QUCEH WORKING PAPER SERIES
http://www.quceh.org.uk/working-papers

THE LONG-TERM DETERMINANTS OF FEMALE HIV INFECTION IN AFRICA: THE SLAVE TRADE, POLYGYNY, AND SEXUAL BEHAVIOR

Graziella Bertocchi (University of Modena and Reggio Emilia)
Arcangelo Dimico (Queen's University Belfast)

Working Paper 2015-04

QUEEN'S UNIVERSITY CENTRE FOR ECONOMIC HISTORY
Queen's University Belfast
185 Stranmillis Road
Belfast BT9 5EE
June 2015

The Long-Term Determinants of Female HIV Infection in Africa: The Slave Trade, Polygyny, and Sexual Behavior#

Graziella Bertocchi* and Arcangelo Dimico**

ABSTRACT

We study the long-term determinants of the high rates of HIV infection in sub-Saharan Africa, particularly among women, with a focus on family structure and sexual behavior as shaped by the demographic shock following the transatlantic slave trade. First we show that, in clusters where polygyny is more widespread, HIV infection rates are higher. By instrumenting polygyny with the demographic shock we can also establish that this link is causal. Next we turn to the channels through which polygyny is likely to affect HIV infection by focusing on sexual behavior, as captured by the intensity of sexual activity and the frequency of extramarital partnerships. We document relevant gender differences in behavior: in clusters affected by a larger demographic shock men (but not women) display a more intense sexual activity, while women (but not men) are more likely to engage in extramarital partnerships. We employ these findings to instrument sexual behavior when estimating its influence on HIV infection and we show that clusters exhibiting more frequent female extramarital partnerships are affected by significantly higher infection rates. We interpret our results as follows. The demographic shock induced by the slave trade represents a "primordial" risk factor which is still shaping contemporary family structure and sexual behavior. Polygyny is associated with unsatisfying marital relationships, particularly for women, with consequent female infidelity and an increased risk of infection, which is further multiplied for women co-habiting within polygynous households.

JEL Codes: I15, J12, N37, O10.

Keywords: HIV, polygyny, slave trade, sexual behavior.

We would like to thank participants at seminars at Queen's University Belfast and the University of Luxembourg, and at the conference on the Deep Causes of Economic Development in Utrecht, the CSAE Conference 2015 on Economic Development in Africa in Oxford and the 2015 Royal Economic Society Conference in Manchester, for helpful comments.

* University of Modena and Reggio Emilia, CEPR and IZA. Address: Dipartimento di Economia Marco Biagi, Viale Berengario, 51, I-41121 Modena, Italy. Phone: +39 (0)59 2056856. Email: graziella.bertocchi@unimore.it. URL: http://www.economia.unimore.it/bertocchi_graziella.

** Queen's University Belfast. Address: Queen's Management School, Riddel Hall, 185 Stranmillis Road, Belfast BT9 5EE, UK. Phone: +44 (0)28 90974167. Email: a.dimico@qub.ac.uk. URL: https://sites.google.com/site/arcangelodimico.

1. Introduction

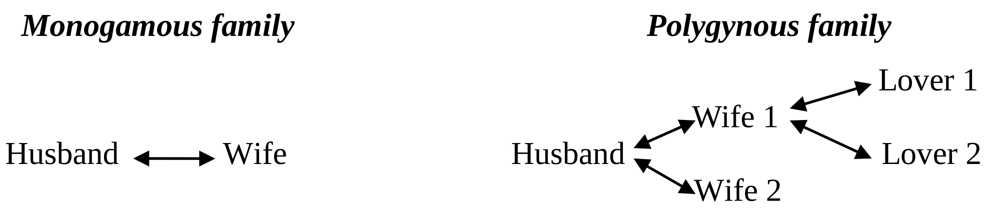
HIV/AIDS is one of the most deathly diseases in Africa. According to the WHO (2012) in 2012 almost 36 million people around the world lived with HIV and 1.6 million people died of AIDS. Almost 75 percent of these deaths (1.2 million) occurred in sub-Saharan Africa. What is peculiar about sub-Saharan Africa is that HIV is much more common among women, who represent nearly 60% of the infected, and 75% of the infected in the 17-24 year-old age group, with young women having a probability of being infected which is almost 5 times larger than for young men (Rehle et al., 2007; Wang, 2010). Violence against women, barriers of access to services, poor education, lack of economic security, and unequal gender norms involving concepts of masculinity that induce unsafe sexual behavior are some of the reasons which can explain this peculiarity within sub-Saharan Africa (UNAIDS, 2013).

In this paper we look at the long-term determinants of HIV infection and its gendered pattern with a focus on the potential influence of the demographic shock that hit Africa as a consequence of the transatlantic slave trade. The available literature has already documented the effect of this historical event on population growth (Fage, 1980; Thornton, 1980; Manning, 1981). The resulting unbalanced sex ratios have also been recognized as one of the factors that favored the diffusion of polygyny in Africa (Thornton, 1983; Dalton and Leung, 2014; Edlund and Ku, 2011; Fenske, 2013). Our present contribution is to uncover the further consequences of the demographic shock for HIV infection as well as for sexual behavior in the present day.

Our hypothesis is that the adoption of polygyny may have affected sexual behavior in the direction of increasing the risk of HIV infection, because of its potential association with more promiscuous sexual habits. On the one hand, within a polygynous family men have sexual relationships with multiple women. In fact, polygyny offers married men the option to take an extra wife rather than engaging in an extramarital relationship. On the other hand, because of the unsatisfying nature of a polygynous marital relationships, women may more frequently become involved in extramarital partnerships (Poku, 2005; Ogundipe-Leslie, 1994; Kwena et al., 2014). Since promiscuous sexual habits represent one of the main channels of transmission of HIV (Halperin and Epstein, 2004; WHO, 2012; Greenwood et al., 2013), one should then expect a statistical association running from the demographic shock – through polygyny and the associated sexual behavior – to HIV infection rates. Moreover, within polygynous families, the epidemic is more likely to be transmitted among co-habitating co-wives, thus multiplying its incidence among women.

Figure 1 illustrates in a simple manner the alternative structures of monogamous and polygynous families. In the former, the husband has sexual relationships with his only wife who, in the absence of competition from other co-wives, is in turn not induced to infidelity. In the latter, the husband has sexual relationships with more than one wife (in practice, most frequently two), some of whom may have been former lovers. At the same time, some of the wives may engage in extramarital partnerships. This pattern magnifies the number of individuals that are directly or indirectly in sexual contact. Our goal is to test the long-term determinants and the contemporary consequences of the cultural and social differences embedded in these alternative marriage institutions.

Figure 1: Family Structure and Sexual Behavior



We start our empirical investigation by revisiting the influence of the slave trades on demographics. We proxy the demographic shock with the rate of change of the population between 1500 and 1860 and we show that across sub-Saharan Africa the slave trades significantly decreased population growth over this period. Based on this finding, we continue our analysis of the long-run determinants of contemporary family structure, sexual habits and HIV infection with a focus on the demographic shock, rather than the intensity of the slave trades, since data on population growth present the advantage of being available at a finer, district level. Moreover, the slave trades have been shown to carry very broad implications, for instance on income and trust (Nunn, 2008; Nunn and Wantchekon, 2011), while the influence we aim at capturing on our outcomes of interest is likely to have been transmitted more specifically through their demographic impact.

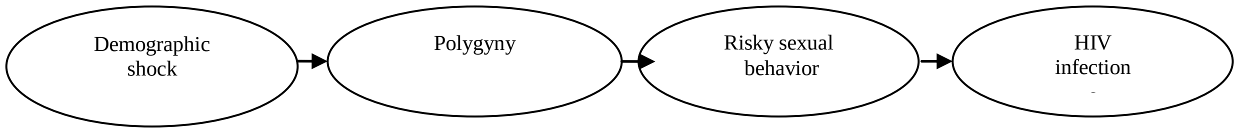
Next, we show that indeed the demographic shock deeply affects contemporaneous family structure, since in areas affected by slower population growth in the period 1500-1860 men are currently more likely to have more than one wife. These findings imply that we can use the demographic shock as an instrument for polygyny when we investigate the relationship between the latter and HIV infection. Through preliminary OLS estimates as well as 2SLS estimates we can therefore establish the presence of a causal link between polygyny and HIV.

The following step of our investigation turns to the channels through which polygyny is likely to affect HIV infection. We focus on sexual behavior, which we capture through two indicators: the intensity of sexual activity and the frequency of extramarital partnerships. We start by assessing the association between these two indicators and the demographic shock, separately for men and women in order to uncover the presence of significant gender differences. Indeed we find that for men a larger demographic shock is associated with a more intense sexual activity, while they are not more likely to have extramarital partnerships. On the other hand, under the same circumstances women have an only marginally more intense sexual activity but, unlike men, they are more likely to have extramarital partnerships. This outcome can be attributed to the fact that neglected and sexually unsatisfied co-wives tend to look for extramarital partners.

Having established a link between the demographic shock on the one hand and the intensity of male sexual activity and female infidelity on the other, we can then employ the demographic shock as an instrument when estimating the effect of sexual behavior on HIV infection, using 2SLS in order to address the obvious reverse causality issue. We find that higher infection rates can indeed be explained by the combined influence of polygyny, i.e., men's decisions to take more than one wife, and women's infidelity, since the fact that unsatisfied women engage in extramarital partnerships increases the rate of transmission of HIV. This outcome can be explained by the fact that, under polygyny, the presence of a wife affected by HIV means that the entire household (husband, co-wives, and children) is likely to be affected, with a consequent significant increase in the spread of the epidemic. Moreover, polygynous cohabitation patterns, with several women sharing the same roof with a single husband, multiply female exposure to the risk of infection.

Figure 2 summarizes the links discussed above. The demographic shock following the slave trade contributed to the diffusion of polygyny. The latter induces risky sexual behavior that in turn represent a vehicle for HIV transmission, with multiplying effects on women's infection rates.

Figure 2: From the Demographic Shock to HIV Infection



The rest of the paper is organized as follows. Section 2 discusses the related literature. Section 3 describes the data and the empirical model. Sections 4 contain our main results. Transmission channels are explored in Section 5. In Section 6 we derive our conclusions and policy implications.

2. Literature Review

The paper is related to several streams of the literature. The first stream looks at the economics of HIV in Africa. Most of this literature is based on randomized field experiments. For instance, Auvert et al. (2005) and Gray et al. (2007) use field experiments to evaluate the effectiveness of prevention policy and find that male circumcision lowers the infection risk for males. For a field experiment in Kenya, Dupas (2011) evaluates the effect of providing information to girls about the HIV status of different groups of men on sexual behavior. De Walque (2007) finds that HIV/AIDS information campaigns in Uganda are more effective among educated people. Ashraf, Bandiera and Jack (2014) compare different incentive schemes for hairdressers to sell condoms as a prevention policy. Duflo, Dupas and Kremer (2015) evaluate school-based HIV/AIDS interventions in Kenya. Contributions not based on field experiments include the following. Young (2005, 2007) simulates the impact of AIDS on fertility and human capital within a household behavior model. Oster (2005) models a dynamic diffusion model of HIV/AIDS and using cross-country information about sexual behavior finds that the vast differences in HIV prevalence rates between the US and Africa can be attributed to high levels of untreated sexually transmitted diseases, while the differences in HIV prevalence rates within Africa can be attributed to differences in sexual behavior. Greenwood et al. (2013) develop a choice-theoretic general equilibrium search model to analyze the Malawian epidemic and calibrate it under the assumption that men and women have different transmission rates. Magruder (2011) explicitly models men and women within a matching model of marital search applied to the analysis of the HIV/AIDS epidemic in South Africa. Reniers and Watkins (2009) empirically document that the spread of HIV in sub-Saharan Africa is lower where the practice of polygyny is common and propose lower coital frequency in polygynous marriages as an explanation.

Our work also builds on previous research on the long-term effects of the slave trade. Over the period 1529-1850 almost 12,500,000 Africans were forced to undertake the Middle Passage (Eltis, 2008). Fage (1980), Thornton (1980), and Manning (1981) argue that the slave trade caused a huge decrease in the rate of growth of population and a skewed distribution of the African population, with women outnumbering men by a ratio which could be larger than three women per man (Thornton, 1980). According to Ciment (2007) the estimated population of sub-Saharan Africa in 1850 was at about 50 million, while absent the slave trade it would have been around 100 million. The resulting unbalanced sex ratio provoked a further reduction of the population growth rate and a deep transformation of family structure and sexual behavior. The institution of polygyny was reinforced. Matrilineal lineage faced a serious thread. Co-wife competition together with a larger wife-husband age gap further contributed to a decline in population growth (Strassmann, 2000; Lamie, 2007).[1] General attitudes toward gender roles were also affected (Teso, 2014). A persistent negative effect of the slave trade on development is found by Nunn (2008), while Nunn and Wantchekon, (2011) highlight its deleterious influence on trust.

Building on Thornton (1983), the specific link between the slave trade and contemporaneous polygyny is investigated by Dalton and Leung (2014), Edlund and Ku (2011), and Fenske (2013).[2] Dalton and Leung (2014) combine the historical data on the slave trades from Nunn (2008a) and Nunn and Wantchekon (2011) with contemporaneous polygyny data from the female Demographic and Health Surveys (DHS). They show that the higher polygyny rates in Western Africa, if compared to Eastern Africa, can be explained by the slave trades, since the preference for male slaves, leading to unbalanced sex ratios, was a distinctive feature of the transatlantic trade out of Western Africa. Using a measure of polygyny given by the number of married women over the number of married (based on United Nations data) Edlund and Ku (2011) argue that polygyny in sub-Saharan Africa is characterized by late but nearly universal marriage for men (general polygyny). Unlike other forms of polygyny involving a large number of wives for high-quality men only, which can be explained by wealth inequality, they find that general polygyny is linked to the slave trade. Within a broad analysis of the determinants of African polygyny, Fenske (2013) uses the DHS household-level codification and confirms the existence of a link with the slave trades (even though he

[1] Historical accounts are provided for instance by Brasio (1958), Labat (1728) and Texeira de Mota (1974).
[2] The economic literature on polygyny includes Becker (1974), Tertilt (2005), Gould, Moav and Simhon (2008), and Lagerlof (2010).

finds that the latter can only predict polygamy across broad regions, i.e., West vs. East Africa).

For an anthropological and historical account of family structure in Africa we refer to Boserup (1970), Goody (1973), and Todd (1984), who point to a number of reasons that can explain the higher incidence of polygyny in Africa, if compared to other continents. Early agricultural conditions which made female labor valuable and reliance on females for production provoked a tendency to multiple marriages[3] which was later reinforced because of the slave trade, to sustain reproduction success and lineage subsistence. In such a context polygyny ensured no shortage of potential husbands and maximized women's chances of pregnancy. At the same time child survival was enhanced through longer breastfeeding patterns and inter-birth intervals (Caldwell and Caldwell, 1987). The importance of ecological factors is confirmed by Lee and Whitbeck (1990) and White and Burton (1988) who find an association between type of agriculture, female labor, and polygyny. The literature has also highlighted the high social costs associated with polygyny, with an emphasis on the fact that it makes it difficult to successfully control the diffusion of sexually transmitted diseases (Caldwell, Orubuloye and Caldwell, 1992), In particular, polygyny has been associated with a higher diffusion of HIV among women. The higher rate of HIV among women is explained by the centrality of men in the household, which increases the likelihood of HIV among women exponentially. Beside the multiplicative effect above, there are other potential reasons why HIV in polygynous groups is higher among women. As mentioned in the Introduction, neglected co-wives tend to look for extramarital partnerships (Poku, 2005; Ogundipe and Leslie, 1994) increasing the probability of HIV infection for the entire household. Transactional sex is more frequent since poor economic conditions induce young women/wives to accept gifts in exchange of sex with older man (Mbirimtengerenji, 2007; Luke and Kurtz, 2002).[4]

Since we interpret the demographic shock as a "primordial" driver of sexual behavior and of the diffusion of sexually transmitted diseases through its persistent influence on social customs and mating patterns, our contribution is also linked to the literature on persistence of culture, social preferences, and sexual behavior. While these traits generally tend to persist

[3] Jacoby (1995) provides an empirical test of this hypothesis.
[4] Evidence from studies on mammals (Ashby and Gupta, 2013) shows that polygyny is normally associated with more sexual activity and a higher rate of sexually transmitted diseases.

over time (Bisin and Verdier, 2001; Alesina and Giuliano, 2013; Giavazzi, Petkov and Schiantarelli, 2014), they can also be altered significantly and persistently by external shocks. For example Francis (2008) shows that the outbreak of HIV in the US in the 1990s caused a permanent change in homosexual relationships. Nunn (2012) provides an extensive review of why cultural values and beliefs are important when studying the process of historical economic development.

Finally the paper shares features with the literature on long-term development which distinguishes between proximate and deep factors of growth (Hall and Jones, 1999; Acemoglu, Johnson, and Robinson, 2002; Galor and Ashraf, 2013). While we consider household composition and sexual behavior as proximate determinants (channels) of HIV, these channels are determined by deeper factors which in our case consist in the demographic shock occurred during the slave trade. As a result the demographic shock indeed represents what in the literature on long-term development would be identified as a deep, "primordial" determinant of HIV, since it is the factor which drives proximate outcomes.

3. Data and Model Specification

3.1. Data

The source of the data we use for information on HIV, polygyny and sexual behavior is the Demographic and Health Survey (DHS).[5] In order to select the countries to be included in our sample we adopt the following criteria. First, since we need to spatially match DHS data with population data, we only consider countries for which Global Positioning System (GPS) records are available. Among countries for which we have GPS data we then choose those for which we also have HIV data (from the HIV dataset of the DHS). For Benin, Nigeria, and Madagascar GPS and DHS, but no HIV data, are available. For Uganda, Tanzania, and Mozambique GPS and HIV datasets are available but the latter can only be matched with the AIDS Indicators Survey (AIS)[6] dataset (instead of DHS) which only provides information on women. In all cases the most recent survey is chosen. We start with the VI wave and then if

[5] DHS surveys are conducted at a cluster level. A cluster contains about 100 households, out of which a predetermined number of households are interviewed. Within each cluster, the number of households selected for male interviews is usually considerably lower than for female. See ICF International (2012) for details on the DHS.

[6] The main difference between the AIS and the HIV datasets is that the former is based on a survey which is conducted to monitor the effectiveness of HIV/AIDS programs. Therefore it reports information on attitudes toward HIV. Instead, the HIV dataset provides information on infection rates based on blood tests.

no data is available for the VI wave we turn to the V or IV waves. Countries and waves included in the final sample are reported in Table A1.[7]

HIV infection is measured by a dummy variable for whether an individual is positive to the blood test for HIV. Our proxy for polygyny is a dummy variable taking value one when the number of co-wives is reported to be above one, and zero otherwise. For the household dataset, the dummy takes a value one when in the household there is at least one individual who reports polygyny, and zero otherwise.[8] Sexual behavior is described by two variables: the intensity of sexual activity and the frequency of extramarital partnerships. We proxy for the former with a dummy variable which is equal to one if the individual has had any sexual activity within the last 4 weeks. For the latter we use the number of sexual partners (spouses excluded) within the last 12 months (standardized with mean equal to zero and standard deviation equal to one). More details on variable definitions and sources are available in Table A2 in the Appendix. Table A3 reports summary statistics.

By merging the HIV dataset with the household dataset and with the male and female datasets, we obtain a sample of 263,517 observations. To be noticed is that this sample is much smaller than the household and the sum of the male and female samples, since the blood test for HIV is not compulsory. This means that in the HIV sample it is highly likely to have selection. It is reasonable to expect that selection causes a downward bias since infected individuals should be less keen to be tested. This is confirmed by Mishra et al. (2006), who find that the rate of HIV infection among individuals not tested for HIV is systematically larger than the rate among individuals (both male and female) tested for HIV. The average rate of HIV infection in our sample is around 6 percent and the overall level of diffusion of polygyny is over 20 percent. Almost 50 percent of individuals report to have had sex within

[7] One problem with DHS data is that the distribution of the population in the sample is not entirely representative of the country population. In fact, rather than sampling individuals independently, the DHS samples individuals as a group (cluster), which causes a larger sample-to-sample variability than sampling individuals directly. In addition regions with a smaller population tend to be over-sampled while regions with larger populations tend to be under-sampled, causing estimates to be biased. For this reason, in order to get point estimates and correct standard errors, we use sampling weights and clusters, so that estimates will be weighted by probability sampling weights provided in each sample and standard errors will be clustered at a primary sampling unit. This procedure is equivalent to estimate an OLS for survey data in which – as described in the DHS Guidance – data is stratified by region and urban areas, organized by primary sampling units, and weighted by probability weights.

[8] Our codification based on the household sample allows to overcome a problem of under-reporting associated with polygyny. Within the same household the husband may report polygyny while information is missing for women, since the latter tend not to respond to this sort of questions, as discussed by Mitsunaga et al. (2005). We also refer to Timæus and Reynar (1998) for a discussion of the correspondence between husbands and wives in survey responses to questions about polygyny.

the last 4 weeks and almost 20 percent of them report to have had extra-marital partners in the last 12 months.

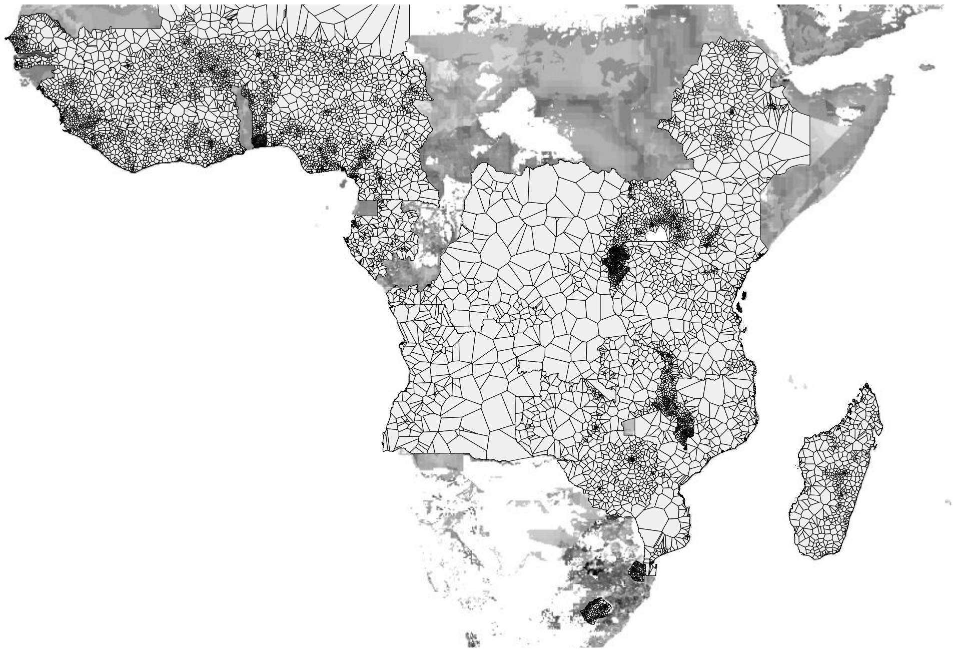
In order to proxy for the demographic shock occurred during the slave trade we use information on the rate of growth of the population in sub-Saharan Africa countries over the period running from 1500 to 1860.[9] The History Database of the Global Environment (HYDE) represents the source for data on population in 1500 and 1860 (see Table A2 in the Appendix). HYDE provides disaggregated spatial data on population over the period 0-2000 AD which is estimated using historical data at the national and subnational level.[10] The reason why we focus on the period from 1500 to 1860 is because this timeframe roughly coincides with the beginning and the end of the slave trade period.

To perform our investigation, we spatially join data on population growth with data from the DHS. In Figure 3 we overlay our map for population growth over the period 1500-1860 with DHS clusters. The darker is the area in the map, the lower the rate of population growth experienced over the time period, where -71.6 percent is the minimum and 162 percent is the maximum. As expected West African regions along the Gold Coast and the Bight of Biafra and Benin are those for which there has been a more significant decline in the population over the slave trade period. On the other hand regions in Eastern Africa have experienced an increase in population though some of these regions (e.g., Madagascar, Mozambique, Zanzibar) were also affected by the transatlantic slave trade. White areas represent missing data which reflect the fact that HYDE does not report data for areas for which no historical sources are available.

[9] The rate of population growth is computed as the difference in log between raster data for population in 1500 and population in 1860.

[10] A detailed description of the methods used to estimate global historical population is provided by Klein Goldewijk (2005) and Klein Goldewijk, Beusen and Janssen (2010). Roughly described, these methods consist in using historical sources (i.e., McEvedy and Jones, 1978; Livi Bacci, 2007; Maddison, 2003; Denevan, 1992) supplemented with sub-national population figures (i.e., Lahmeyer, 2004). These records are then plugged into population models in order to estimate data at the subnational level. This method is similar to that used by Fage (1980), Thornton (1980), Manning (1981), and Frankema and Jerven (2014) to estimate population during the slave trade and to study the impact of the latter on population growth and polygyny.

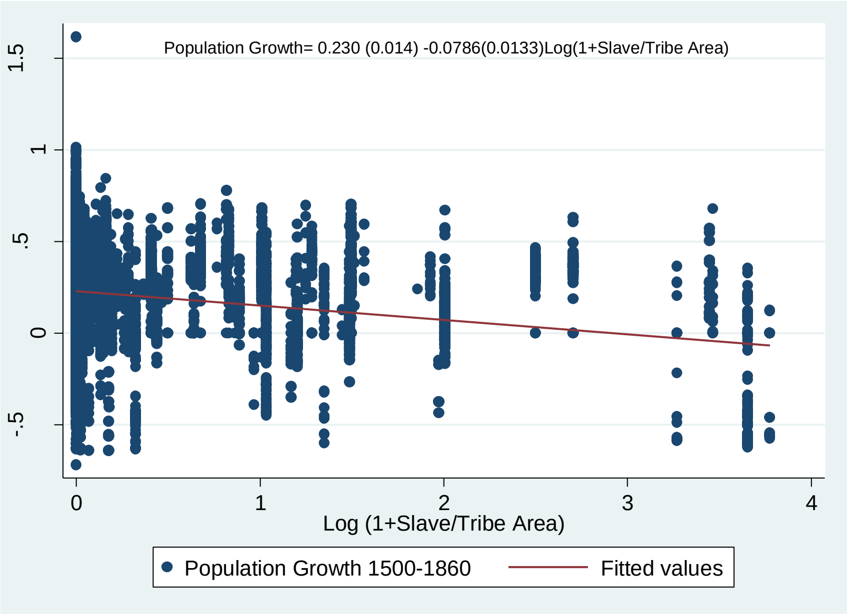
Figure 3: African Population Growth in 1500-1860



Point GPS coordinates are transformed in polygons using a Voronoi algorithm.
Darker shades indicate lower rates of population growth. White areas represent missing values.

To illustrate the link between population growth and the slave trades, in Figure 4 we plot our measure of population growth during the slave trade period against data on the number of slaves at an ethnic level normalized by the total ethnic surface area. Data are from Nunn and Wantchekon (2011). The plot shows a significant and negative pairwise correlation between the transatlantic slave trades and the change in the population over the specified time period. On average the rate of population growth over the period 1500-1860 decreases by over 0.07 per a one standard deviation change in the log of the number of slaves. As shown in Table A4 in the Appendix, controlling for a full set of geographical covariates this effect becomes over 0.03 per a one percent standard deviation in the slaves variable (Model 2). The effect increases significantly when we estimate the same relationship with an IV estimator in which – as in Nunn and Wantchekon (2011) and Dalton and Leung (2014) – we use distance from the coast as an instrument for the slave trades: the estimated decrease in population growth is now over 0,12 per a one standard deviation in the log of the slave trades.

Figure 4: Population Growth and the Slave Trades



There are two possible objections to the use of our proxy for the demographic shock. The first objection is why we choose to use data on population growth over the period 1500-1860 rather than directly using data on slave trades from Nunn and Wantchekon (2011). The second objection is related to the quality of our data.

The answer to the first objection is two-fold. First, polygyny in West Africa developed because of the demographic shock and not because of the slave trade itself. The latter could in fact correlate with other social effects (e.g., trust). Therefore, we believe a demographic variable should capture the effect of the shock more accurately than an overall measure of the intensity of the slave trades. The second and more important reason why we prefer to use data on population growth during the slave trade period is much more practical. Data from Nunn and Wantchekon (2011) are at an ethnic-group level and because of that they display almost no variation at a country/district level, while in our analysis district fixed effects are important to capture the effect of potential omitted variables (e.g., the presence of HIV treatment centres in the district). This problem is also noticed by Fenske (2013), who reports that the effect of the slave trades on polygyny becomes completely insignificant once controlling for country fixed effects.

Regarding the second objection, which is related to the quality of the data from HYDE, we consider the relationship between the intensity of the slave trades and the demographic shock, illustrated in Figure 4, as well as the fact that HYDE only reports data for areas for which there are historical sources, as signals of reliability of the data. However to have further reassurance about their quality we also undertake a series of tests.

According to Eltis and Richardson (1997) slave raids were more frequent in regions with a higher population density. Therefore, in areas affected by the slave trade we should expect a negative relationship between population in 1500 and subsequent population growth. On the other hand, in regions not affected by the slave trade you are likely to observe a kind of Malthusian trap where more wealth (which we proxy by population in 1500) leads to higher population growth. For this reason we split our sample between countries on the West Coast and countries on the East Coast[11] in order to evaluate whether the underlying factor (i.e., population in 1500) which can explain population growth over the next centuries exerts a different influence across these two sub-samples. While countries on the East Coast have been affected by the Indian Ocean slave trade, according to the existing literature this wave of trade did not exert significant demographic effects, because it was much more limited in size and because women tended to be preferred over men. Table 1 shows results from this experiment and as expected population in 1500 has a negative and significant effect on the demographic shock in West Africa (Model 1). On the other hand, the effect of population in 1500 on the demographic shock for East African countries is positive but not significant (Model 2). A similar pattern is found for distance from the coast, whose effect for West African countries is positive (the further the cluster is from the coast the larger is population growth) given that clusters closer to the coast are more likely to be affected by the shock (Model 3). In East Africa distance from the coast is instead not significant given the limited extent of the slave trade (Model 4).

Taken together, this evidence clearly documents the presence of a demographic shock significantly linked to the slave trades. In the following analysis we investigate the influence of this shock on contemporaneous family structure and sexual behavior and, through these channels, on HIV.

[11] Madagascar is excluded because it was affected by the transatlantic slave trade.

Table 1: Heterogeneity of the Demographic Shock

	Dependent Variable: Population Growth 1500-1860			
	Model 1	***Model 2***	***Model 3***	***Model 4***
Estimation Method: OLS	***West Africa***	***East Africa***	***West Africa***	***East Africa***
Population in 1500	-0.0488**	0.00382		
	(0.0157)	(0.0216)		
Distance from the Coast			0.0177***	0.000112
			(0.00557)	(0.00519)
Observations	4,175	4,032	4,175	4,032
RMSE	0.103	0.136	0.106	0.136

Country clustered robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.
Additional controls include:
1) Cluster level controls: distance from the coast, soil fertility, annual precipitation, annual temperature, net primary productivity, slave density, water availability, altitude and forest land cover.
2) District fixed effects.

3.2. The Empirical Model

In order to investigate the links among the relevant variables we estimate variants of the following model:

$$Y_{i,j,d,c} = \sum_{d=1}^{D} \pi_d .1_{D_i=d} + \beta_1 X_{i,j,d,c} + \beta_2 Z^1{}_{j,d,c} + \beta_3 Z^2{}_{i,j,d,c} + \nu_{i,j,d,c} \qquad (1)$$

where $Y_{i,j,d,c}$ represents an outcome variable for individual *i* in DHS cluster *j* and district *D* and $X_{i,j,d}$ represent the main regressor, which depending on the specification will be either polygyny, or the level of HIV infection at the individual level, or sexual behavior again at the individual level, or population growth over the period 1500-1860 at a cluster level ($X_{j,d}$). We include district ($\sum_{d=1}^{D} \pi_d .1_{D_i=d}$) fixed effects in order to control for institutional and policy factors (e.g., health programs). The model also includes a set of cluster-level historical, geographical, and economic controls denoted by $Z^1{}_{j,d}$ (e.g., distance from the coast, soil fertility, annual precipitation, annual temperature, net primary productivity, mean GDP in 2000, population density in 2000, water availability, altitude, and forest land cover) and individual and household controls $Z^2{}_{i,j,d}$ (e.g., age, gender, individual income, education, occupation, religion, age and gender of the household head, cohabitation duration, and ideal number of children). A proxy for the transatlantic slave trades (total number of slaves at an ethnic group level over ethnic homeland surface area) is used to control for a possible effect through trust (Nunn and Wantchekon, 2011). Description, sources and summary statistics for each variable are reported in the Appendix (Tables A1-A3).

4. HIV, Polygyny, and the Demographic Shock

In Table 2 we start by examining the relationship between HIV infection and polygyny. The dependent variable is a dummy for whether the individual is positive to the blood test for HIV. As mentioned in the Data section, the HIV data are likely to be affected by selection, since infected individuals are less keen to take the test. However, since selection tends to be more severe in areas with a higher incidence of HIV, this will tend to under-estimate the effect of polygyny on HIV. The main regression is our proxy for polygyny. In Model 1 we merge the household dataset with the male and female datasets and we find a significant and positive effect of polygyny on HIV. In Model 2 we confine estimates to males only and, although the coefficient is smaller, the effect is still significant at a 1 percent. Finally in Model 3 we confine estimates to females only and the marginal effect of polygyny on HIV is almost equal to 1.9 percent (almost 0.6 percent larger than the estimated effect among men).

Table 2: HIV Infection and Polygyny - OLS

	Dependent Variable: HIV (1 if positive)		
Estimation Method: OLS	***Model 1***	***Model 2***	***Model 3***
Polygyny	0.0170***	0.0132***	0.0188***
	(0.00198)	(0.00255)	(0.00259)
Observations	164,484	71,219	93,265
RMSE	0.221	0.199	0.237
Sample	Entire Sample	Men	Women

Primary sampling unit clustered robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Additional controls include:

1) Cluster level controls: distance from the coast, soil fertility, annual precipitation, annual temperature, net primary productivity, mean GDP 2000, population density 2000, slave density, water availability, altitude and forest land cover..

2) Individual level controls: age, gender, individual income, education, occupation and religion.

3) Household level controls: gender of the household head, relationship to the head, relationship structure and household wealth.

4) District fixed effects.

However these effects are likely to be affected by reverse causality, since while it is reasonable to expect that a higher diffusion of polygyny may increase the spread of HIV, at the same time a higher incidence of HIV may also affect negatively the probability of polygyny. Table 3 reports the link between population growth between 1500 and 1860 and polygyny. In Model 1 we employ the entire sample and we find a positive relationship between the demographic shock and polygyny. The estimated effect is significant at a 1 percent level, which suggests that individuals in clusters which have been more exposed to the historical shock show a higher incidence of polygyny in the present day.[12] The relationship is still significant (at a 5 percent level) in Model 2 over the male sample. Over the female sample Model 3 also shows a significant coefficient, despite the fact that, as previously discussed, women tend to under-report polygyny.

Table 3: Polygyny and the Demographic Shock - OLS

	Dependent Variable: Polygyny		
Estimation Method: OLS	***Model 1***	***Model 2***	***Model 3***
Population Growth 1500-1860	-0.0434***	-0.0395**	-0.0468*
	(0.0163)	(0.0183)	(0.0179)
Observations	164,578	71,264	93,314
RMSE	0.341	0.31087	0.3602
Sample	Entire Sample	Men	Women

Primary sampling unit clustered robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.
Additional controls include:
1) Cluster level controls: distance from the coast, soil fertility, annual precipitation, annual temperature, net primary productivity, mean GDP 2000, population density 2000, slave density, water availability, altitude and forest land cover.
2) Individual level controls: age, gender, individual income, education, occupation and religion.
3) Household level controls: gender of the household head, relationship to the head, household income and religion.
4) District fixed effects.

In Table A5 we replicate Model 2 of Table 3 over the same sample (which is the most reliable since it is not likely to suffer from under-reporting) while using the slave trades in place of the demographic shock as a regressor. The slave trades variable is significant without district (or country) effects (Model 1) but it turns insignificant when district fixed effects are included (Model 2). These results are consistent with Fenske (2013) and represent one of the reasons, as anticipated in the data section, why we prefer to use population data rather than the slave trades data from Nunn and Wantchekon (2011), since the latter display almost no variation at a country/district level. However, since the demographic shock affecting

[12] To be noticed is that, since we control for age for each individual in the household (i.e., wife and husband), we account for the potential influence on polygyny of the age gap, as suggested by Edlund and Ku (2011).

population growth over the period 1500-1860 largely depends on the slave trades, as shown in the previous section, our results so far confirm the existing evidence (i.e., Dalton and Leung, 2013; Edlund and Ku, 2011; Fenske, 2013) on the link between the slave trades and polygyny.

Taken together, from an empirical point of view results in Tables 2 and 3 do suggest a structural relationship between the demographic shock, polygyny, and HIV. Therefore the demographic shock may represent the underlying factor which may affect the spread of HIV through its effect on polygyny. This structural relationship between the demographic shock, polygyny, and HIV can be represented by the following simultaneous equation:

$$HIV_{i,j,d,c} = \alpha + \beta_1 \, Polygyny_{i,j,d,c} + \beta_2 X_{i,j,d,c} + u_{i,j,d,c} \quad (2)$$

$$Polygyny_{i,j,d,c} = \lambda + \gamma Demographic\ Shock_{j,d,c} + v_{i,j,d,c} \quad (3)$$

where in 2SLS estimates Equation 3 represents the first stage regression while Equation 2 represents the second stage.

The identification of the effect of polygyny on HIV in Equation 2 relies on the exogeneity of the demographic shock which of course cannot be tested. Instrument validity requires the demographic shock to affect HIV infection only through polygyny, while alternative channels of influence need to be excluded. Instrument validity could be violated if, for example, the demographic shock has increased the disease hazard in an area (e.g., through an increase in the incidence of poverty or changes in the population's behavior) which may imply that the instrument captures this higher environmental hazard. For this reason in Table 4 we run placebo regressions in which we test the effect of the demographic shock on some measures of disease/malnutrition provided by the DHS. In Model 1 we test whether the demographic shock affects iodine disorder deficiency, which is one of the most frequent sources of impaired cognitive development, being often related to poor nutrition of women during pregnancy. In Model 2 and Model 3 we check whether the instrument has any effect on the individual age/height percentile and body mass index, which are also normally used as measures of health and malnutrition. In Model 4 and Model 5 we test the effect on haemoglobin and anaemia levels,[13] which are normally related to malaria and sickle cells.

[13] Anaemia and haemoglobin levels are of course highly correlated.

Finally, in Model 6 we test the effect on GDP at a cluster level. For all models in Table 4 we find an insignificant effect of the demographic shock on each of these proxies of health/malnutrition/poverty, which should attenuate potential concerns related to alternative channels through which the demographic shock may affect HIV.

Table 4: Placebo Regressions

Dependent Variables:	***Iodine Test***	***Age/Height Percent.***	***Body/Mass Index***	***Haemoglobin Level***	***Anaemia***	***Mean GDP***
	Model 1	***Model 2***	***Model 3***	***Model 4***	***Model 5***	***Model 6***
Population Growth 1500-1860	0.312	0.0329	-0.00955	-0.00427	0.0091	-0.221
	(0.423)	(0.0473)	(0.00645)	(0.00577)	(0.0196)	(0.177)
Observations	73,600	132,480	132,547	140,063	140,048	9,401
R-squared	5.983	1.603	0.159	0.167	0.603	0.495

Primary sampling unit clustered robust standard errors in parentheses for Models 1-5. Country clustered robust standard errors in parentheses for Model 6. *** p<0.01, ** p<0.05, * p<0.1.

Additional controls include:

1) Cluster level controls: distance from the coast, soil fertility, annual precipitation, annual temperature, net primary productivity, mean GDP 2000 (except in Model 6), population density 2000 (except in Model 6), slave density, water availability, altitude and forest land cover.
2) District fixed effects.

In Table 5 we run additional placebo regressions to exclude that our proxy for the demographic shock may be affected by pre-colonial institutions which, according to Gennaioli and Rainer (2007) and Michalopoulos and Papaioannou (2013), have an impact on development. We regress our proxy for the demographic shock at a cluster level on proxies for jurisdictional hierarchy of the local community (Model 1) and jurisdictional hierarchy beyond the local community (Model 2) – both provided by Murdock (1967)[14] – and we do not find any statistical association between these variables, which should exclude any institutional effect of the demographic shock on institutions. Effects of current institutions are automatically ruled out by fixed effects.

[14] See Murdock (1967) for the codification of the two variables.

Table 5: Placebo Regressions

Dependent Variable:	***Population Growth 1500-1860***	
	Model 1	***Model 2***
Jury. Hierarchy Loc. Community - 3 Levels	-0.0125	
	(0.0223)	
Jury. Hierarchy Loc. Community- 4 Levels	-0.0136	
	(0.0212)	
Jury. Hierarchy Beyond Loc. Community- 1 Level		-0.0118
		(0.0201)
Jury. Hierarchy Beyond Loc. Community- 2 Levels		-0.00141
		(0.0223)
Jury. Hierarchy Beyond Loc. Community- 3 Levels		-0.0096
		(0.0287)
Observations	5,725	5,717
RMSE	0.106	0.106

Primary sampling unit clustered robust standard errors in parentheses.*** p<0.01, ** p<0.05, * p<0.1.

Additional controls include:

1) Cluster level controls: distance from the coast, soil fertility, annual precipitation, annual temperature, net primary productivity, slave density, water availability, altitude and forest land cover.

2) District fixed effects.

Since our analysis does provide support for the use of the demographic shock as a suitable instrument for polygyny, we now proceed by re-evaluating the effect of polygyny on HIV according to the above system of simultaneous equations.

Table 6 shows 2SLS results for the effect of polygyny on HIV for the full sample. In order to maximize the sample size, Model 1 does not include household and individual effects. First-stage statistics confirm results in Table 2 by showing a significant effect of our proxy for the demographic shock on polygyny. The estimated effect of polygyny on HIV for Model 1 is almost 0.81, meaning that the probability of HIV for polygynous households is almost 81 percent higher. This effect is consistent with the distribution of household members between polygamous and monogamous families in our sample, given that the typical polygamous family is composed by two co-wives and the husband. This means that in a polygamous household there is an almost 100 percent higher unconditional probability of being affected, since the number of wives is twice their number in a monogamous household. In Model 2 we include household and individual controls and the effect of the demographic shock on polygyny in the first-stage statistics is still significant. The estimated effect of polygyny on

HIV for this model is to increase the probability of HIV by almost 53 percent. Compared to Model 1 the probability decreases by about 28 percent.

Table 6: HIV Infection and Polygyny - IV

	Dependent Variable: HIV (1 if positive)		
Estimation Method: 2SLS	***Model 1***	***Model 2***	***Model 3***
Polygyny	0.807*	0.530*	0.541**
	(0.452)	(0.276)	(0.210)
		First Stage Statistics	
Population Growth 1500-1860	-0.03620**	-0.0434***	-0.0429***
	(0.0174)	(0.0163)	(0.0163)
Rugged Terrain			-0.00620**
			(0.0025)
Anderson LR-Stat.	26.136	34.946	56.484
Cragg-Donald F-Stat.	26.136	34.939	28.238
Stock and Yogo Critical Val.	16.38	16.38	19.93
Hansen J-Stat. (p-value)	NA	NA	0.8878
Endogeneity Test (p-value)	0.0007	0.0121	0.0003
Observations	181,585	164,484	164,092
RMSE	0.3571	0.2819	0.2843
Sample	Entire Sample	Entire Sample	Entire Sample

Primary sampling unit clustered robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Additional controls include:

1) Cluster level controls: distance from the coast, soil fertility, annual precipitation, annual temperature, net primary productivity, mean GDP 2000, population density 2000, slave density, water availability, altitude and forest land cover.
2) Individual level controls: age, gender, individual income, education, occupation and religion (not included in Model 1).
3) Household level controls: gender of the household head, relationship to the head, relationship structure and household wealth (not included in Model 1).
4) District fixed effects.

To gain further reassurance about the potential exogeneity of our instrument, next we introduce a second instrument in order to get a test for over-identification which may shed further light (even though the rejection of the null denotes instrument endogeneity for sure, while the non-rejection does not completely exclude it). Rugged terrain is the second instrument which we use in conjunction with the proxy for the demographic shock. Goody (1977) argues that polygyny is normally associated with hoe agriculture and according to Alesina, Giuliano and Nunn (2013) the diffusion of plough agriculture is associated to soil characteristics (i.e., terrain slope, depth and so on). Therefore one should expect that rugged terrain affects the diffusion of hoe/plough agriculture and therefore polygyny. In Model 3 rugged terrain is also found to have a significant effect on polygyny. The Hansen test does not reject the null of exact identification and the coefficient on polygyny remains relatively stable (the slightly larger size if compared to Model 2 can be attributed to the smaller number of observations), which according to Wooldridge (2002) is an additional signal of instrument exogeneity. Moreover, because of the increased efficiency of Model 3 (the standard error is much smaller), the effect of polygyny on HIV results significant at a 5 percent level.

Finally in Table 7 we replicate 2SLS separately for men and women. In Model 1 and Model 3 we confine estimates to men only and we find that the effect of polygyny on HIV among men is marginally significant at best. In Model 2 and Model 4 we confine estimates to women only and the effect is marginally significant in Model 2 and significant at a 5 percent level in Model 4 (where we include household and individual controls). In addition, the marginal effect of polygyny on HIV among women is almost 20 percent higher than the effect among men.

To conclude, we can interpret our results as follows. Even after we instrument our indicators for polygyny using the demographic shock following the slave trade, we find evidence of a significant and positive effect of polygyny on the probability of HIV infection. We can therefore interpret the demographic shock as a "primordial" risk factor for HIV. Moreover, its effect is much larger and more precise among women. This is likely to happen since, when multiple co-wives live together in the same household, an infected co-wife can represent the source of transmission to all the other, multiplying the incidence of HIV among women.

Table 7: HIV Infection and Polygyny by Gender - IV

	Dependent Variable: HIV (1 if positive)			
Estimation Method: 2SLS	***Model 1***	***Model 2***	***Model 3***	***Model 4***
Polygyny	0.4348	0.5796*	0.4180*	0.6033**
	(0.3212)	(0.3098)	(0.226)	(0.249)
	First Stage Statistics			
Population Growth 1500-1860	-0.0396**	-0.0468***	-0.0394**	-0.0461**
	(0.0182)	(0.0179)	(0.0182)	(0.0179)
Rugged Terrain			-0.0058**	-0.0063**
			(0.0024)	(0.030)
Anderson LR-Stat.	15.274	20.267	25.367	31.350
Cragg Donald F-Stat.	15.266	20.258	12.677	15.669
Stock and Yogo Critical Val.	16.38	16.38	19.93	19.93
Hansen J-Stat. (p-value)	NA	NA	0.9908	0.8582
Endogeneity Test (p-value)	0.118	0.0131	0.0295	0.0005
Observations	71,219	93,265	70,983	93,109
RMSE	0.2383	0.3111	0.2357	0.3168
Sample	Men	Women	Men	Women

Primary sampling unit clustered robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Additional controls include:

1) Cluster level controls: distance from the coast, soil fertility, annual precipitation, annual temperature, net primary productivity, mean GDP 2000, population density 2000, slave density, water availability, altitude and forest land cover.
2) Individual level controls: age, individual income, education, occupation and religion (not included in Models 1 and 2).
3) Household level controls: gender of the household head, relationship to the head, relationship structure and household wealth (not included in Models 1 and 2).
4) District fixed effects.

5. HIV, Sexual Behavior, and the Demographic Shock

So far we have shown that there exists a link between the demographic shock, polygyny, and ultimately on HIV infection, and that the causal effect of polygyny on HIV is particularly strong among women. In this section we try to evaluate possible channels which can explain such a relationship. Since sexual behavior represents one of the main channels of transmission of HIV, we therefore focus on the role played by sexual behavior, as shaped by polygyny and in turn by the demographic shock. The general idea is that mating habits and marriage systems shape sexual behavior, so that sexual relationships in polygamous households may differ from those in monogamous households, thus leading to a differential effect on HIV.

Since the relationship between sexual behavior and HIV is clearly subject to double causation, in order to verify whether the demographic shock can be a suitable instrument for sexual behavior in Table 8 we start by looking at their relationship using OLS. To evaluate behavioral differences between men and women we use the DHS male and female datasets separately. We employ two alternative measures of sexual behavior: the intensity of sexual activity (coded as a dummy variable which is positive if the individual has had any sexual activity within the last 4 weeks) and the frequency of extramarital partnerships (proxied by the standardized number of sexual partners, spouse excluded, within the last 12 months).

In Model 1 and Model 2 of Table 8 we show estimates for the effect of our proxy for the demographic shock on male sexual activity and number of sexual partners other than the wife/wives. We find a statistically significant and positive relationship between the demographic shock and sexual activity. In other words, men in regions which have been more heavily exposed to the slave trade tend to be sexually more active today. However, despite the more active sexual activity, men do not show a higher propensity to extramarital partnerships (Model 2). In fact the effect of the demographic shock on extramarital partnerships is not significant.

In Model 3 and Model 4 we turn to the female dataset. Given that individuals in areas exposed to a more severe shock tend to have a higher probability to be in a polygynous relationship, one should expect a negative effect of the demographic shock on sexual activity among women, as a consequence of their placement in a polygynous household. However in Model 3 we do not find any significant statistical relationship between the demographic shock and sexual activity, which suggests that women in areas with a higher probability of polygyny do not have significantly lower sexual activity than women in monogamous families. However, in Model 4 we find a significant relationship between the demographic shock and the number of sexual partners other than the husband. Therefore women in clusters more heavily exposed to the demographic shock during the slave trade are more likely to be in a polygynous relationship, but this does not decrease their sexual activity, which can in part be explained by the fact that in those regions women have a higher probability of extramarital partnerships. The latter sort of behavior can be attributed to the fact that neglected co-wives tend to look for extramarital partners (Poku, 2005; Ogundipe-Leslie, 1994). This result also confirms existing evidence from surveys in Tanzania, Lusaka, and Kenya. For example, Kwena et al. (2014) conduct a survey in Kisumu County (Kenya) and find that lower sexual satisfaction within the marriage is associated with increased likelihood of extramarital partnerships. Using WHO/GPA surveys Caraël (1995) finds that in Lusaka (Zambia's capital) and Tanzania the share of women in a polygynous relationship who report more than one regular partners is around 10 percent. In Lesotho, where polygyny is prevalent, the percentage of women with more than one regular partner is almost as high as 39 percent. These preliminary results do highlight the presence of differential patterns across genders.[15]

Table 8: Sexual Behavior and the Demographic Shock - OLS

Estimation Method: OLS	***Model 1***	***Model 2***	***Model 3***	***Model 4***
Dependent Variables:	***Sexual Activity***	***Extramarital Partners***	***Sexual Activity***	***Extramarital Partners***
Population Growth 1500-1860	-0.0258**	0.0334	0.0004	-0.0349**
	(0.0116)	(0.0359)	(0.0094)	(0.0139)
Sample	Men	Men	Women	Women
Observations	121,092	120,041	279,778	275,491
RMSE	0.4064	1.044	0.4261	0.7395

Primary sampling unit clustered robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.
Additional controls include:
1) Cluster level controls: distance from the coast, soil fertility, annual precipitation, annual temperature, net primary productivity, mean GDP 2000, population density 2000, slave density, water availability, altitude and forest land cover.
2) Individual level controls: age, gender of the household head, cohabitation duration, ideal number of children, education and religion.
3) District fixed effects.

Since OLS estimates are likely to be affected by omitted variable bias, to validate the idea that sexual behavior represents the channel through which the demographic shock influences HIV, we replicate our analysis using an IV estimator, in order to exclude potential confounders what may affect the estimates. The instrument we select is population in 1500, since slave raids were more frequent in regions with a higher population density (Eltis and Richardson, 1997), which in turn implies that the demographic shock is likely to be more

[15] Over the full sample merging the household, male, and female datasets, we obtain a negative and (at one percent) significant coefficient for sexual activity and a negative but insignificant coefficient for extramarital partnerships. We do not report these estimates for brevity.

severe in regions with a larger population in 1500. Therefore, rather than a positive relationship between a measure of pre-industrial wealth and population growth, as it is likely to occur in an Malthusian state, we should expect a negative relationship between population in 1500 and subsequent population growth. In Table 9, Models 1 and 2, we focus on the male dataset. For both models diagnostic tests and first-stage statistics confirm the relevance of the instrument. The coefficient in Model 1 is larger than the coefficient on the corresponding model estimated by OLS. On average, a one percent change in population growth over the slave trade period now increases the probability of having sexual activity by almost 32 percent. The insignificant effect of the shock on extramarital partnerships among men in Model 2 is also confirmed.[16] To sum up, for men in regions more heavily exposed to the demographic shock – where polygyny is more frequently practiced – the demographic shock also affects sexual activity. Moreover, the higher intensity of sexual activity is not imputable to extramarital partnerships. However, the latter outcome may be related to the fact that when men want a new partner rather than cheating they have the option to marry her, given that in these regions polygyny is possible.

In Models 3 and 4 of Table 9 we turn to the female dataset. IV estimates again largely confirm OLS results. Despite polygyny, women in areas affected by the demographic shock do not seem to have a less intense sexual activity, with the coefficient being significant at a 10 percent level (i.e., sexual activity increases by almost 14 percent per a one percentage change in population growth, but the effect is marginally significant). In addition the significant effect on the number of partners (not including the spouse) in the last 12 months is still significant.

[16] To derive statistics for the Hansen test, in unreported regressions we also use as a second instrument the share of land covered by forest, since archaeological evidence (Stilwell, 2014) reports that sedentary tribes in dense forests used to rely on slave labor even before 1500, which implies that slave raids were more likely to occur in areas with a larger forest land cover. The Hansen test does not reject the null of exact identification and the coefficients remain relatively stable.

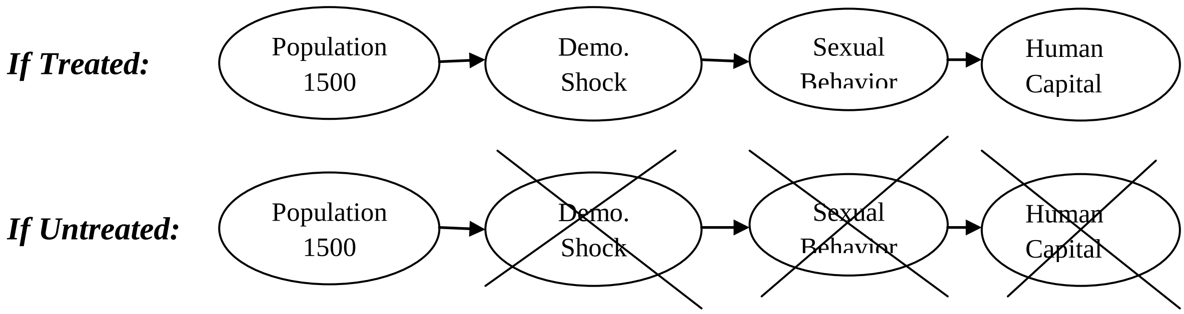
Table 9: Sexual Behavior and the Demographic Shock - IV

Estimation Method: 2SLS	***Model 1***	***Model 2***	***Model 3***	***Model 4***
Dependent Variable	***Sexual Activity***	***Extramarital Partners***	***Sexual Activity***	***Extramarital Partners***
Population Growth 1500-1860	-0.3156**	0.2837	-0.1395*	-0.2796**
	(0.1583)	(0.371)	(0.0725)	(0.1422)
	First Stage Statistics			
Population in 1500	-0.0258***	-0.0232***	-0.0343***	-0.0293***
	(0.00597)	(0.00577)	(0.0054)	(0.0053)
Anderson LR-Stat.	1116.595	902.575	5474.893	4008.684
Cragg Donald F-Stat.	1121.000	905.358	5528.062	4037.434
Stock and Yogo Critical Val.	16.38	16.38	16.38	16.38
Endogeneity Test (p-value)	0.0347	0.4685	0.0353	0.0613
Observations	121,092	120,041	279,778	275,491
RMSE	0.4077	1.043	0.4262	0.7397
Sample	Men	Men	Women	Women

Primary sampling unit clustered robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.
Additional controls include:
1) Cluster level controls: distance from the coast, soil fertility, annual precipitation, annual temperature, net primary productivity, mean GDP 2000, population density 2000, slave density, water availability, altitude and forest land cover.
2) Individual level controls: age, gender of the household head, cohabitation duration, ideal number of children, education and religion.
3) District fixed effects.

To further check the validity of the instrument used in Table 9, in Table 10 we run placebo regressions in which we test the effect of population in 1500 on current measures of GDP and population density. Figure 4 represents graphically the idea of our falsification test.

Figure 4: Falsification Test



In order to be a valid instrument, population in 1500 should exert an effect on sexual behavior only through the demographic shock, while there should be no effect on sexual behavior through, for example, development or population pressure. So we can split the sample between clusters that were exposed to the transatlantic slave trade (i.e., West Africa countries) and clusters that were not (i.e., East Africa countries[17]). If a country is exposed to the transatlantic slave trade we should observe an effect of population in 1500 on the demographic shock and therefore on sexual behavior, mortality (where the latter effect runs through HIV, according to our previous results), human capital accumulation, and ultimately development. On the other hand there should be no relationship between population in 1500, the demographic shock, and development if the country is not heavily exposed to the trade given that the relationship between these variables should be random.

Results in Table 10 show that indeed the effect of population in 1500 on GDP and population density in 2000 (i.e., two proxies of development) is significant only for the treated area. This suggests that the effect of the demographic shock on GDP and population density may be present only for the areas exposed to the slave trade, consistently with our previous results on HIV (i.e., HIV increases mortality and decreases human capital accumulation). On the other hand the relationship is not significant for areas not exposed to the slave trade given that the relationship should be random as highlighted in Figure 4. We repeat the same exercise with distance from the coast in place of population in 1500 and we find that distance from the coast has a significant effect on population density and mean GDP in 2000 for both samples (consistent with Sachs and Werner, 1997) which from our perspective dismisses the

[17] Madagascar is not included in this sample because quite heavily exposed to the transatlantic slave trade.

variable as a potential instrument. The endogeneity of distance from the coast is also confirmed by over-identification tests.

Table 10: Placebo Regressions for the Exogeneity of Population in 1500

Dependent Variables:	***Log (1+Mean GDP 2000)***		***Log (1+ Pop Density 2000)***	
	Model 1	***Model 2***	***Model 3***	***Model 4***
Estimation Method: OLS	***West Africa***	***East Africa***	***West Africa***	***East Africa***
Population in 1500	-0.258***	0.0278	0.547***	0.0530
	(0.0528)	(0.0274)	(0.0860)	(0.0688)
Observations	4,175	4,032	4,175	4,032
RMSE	0.521	0.459	1.428	1.442

Country clustered robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.
Additional controls include:
1) Cluster level controls: distance from the coast, soil fertility, annual precipitation, annual temperature, net primary productivity, slave density, water availability, altitude and forest land cover.
2) District fixed effects.

In Table 11, to investigate how sexual behavior affects HIV infection, we preliminarily run OLS regressions, despite the fact that we clearly face a serious endogeneity problem due to reverse causality. More sexual activity is associated with more infection but on the other hand infected people tend to reduce sexual activity. The latter link is what explains the counter-intuitive negative association between sexual activity and HIV emerging in Models 1 and 3 both for men and women. Similar considerations may apply to the results for extramarital affairs.

Table 11: HIV Infection and Sexual Behavior - OLS

	Dependent Variable: HIV (1 if positive)			
Estimation Method: OLS	***Model 1***	***Model 2***	***Model 3***	***Model 4***
Sexual Activity	-0.00439**		-0.00517**	
	(0.00210)		(0.00242)	
Extramarital Partners		0.0010		0.0040***
		(0.0008)		(0.0015)
Observations	92,708	93,081	121,554	121,927
RMSE	0.200	0.2001	0.2580	0.2582
Sample	Men	Men	Women	Women

Primary sampling unit clustered robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Additional controls include:

1) Cluster level controls: distance from the coast, soil fertility, annual precipitation, annual temperature, net primary productivity, mean GDP 2000, population density 2000, slave density, water availability, altitude and forest land cover.

2) Individual level controls: age, individual income, education, occupation and religion.

3) Household level controls: gender of the household head, relationship to the head, relationship structure and household wealth.

4) District fixed effects.

Next we run analogous IV regressions using the demographic shock as an instrument, even though we can do so only for models for which the instrument is sufficiently strong. As shown by Tables 8 and 9, we can rely on a sufficiently significant first-stage relationship only for sexual activity over the male sample and for extramarital partners over the female sample. We show the resulting estimates in Table 12. In Models 1 and 3, in order to maximize the estimated sample size, we do not include household and individual controls. First-stage statistics confirm the effect of the proxy for the demographic shock on both measures of sexual behavior. The second stages show that, for males, a more intense sexual activity has a significant and positive effect on HIV, which increases its probability by almost 32 percent (Model 1). For females, Model 3 also shows a significant and positive effect of extramarital partnerships on HIV. A standard deviation in the standardized number of partners (roughly one more partner) increases the probability of HIV by almost 29 percent. In Models 2 and 4 we replicate the same regressions including individual controls. Under these specifications, however, the coefficient for male sexual activity is larger but loses significance, while the coefficient for female infidelity displays a p-value of 0.051.

Table 12: HIV Infection and Sexual Behavior - IV

	Dependent Variable: HIV (1 if positive)			
Estimation Method: 2SLS	Model 1	Model 2	Model 3	Model 4
Sexual Activity	0.316*	0.370		
	(0.172)	(0.277)		
Extramarital Partners			0.285**	0.240*
			(0.133)	(0.123)
	First Stage Statistics			
Population Growth 1500-1860	-0.0586***	-0.0367***	-0.0922***	-0.0914***
	(0.01819)	(0.0135)	(0.0288)	(0.0285)
Anderson LR-Stat.	18.920	11.419	22.560	21.978
Cragg Donald F-Stat.	18.920	11.414	22.560	21.970
Stock and Yogo Critical Val.	16.38	16.38	16.38	16.38
Endogeneity Test (p-value)	0.0754	0.1334	0.0078	0.020
Observations	92,810	92,708	123,283	121,927
RMSE	0.2471	0.2489	0.3475	0.3206
Sample	Men	Men	Women	Women

Primary sampling unit clustered robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Additional controls include:

1) Cluster level controls: distance from the coast, soil fertility, annual precipitation, annual temperature, net primary productivity, mean GDP 2000, population density 2000, slave density, water availability, altitude and forest land cover.
2) Individual level controls: age, individual income, education, occupation and religion (not included in Models 1 and 3).
3) Household level controls: gender of the household head, relationship to the head, relationship structure and household wealth (not included in Models 1 and 3).
4) Country and district fixed effects.

To conclude, we can interpret our findings as follows. Confirming the results from Section 4, the demographic shock again emerges as a "primordial" risk factor for HIV, since it still affects sexual behavior. This influence is particularly strong for women and for their attitude toward marital infidelity. Through this channel, the effect of polygyny uncovered in Section 4 is magnified because of the co-residence patterns involving multiple co-wives.

6. Conclusion

The papers has shown how the demographic shock affecting sub-Saharan Africa as a consequence of the slave trade is associated with contemporaneous outcomes, namely, the prevalence of polygyny, sexual behavior, and HIV infection rates. In clusters which were more heavily affected by the shock, we observe a larger fraction of polygynous families, more active male sexual activity, a larger probability of female extramarital partnerships, and a higher female HIV infection rate. Our findings can be interpreted as the outcome of unsatisfying marital relationships for women within polygynous families, which encourage women's infidelity, thus increasing their exposure to the risk of infection, with multiplying effects through co-wives cohabitation.

Policy reports by international organizations (UNAIDS, 2013) stress the disproportionate impact of HIV on women, especially in sub-Saharan Africa, and search for tools to allow them to play an active role in the struggle against AIDS. Harmful gender norms, inequalities between the sexes, and women's rights violations are among the reasons why women and girls are more vulnerable to the virus. Women's empowerment, starting from access to comprehensive sexuality education and knowledge, has been indicated as a crucial intermediate goal. Since sexual habits involving multiple partners are a central driver of the spread of the epidemic, it is crucial to understand their determinants, including the determinants of extramarital partnerships. Public opinion as well as policymakers are well aware of the potential link between the latter and the practice of polygyny. A recent marriage bill legalizing polygamy in Kenya has ignited controversies as HIV activists claim it may increase the spread of the disease (Maroncha, 2014). Polygamy is commonly accepted in Kenya according to customary law, but it was previously outlawed. Both the unsatisfying nature of polygynous marriages, leading to female infidelity, and cohabitation among co-wives, have been explicitly recognized as drivers of HIV spread within the debate that followed the enactment of the bill. To document and quantify the strength of the links between HIV incidence on the one hand and sexual behavior and polygyny on the other, and to point to the demographic disruption brought about by the slave trade as a "primordial" risk factor, is the main contribution of this paper.

References

Acemoglu, D., Johnson, S. and Robinson, J.A., 2002. Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution. Quarterly Journal of Economics 117, 1231-1294.

Alesina, A. and Giuliano, P., 2013. Culture and Institutions. NBER Working Papers No. 19750.

Ashby, B. and Gupta S., 2013. Sexually Transmitted Infections in Polygamous Mating Systems. Philosophical Transactions of the Royal Society B: Biological Sciences 368, 20120048.

Ashraf, N., Bandiera, O. and Jack, K., 2014. No Margin, No Mission? A Field Experiment on Incentives for Public Service Delivery. Journal of Public Economics, in press.

Auvert, B., Taljaard, D., Lagarde, E., Sobngwi-Tembekou, J., Sitta, R. and Puren, A., 2005. Randomized, Controlled Intervention Trial of Male Circumcision for Reduction of HIV Infection Risk: The ANRS 1265 Trial. PLoS Med 2, 1112-1122.

Becker, G.S., 1974. A Theory of Marriage: Part ii. Journal of Political Economy 82, S11–S26, Part II.

Bisin, A. and Verdier, T., 2001. The Economics of Cultural Transmission and the Dynamics of Preferences. Journal of Economic Theory 97, 298-319.

Boserup, E., 1970. Women's Role in Economic Development. George Allen and Unwin, London.

Brasio, A., 1958. Monumenta Missionaria Africana. Coligida e Anotada pelo Padre Antonio Brasio. Agência Geral do Ultramar, Lisbon.

Caldwell, J.C. and Caldwell, P., 1987. The Cultural Context of High Fertility in Sub-Saharan Africa. Population and Development Review 13, 409-437.

Caldwell, J.C., Orubuloye, I.O. and Caldwell, P., 1991. The Destabilization of the Traditional Yoruba Sexual System. Population and Development Review 17, 229-262.

Caraël, M., 1995. Sexual Behaviour. In John Cleland, J.G. and Ferry, B. (eds.), Sexual Behaviour and AIDS in the Developing World. Francis and Taylor, Oxon, 75-123.

Ciment, J., 2007. Atlas of African-American History, Revised Edition. Facts On File, New York.

Dalton, J.T. and Leung, T.C., 2014. Why Is Polygyny More Prevalent in Western Africa? An African Slave Trade Perspective. Economic Development and Cultural Change 62, 599-632.

Denevan, W., 1992. The Native Population of the Americas in 1492 (2nd ed.). The University of Wisconsin Press, Madison.

De Walque, D., 2007. How Does the Impact of an HIV/AIDS Information Campaign Vary with Educational Attainment? Evidence from Rural Uganda. Journal of Development Economics 84, 686-714.

Duflo, E., Dupas, P. and Kremer, M., 2015. Education, HIV, and Early Fertility: Experimental Evidence from Kenya. CEPR Discussion Paper No. 10338.

Dupas, P., 2011. Do Teenagers Respond to HIV Risk Information? Evidence from a Field Experiment in Kenya. American Economic Journal: Applied Economics 3, 1-34.

Edlund, E. and Ku, H., 2011. The African Slave Trade and the Curious Case of General Polygyny. MPRA Paper No. 52735.

Eltis, D., 2008. The U.S. Transatlantic Slave Trade, 1644-1867: An Assessment. Civil War History 54, 347-378.

Eltis, D. and Richardson, D. (eds.), 1997. Routes to Slavery: Direction, Ethnicity and Mortality in the Transatlantic Slave Trade. Routledge, Oxon.

Fage, J.D., 1980. Slaves and Society in Western Africa, c, 1445 - c. 1700. Journal of African History 21, 289-310.

Fenske, J., 2013. African Polygamy: Past and Present. Mimeo, University of Oxford.

Francis, A.M., 2008. The Economics of Sexuality: The Effect of HIV/AIDS on Homosexual Behavior in the United States. Journal of Health Economics 27, 675-689.

Frankema, E. and Jerven, M., 2014. Writing History Backwards or Sideways: Towards a Consensus on African Population, 1850-present. Economic History Review 67, 907-931.

Galor, O. and Ashraf, Q., 2013. The Out of Africa Hypothesis, Human Genetic Diversity and Comparative Development. American Economic Review 103, 1-46.

Gennaioli, N. and Rainer, I., 2007. The Modern Impact of Precolonial Centralization in Africa. Journal of Economic Growth 12, 185-234.

Giavazzi, F., Petkov, I. and Schiantarelli, F., 2014. Culture: Persistence and Evolution. NBER Working Paper No. 20174.

Goody, J., 1973. The Character of Kinship. Cambridge University Press, Cambridge.

Gould, E.D., Moav, O. and Simhon, A., 2008. The Mystery of Monogamy. American Economic Review 98, 333-357.

Ghosh, T., Powell, R.L., Elvidge, C.D., Baugh, K.E., Sutton, P.C. and Anderson, S., 2010. Shedding Light on the Global Distribution of Economic Activity. Open Geography Journal, 3, 147-160.

Gray, R.H., Kigozi, G., Servadda, D., Makumbi, F., Watya, S., Nalugoda, F. and Kiwanuka, N., 2007. Male Circumcision for HIV Prevention in Men in Rakai, Uganda: A Randomized Trial. Lancet 369, 657-666.

Greenwood, J., Kircher, P., Santos, C. and Tertilt, M., 2013. An Equilibrium Model of the African HIV/AIDS Epidemic. NBER Working Paper No. 18953.

Hall, R.E. and Jones, C.I., 1999. Why Do Some Countries Produce So Much More Output Per Worker Than Others? Quarterly Journal of Economics 114, 83-116.

Halperin, D.T. and Epstein, H., 2004. Concurrent Sexual Partnerships Help to Explain Africa's High HIV Prevalence: Implications for Prevention. Lancet 364, 4-6.

ICF International, 2012. Survey Organization Manual for Demographic and Health Surveys. MEASURE DHS. ICF International, Calverton.

Jacoby, H.G., 1995. The Economics of Polygyny in Sub-Saharan Africa: Female Productivity and the Demand for Wives in Côte d'Ivoire. Journal of Political Economy 103, 938-971.

Klein Goldewijk, K., 2005. Three Centuries of Global Population Growth: A Spatial Referenced Population Density Database for 1700 – 2000. Population and Environment 26, 343-367.

Klein Goldewijk, K. , Beusen, A. and Janssen, P., 2010. Long Term Dynamic Modeling of Global Population and Built-up Area in a Spatially Explicit Way, HYDE 3 .1. Holocene 20, 565-573.

Kwena, Z., Mwanzo, I., Shisanya, C., Camlin, C., Turan, J., Achiro, L. and Bukusi, E., 2014. Predictors of Extra-Marital Partnerships among Women Married to Fishermen along Lake Victoria in Kisumu County, Kenya. PLoS ONE 9, e95298.

Labat, J.-B., 1728. Nouvelle Relation de l'Afrique Occidentale. Guillaume Cavelier, Paris.

Lagerlof, N.-P., 2010. Pacifying Monogamy. Journal of Economic Growth 15, 235-262.

Lahmeyer, 2004. Populstat Database. Growth of the Population per Country in a Historical Perspective, Including their Administrative Divisions and Principal Towns. www.populstat.info/.

Lamie, K., 2007. The Impact of the Atlantic Slave Trade on West Africa: Polygyny and Female Reproductive Success. Nebraska Anthropologist 22, 16-32.

Lee, G.R. and Whitbeck, L.B. 1990. Economic Systems and Rates of Polygyny. Journal of Comparative Family Studies 21, 13-24.

Livi Bacci, M., 2007. A Concise History of World Population (4th ed.). Blackwell, Malden.

Luke, N. and Kurtz, K.M., 2002. Cross-Generational and Transactional Sexual Relations in Sub-Saharan Africa: Prevalence of Behavior and Implications for Negotiating Safer Sexual Practices, ICRW/PSI Report.

Maddison, A., 2003. The World Economy: Historical Statistics. OECD Development Centre, Paris.

Magruder, J., 2011. Marital Shopping and Epidemic AIDS. Demography 48, 1401-1428.

Manning, P., 1981. The Enslavement of Africans: A Demographic Model. Canadian Journal of African Studies 15, 499-526.

Maroncha, L., 2014. Polygamy in Kenya: A Legal Way to Spread HIV? Key Correspondents: Reporting for Action on HIV, May 1.

McEvedy, C. and Richard, J., 1978. Atlas of World Population History. Penguin, Middlesex.

Mbirimtengerenji, N.D., 2007. Is HIV/AIDS Epidemic Outcome of Poverty in Sub-Saharan Africa? Croatian Medical Journal 48, 605-661.

Michalopoulos, S. and Papaioannou, E., 2013. Pre-Colonial Ethnic Institutions and Contemporary African Development. Econometrica 81, 113-152.

Mishra, V., Vaessen, M., Ties Boerma, J.T., Arnold, F., Way, A., Barrere, B., Cross, A., Hong, R. and Sangha, J., 2006. HIV Testing in National Population-based Surveys: Experience from the Demographic and Health Surveys. Bulletin of the World Health Organization 84, 537-545.

Mitsunaga, T.M., Powell, A.M., Heard, N.J. and Larsen, U.M., 2005. Extramarital Sex Among Nigerian Men: Polygyny and Other Risk Factors. Journal of Acquired Immune Deficiency Syndromes 39, 478-488.

Murdock, G.P., 1967. Ethnographic Atlas. University of Pittsburgh Press, Pittsburgh, PA.

Nunn, N., 2012. Culture and the Historical Process. Economic History of Developing Regions 27, 108-126.

Nunn, N., 2008. The Long Term Effects of Africa's Slave Trades. Quarterly Journal of Economics 123, 139-176.

Nunn, N. and Wantchekon, L., 2011. The Slave Trade and the Origins of Mistrust in Africa. American Economic Review 101, 3221-3252.

Ogundipe-Leslie, M., 1994. Re-creating Ourselves: African Women & Critical Transformations. African World Press, Trenton.

Oster, E., 2005. Sexually Transmitted Infections, Sexual Behavior, and the HIV/AIDS Epidemic. Quarterly Journal of Economics 120, 467-515.

Poku, N.K., 2005. Aids in Africa: How the Poor Are Dying. Polity Press, Cambridge.

Rehle, T., Shisana, O., Pillay, V., Zuma, K., Puren, A. and Parker, W., 2007. National HIV Incidence Measures: New Insights into the South African Epidemic. South African Medical Journal 97, 194-199.

Reniers, G. and Watkins, S., 2010. Polygyny and the Spread of HIV in Sub Saharan Africa: A Case of Benign Concurrency. AIDS 24, 299–307.

Sachs, J. D. and Warner, A. M., 1997. Fundamental Sources of Long Run Growth. American Economic Review 87, 184-188.

Stilwell, S., 2014. Slavery and Slaving in African History. Cambridge University Press, New York.

Stock, J.H. and Yogo, M., 2005. Testing for Weak Instruments in Linear IV Regression. In Andrews, D.W.K. and Stock, J.H. (eds.), Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg. Cambridge University Press, Cambridge, 80-108.

Strassmann, B., 2000. Polygyny, Family Structure, and Child Mortality: A Prospective Study among the Dogon of Mali. In Cronk, L., Chagnon, N. and Irons, W. (eds.), Adaptation and Human Behavior. Aldine de Gruyer, New York, 49-67.

Teso, E., 2014. The Long-Term Effect of Demographic Shocks on the Evolution of Gender Roles: Evidence from the Trans-Atlantic Slave Trade. Mimeo, Harvard University.

Texeira de Mota, A., 1974. As Viagens do Bispo D. Frei Vitorino Portuense à Guiné e a Cristianização dos Reis de Bissau. Junta de Investigações Científicas do Ultramar, Lisbon.

Tertilt, M., 2005. Polygyny, Fertility, and Savings. Journal of Political Economy 113, 1341-1371.

Thornton, J., 1983. Sexual Demography: The Impact of the Slave Trade on Family Structure. in Robinson, C. and Klein, M. (eds.),Women and Slavery in Africa, University of Wisconsin Press, Madison, 39-48.

Thornton, J., 1980. The Slave Trade in Eighteenth Century Angola: Effects on Demographic Structures. Canadian Journal of African Studies 14, 417-428.

Timæus, I.M. and Reynar, A., 1998. Polygynists and their Wives in Sub-Saharan Africa: An Analysis of Five Demographic and Health Surveys. Population Studies 52, 145-162.

Todd, E., 1984. L'Enfance du Monde, Structures Familiales et Développement. Seuil, Paris.

UNAIDS, 2013. Global Report: UNAIDS Report on the Global AIDS Epidemic. UNAIDS, Geneva.

Wang, S.K., 2010. Violence & HIV/AIDS: Violence against Women as a Cause and Consequence of HIV/AIDS. Duke Journal of Gender Law & Policy 17, 313-332.

White, D.R. and Burton, M.L., 1988. Causes of Polygyny: Ecology, Economy, Kinship, and Warfare. American Anthropologist 90, 871-887.

WHO, 2012. World Health Statistics 2012. World Health Organization, France.

Young, A., 2007. In Sorrow to Bring Forth Children: Fertility amidst the Plague of HIV. Journal of Economic Growth 12, 283-327.

Young, A., 2005. The Gift of the Dying: The Tragedy of AIDS and the Welfare of Future African Generations. Quarterly Journal of Economics 120, 423-466.

Appendix

Table A1: DHS Countries and Waves

FEMALE SAMPLE	***MALE SAMPLE***	***HIV SAMPLE***
BF6	BF6	BF6
BJ6	BJ6	**
BU6	BU6	BU6
CD5	CD5	CD5
CI6	CI6	CI6
CM6	CM6	CM6
ET6	ET6	ET6
GA6	GA6	GA6
GH4	GH4	GH4
GN4	GN4	GN4
KE5	KE5	KE5
LB5	LB5	LB5
LS5	LS5	LS5
MD5	MD5	**
ML5	ML5	ML5
MW5	MW5	MW5
MZ5	*	MZ5
NG5	NG5	**
RW6	RW6	RW6
SL5	SL5	SL5
SN4	SN4	SN4
SZ5	SZ5	SZ5
TZ5	*	TZ5
UG6	*	UG6
ZM5	ZM5	ZM5
ZW6	ZW6	ZW6

* The male sample is not part of the AIS dataset which for these countries is the only available with GPS coordinates.

** No HIV sample are available for these countries.

Table A2. Variable Description and Sources

Variable	***Description***	***Source***
HIV, Individual, and Household Variables	Men, Women, and Household Member Datasets	Demographic and Health Survey (DHS) HIV and DHS samples
Clusters	Voronoi Transformation of DHS GPS Coordinates	DHS GPS Dataset
Population Growth 1500-1860	Log Population 1860 – Log Population 1500	The History Database of the Global Environment (HYDE)
Altitude	FAO GAEZ Data on Median Altitude	FAO GAEZ (Terrain Resources)
Water Availability	Share of Inland Water Bodies	FAO GAEZ (Water Resources)
Forest Land Cover	Share of Forest Land	FAO GAEZ (Land Cover)
Soil Fertility	Natural Soil Fertility	FAO GAEZ (Soil Resources)
Precipitation	Total Annual Precipitation (1960-2000)	FAO GAEZ (Agro Climatic Resource – Moisture Regimes)
Temperature	Annual Temperature (1960-2000)	FAO GAEZ (Agro Climatic Resources – Thermal Regimes)
Net Primary Productivity	Index of Vegetation	FAO GAEZ (Agro Climatic Resources – Growing Period)
Soil Ruggedness	Computed Using DEM Terrain Analysis in QGIS	Harmonized World Soil Database (International Institute for Applied Systems Analysis - IIASA)
Population Density	Spatial Data on Population Density	Gridded Population of the World – SEDAC/CIESIN
Distance from the Coast	Mean Distance from the Closest Coast	NASA Ocean Biology Processing Group
Mean GDP	Spatial Data on Total Economic Activity	National Geophysical Data Center (Ghosh et al., 2010)
Transatlantic Slave Trades	Total Number of Slaves per Ethnic Group/Tribe Land	Nunn and Wantchekon (2011)

Table A3: Summary Statistics

HIV and Household Dataset

Variable	Obs	Mean	Std. Dev.	Min	Max
HIV	263517	.0633773	.2436408	0	1
Sexual Activity (1 if had sex in the last 4 wks)	240510	.5033637	.4999897	0	1
Polygyny	203927	.2045095	.4033437	0	1
Extramarital Partners (standardized - 1sd)	241519	.0022969	-0.103079	15.76942	1
Gender	235698	.5938659	.4911112	0	1
Distance from the Coast (standardized - 1 sd)	269903	-.0552175	.9634542	-.8523329	3.533765
Soil Fertility - Ph (standardized - 1 sd)	270457	-.0303074	.9990131	-1.340364	4.595398
Annual Precipitation (standardized - 1 sd)	270457	-.0137997	1.008733	-1.376884	3.375861
Annual Temperature (standardize d - 1 sd)	270457	-.0476981	1.005373	-4.417325	1.864606
Net Primary Productivity(standardized - 1 sd)	270457	-.0075186	.9438511	-5.672748	1.399589
Mean GDP 2000 (Log 1+X)	270457	.4579312	.9416462	0	5.087565
Population Density 2000 (Log 1+X)	270457	3.471484	1.851666	0	9.426983
Slaves/Tribe Area (Log 1+X)	270457	.1890907	.4842291	0	3.655909
Water Availability (standardized - 1sd)	270457	-.008009	.9842829	-.2067334	9.538667
Altitude (standardized - 1 sd)	270408	.0081215	1.019178	-1.071481	4.352947
Forest Land Cover (log 1 + X)	270457	2.809252	1.144777	0	4.566624
Population Growth 1500-1860	270457	.2272007	.2537318	-.469189	1.617584
Rugged Terrain (standardized - 1 sd)	269648	.0325239	1.011561	-.9551349	7.677637

Table A4: The Demographic Shock and the Transatlantic Slave Trades

	Dependent Variable: Population Growth 1500-1860		
	Model 1	**Model 2**	**Model 3**
Estimation Method:	***OLS***	***OLS***	***IV***
Transatlantic Slave Trades	-0.0786***	-0.0366***	-0.123***
	(0.0133)	(0.0130)	(0.0446)
			First Stage Statistics
Distance from the Coast			-0.1955***
			(0.0338)
Anderson Cannon LR-Stat.			487.268
Cragg Donald F-Stat.			497.268
Stock and Yogo Critical Val.			16.38
Endogeneity Test (p-value)			0.0467
Observations	11,649	11,617	11,586
RMSE	0.253	0.231	0.24.14

Clustered robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Additional controls in Models 2 and 3 include: soil fertility, annual precipitation, annual temperature, net primary productivity, water availability, altitude and rugged terrain.

Table A5: Polygyny and the Transatlantic Slave Trades - OLS

	Dependent Variable: Polygyny	
Estimation Method: OLS	***Model 1***	***Model 2***
Transatlantic Slave Trades	0.0158***	-0.0036
	(0.0053)	(0.0092)
Observations	71,264	71,264
RMSE	0.3183	0.3109
District Fixed Effects	No	Yes
Sample	Men	Men

Primary sampling unit clustered robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Additional controls include:

1) Cluster level controls: distance from the coast, soil fertility, annual precipitation, annual temperature, net primary productivity, mean GDP 2000, population density 2000, slave density, water availability, altitude and forest land cover.

2) Individual level controls: age, individual income, education, occupation and religion.

3) Household level controls: gender of the household head, relationship to the head, household income and religion.