of Surviving Children | Number of Surviving Children | Number of Surviving Children |
|  |  |  |  |  |  | (6) | (7)  |
| Group: | Age15-19 | Age20-24 | Age25-29 | Age30-34 | Age35-39 | Age40-44 | Age45-49 |
|  |  |  |  |  |  |  |  |
| ln(HIV) | 0.0023 | 0.0056 | -0.0126\* | -0.0267\*\* | -0.0264\* | -0.0436\*\* | -0.0354\* |
|  | (0.002) | (0.004) | (0.006) | (0.009) | (0.012) | (0.014) | (0.017) |
|  |  |  |  |  |  |  |  |
| Mean Number of Surviving Children  | 0.2293 | 1.1486 | 2.2567 | 3.2957 | 4.1715 | 4.7813 | 5.0508 |
| SD Number of Surviving Children | 0.2990 | 0.6607 | 0.9633 | 1.2864 | 1.5686 | 1.8282 | 1.9846 |
| Observations | 5,321 | 5,314 | 5,264 | 5,116 | 4,903 | 4,627 | 4,301 |
| Panel B: IV regressions |
| Dependent Variable:  | Number of Surviving Children | Number of Surviving Children | Number of Surviving Children | Number of Surviving Children | Number of Surviving Children | Number of Surviving Children | Number of Surviving Children |
|  |  |  |  |  |  | (6) | (7)  |
| Group: | Age15-19 | Age20-24 | Age25-29 | Age30-34 | Age35-39 | Age40-44 | Age45-49 |
|  |  |  |  |  |  |  |  |
| ln(HIV) | 0.0270 | 0.4264\*\* | 0.5224\*\* | 0.8390\*\* | 1.0408\*\* | 1.4494\*\* | 0.4285 |
|  | (0.034) | (0.109) | (0.148) | (0.216) | (0.296) | (0.459) | (0.323) |
|  |  |  |  |  |  |  |  |
| Mean Number of Surviving Children  | 0.2293 | 1.1486 | 2.2567 | 3.2957 | 4.1715 | 4.7813 | 5.0508 |
| SD Number of Surviving Children  | 0.2990 | 0.6607 | 0.9633 | 1.2864 | 1.5686 | 1.8282 | 1.9846 |
| Observations | 5,321 | 5,314 | 5,264 | 5,116 | 4,903 | 4,627 | 4,301 |

Notes: This table shows the effect of HIV on the number of surviving children per woman by age group. An observation is a survey cluster. All specifications include the latitude and longitude of the centroid of a cluster, a dummy variable that indicates countries in eastern Africa (Ethiopia, Kenya, Malawi, Zambia, and Zimbabwe), a dummy variable that indicates countries in southern Africa (Lesotho, Swaziland), the mean age of women, the mean years of schooling of women, a percentage of women who work, a percentage of a Muslim household, the mean number of durable goods (electricity, radio, TV, refrigerator, bicycle, motorcycle, and car) of a household in a cluster, dummies for deciles of longitude and latitude, and a dummy that indicates whether a cluster is an urban area or not. Robust standard errors in parentheses. \*\* p<0.01, \* p<0.05, + p<0.10

1. \* Chin: Department of Economics, Baylor University. Wilson: White House Social and Behavioral Sciences Team and Department of Economics, Reed College, nwilson@reed.edu. Ran Duan and Mark Jarrett provided timely research assistance. The findings, interpretations, and conclusions expressed in this paper are those of the authors and do not necessarily represent the views of the aforementioned individuals or agencies. All errors are our own. [↑](#footnote-ref-1)
2. Oster (2012) and Chin (2013) conduct placebo tests helping to support the validity of this instrumental variable in health behavior regressions and we present similar placebo results. [↑](#footnote-ref-2)
3. Oster (2012) uses distance to the origin as an instrument for HIV prevalence in risky sexual behavior regressions. Chin (2013) uses distance to the origin as an instrument for HIV prevalence in intimate partner violence regressions. [↑](#footnote-ref-3)
4. These 14 countries are also the countries used in Oster (2012). Chin (2013) uses a smaller number of countries because intimate partner violence modules are not offered in all DHS countries. [↑](#footnote-ref-4)
5. HIV prevalence-age profiles for women in high HIV countries typically peak around age 30, as do pregnancy-age profiles, further suggesting that HIV/AIDS may have a larger effect on the survival of later births. [↑](#footnote-ref-5)
6. The cross-sectional analyses of country-and region-level data in Kalemli-Ozcan (2012) suggest that HIV has increased fertility, whereas the time series analysis in Kalemli-Ozcan (2012) suggests positive and negative effects. [↑](#footnote-ref-6)
7. These countries (and survey rounds) are: Burkina Faso 2003, Cameroon 2004, Ghana 2003, Guinea 2005, Kenya 2003, Lesotho 2004, Malawi 2004, Mali 2006, Senegal 2005, Swaziland 2006, Zambia 2007, Zimbabwe 2005. Following Oster (2012), we omit Liberia and Democratic Republic of the Congo because of ongoing large-scale civil conflict in these countries during the period over which we observe fertility behavior. [↑](#footnote-ref-7)
8. The controls in $X\_{jc}^{'}$ are the same as those in the previous literature using distance to the origin of the pandemic as an instrumental variable (Oster 2012, Chin 2013). One exception is that unlike Oster (2012), we do not control for number of children ever born because fertility is our outcome of interest. [↑](#footnote-ref-8)
9. As in Oster (2012), we address a major barrier to using ln(HIV) by assigning HIV prevalence of 0.1% to clusters where zero individuals tested HIV positive in the cluster. [↑](#footnote-ref-9)
10. These estimates are very similar to those in Oster (2012) and slightly smaller in absolute value than those in and Chin (2013). Chin (2013) uses a smaller set of countries (i.e. the DHS surveys that also have IPV modules): Kenya 2008, Liberia 2007, Malawi 2010, Mali 2006, Zambia 2007, and Zimbabwe 2010. [↑](#footnote-ref-10)
11. The sample (i.e., countries and survey rounds) for the falsification test is: Burkina Faso 1992, Cameroon 1991, Ethiopia 2000, Ghana 1993, Guinea 1999, Malawi 2006, Mali 1995, Senegal 1992-93, Zimbabwe 1999. [↑](#footnote-ref-11)
12. As an additional robustness check, we re-estimate the main results omitting one country at a time. Throughout, we find a positive and statistically significant effect of HIV prevalence on TFR. This suggests that it is not a particular country that is driving our results. [↑](#footnote-ref-12)