

Women’s Position in Ancestral Societies and Female HIV: The Long-Term Effect of Matrilineality in Sub-Saharan Africa *

JOB MARKET PAPER

Jordan Loper[†]

August 28, 2019

[Most Recent Version Here](#)

Abstract

Can contemporaneous female HIV rates be traced back to women’s position in ancestral societies ? Using within-country variation across about 280,000 individuals in their ethnic group’s ancestral kinship organization, in 18 Sub-Saharan African countries, I find that females originating from ancestrally matrilineal ethnic groups are today more likely to be infected by HIV than their patrilineal counterparts. I show that this finding is robust to the inclusion of subnational fixed effects, as well as a large set of cultural, historical, geographical, and environmental factors that could be confounding the estimates. I also use a number of alternative estimation strategies, including and instrumental variable (IV) strategy as well as a geographic regression discontinuity design (RDD), and I find results consistent with my main findings. Contrary to patrilineal kinship organizations, in matrilineal kinship organizations males’ biological children are not their heirs, and they integrate the kin group of their mothers. Matrilineal males are therefore relatively less likely to engage in a long-term committed relationship to ensure the survival of their offsprings, and seeking for better genes through sexual variety may have been a preferred reproductive strategy for ancestral matrilineal females. A prediction that emerges from the evolutionary psychology literature is that, driven by the reproductive success of their ancestors, matrilineal females should exhibit a stronger preference for promiscuous sexual behaviours, relative to their patrilineal counterparts, which may explain their higher contemporaneous HIV rates. I provide evidences supporting this prediction: I find that matrilineal females adopt riskier sexual and contraceptive behaviours, and are mainly infected through extra-marital routes of infection. However, I also find that matrilineal individuals adopt more protective contraceptive behaviours when they have internalized the risk of transmission of the virus. I finally build an epidemiological model and conduct a simulation exercise to illustrate how these differences between matrilineal and patrilineal individuals’ sexual and contraceptive behaviours translate into different gender-specific HIV rates dynamics.

Keywords: Kinship systems, matrilineality, cultural persistence, HIV, sexual behaviour, gender

JEL Classification: D13, D91, I12, I15, J12, N37, Z13

*I thank Alberto Alesina, Isaiah Andrews, Duman Bahrami-Rad, Anke Becker, Nicolas Berman, Melissa Dell, Yannick Dupraz, Catherine Guirkingier, Joseph Henrich, Asim Khwaja, Jacob Moscona, Marc Sangnier, Heather Sarsons, Jonathan Schulz, Liyang Sun, Michèle Tertilt, Edoardo Teso, Romain Wacziarg and participants at various seminars and conferences for helpful comments and suggestions. I also especially thank Nathan Nunn and Roberta Ziparo for their warm support and guidance. This work was supported by the French National Research Agency Grant ANR-17-EURE-0020. Part of this paper was written while I was visiting Harvard University and I thank them for their hospitality.

[†]Aix-Marseille Univ., CNRS, EHESS, Centrale Marseille, IRD, AMSE, Marseille, France (email: jordan.loper@univ-amu.fr)

1 Introduction

Every four minutes three young women become infected with HIV in the world. In 2018, the number of women in the world living with HIV amounted to 18.8 millions (UNAIDS (2018b)). With 14.5 millions of women living with HIV¹, representing about 80% of the worldwide female HIV positive population, Sub-Saharan Africa is the most affected region in the world². Uniquely, it is the only place in the world where more women than men live with HIV: adult female HIV positive population was about 1.5 larger than male HIV positive population³, and adult females represented about 57% of new adult infections in 2017⁴. This phenomenon has been referred to “Feminization of HIV” in Sub-Saharan Africa. Widely-acknowledged, this phenomenon has very recently raised researchers’ interest in understanding its long-term origins and subsequent variation across the continent (Anderson (2018); Bertocchi and Dimico (2018)).

In parallel, researchers are also increasingly coming to understand the role of transmitted cultural traits on contemporaneous individuals’ behaviour (Giuliano and Nunn (2017)). As such, ancestral kinship organizations, fundamental social institutions which determine for many societies group membership and social obligations to group members, have been recently emphasized as deep-rooted determinants of individuals’ behaviour (Moscona et al. (2019)). As an example, building on the anthropological literature, Lowes (2018a) provides experimental evidences that *matrilineality* undermines spousal cooperation within household.

While in *patrilineal* kinship organizations children integrate their father’s kinship group and inheritance can only be passed on to children of male group members, in *matrilineal* kinship organizations group membership and inheritance are traced through female members, which means that children integrate their mother’s kinship group and inheritance is restricted to the children of the female group members (i.e. males do not transmit to their biological children but to their sister’s children). Further, while in *patrilineal* societies, a wife is effectively incorporated into the lineage of her husband because she is not relevant to her kin group for determination of lineage or inheritance, thereby reducing her ability to rely on her own kin group in the case of separation or conflict, in *matrilineal* societies, husbands and wives maintain strong allegiances with their own kin group, allowing matrilineal wives to benefit from an improved outside option relative to their patrilineal counterparts, thereby reducing matrilineal men’s authority over their spouses.

In this paper, I examine ancestral matrilineality as a deep-rooted determinant of contemporaneous HIV prevalence among female populations in Sub-Saharan Africa. The latest theories from the evolutionary psychology literature suggest that gender-specific sexual preferences have been shaped in the long-run by our ancestors reproductive success (Buss (2016); Diamond (1998)). As such, producing large amounts of sperm, males exhibit on average a relatively stronger preference for sexual variety than their female counterparts, who, producing limited number of ovula over their lifetime and bearing the larger share of the initial parental investment (i.e. gestating, bearing, nursing, nurturing and protecting a child) are seeking for a sexual mate that will be willing to commit in the long-term. However, I argue that, relative to their patrilineal counterparts, matrilineal males have a weaker propensity to commit in the long-term to ensure the survival of their offsprings, since these latter will not integrate

¹11.2 millions in “Eastern and Southern Africa” and 3.3 millions in “West and Central Africa”

²These numbers are from <http://aidsinfo.unaids.org/>

³9.5 millions males lived with HIV in Sub-Saharan Africa

⁴Among the 1.02 millions newly infected adult individuals in Sub-Saharan Africa in 2017, 580,000 were females.

their own kin group but their wife kin group, in addition of not inheriting from them but from their mother's brothers. In fact, in matrilineal societies husband has allegiances to his sisters, whose children he must support because they are his heirs. Therefore, ancestral matrilineal societies may have constituted specific contexts in which substituting long-term committed sexual relationships by sexual variety may have been a more beneficial sexual strategy for females' reproductive success, allowing them to have access to better genes for their offsprings ("Better Genes Theories", discussed in [subsection 2.3](#)). Furthermore, benefiting from better marriage outside option as well as being inherently valued more relative to their patrilineal counterparts, I expect matrilineal females to benefit from a relatively greater sexual autonomy and ability to implement their own sexual preferences. Consequently, I test the hypothesis that ancestral matrilineality has shaped more promiscuous contemporaneous females' sexual and contraceptive behaviour, and therefore regrettably driven higher HIV prevalence among these specific populations.

To do so, I use a representative random sample of individuals reporting their ethnicity and who were tested for HIV across 18 Sub-Saharan African countries (in fact, when possible, I use several surveys for some countries, amounting to 32 country-*Demographic Health Surveys, DHS*), and linking them to their ancestral ethnic group in the *Ethnographic Atlas*, a worldwide anthropological database containing ethnographic information on cultural aspects and ways of life of ethnic groups prior to industrialization and colonial contact, I estimate the probability of testing positive for HIV as a function of whether an individual originates from an ancestrally matrilineal ethnic group, and I indeed find a higher probability of testing positive for HIV for females originating from ancestrally matrilineal ethnic groups. Interestingly, and in line with my conceptual framework, I find that this effect is specific to female populations.

There is still the concern that the variation in ethnic group's ancestral kinship organizations exploited here may be correlated with important unobservables which are driving my main result, and has nothing to do with the channel of causality argued for in this paper. To address these concerns, in my key identification strategy I exploit the variation in ethnic groups' ancestral kinship organizations within subnational regions in Sub-Saharan countries, which enables me to include subnational region-survey fixed effects into my estimations. Doing so allows me to control for a large set of potential national and subnational legal, institutional and economic confounding factors, that may spuriously drive the observed relationship between ancestral matrilineality and contemporaneous HIV prevalence rates. Furthermore, the inclusion of a host of individual-level controls allows me to account for differential socio-economic composition between matrilineal and patrilineal individuals, in particular in education, marital behaviour and religion, which are plausible drivers of sexual and contraceptive behaviour ([Case and Paxson \(2013\)](#)). Additionally, I include a large set of ethnic-group level historical controls, computed from the *Ethnographic Atlas*, which are aimed at capturing ancestral matrilineality's ethnical covariates that may be other long-term confounding factors affecting contemporaneous gender norms and subsequent female's sexual and contraception behaviour. Furthermore, exploiting village-level geographic informations provided in DHS, I also include a host of fine-grained geographical controls meant at capturing alternative geographical channels which have been shown to shape contemporaneous variation in HIV prevalence rate in Sub-Saharan Africa. Finally, I show that my main results are not driven by differential selection in HIV testing between matrilineal and patrilineal individuals, nor reflect differences in general health status, but are specific to sexually transmitted diseases.

Despite the great care provided in controlling for a large range of observables, I go one step further by formally

testing for omitted bias, using [Altonji et al. \(2005\)](#) ratios and estimating [Oster \(2017\)](#) bias-adjusted lower bound coefficients. All these exercises provide very little support to the presence of an omitted variable bias in my OLS estimates.

Though I show that my OLS estimates are very unlikely to be affected by omitted variable bias, reverse causality may remain a concern if societies in which females had more promiscuous sexual behaviours subsequently adopted matrilineal kinship organizations. In order to provide further support to the causal interpretation of my main individual-level OLS findings, I exploit data on the GPS location of DHS villages as well as the digitized Murdock's map of ancestral ethnic groups in Africa ([Figure 2](#)), and implement three alternative identification strategies, all of them having the common feature of allowing me to estimate the effect of the variation in cultural trait within similar environments. First, I implement an instrumental variable (IV) strategy, instrumenting for individual's ethnic group's ancestral matrilineality by the distance between the location of DHS villages and the nearest ancestral matrilineal ethnic boundary, and controlling for a large array of potential confounders. My IV-2SLS estimates provide additional support to the main findings found in my OLS estimations.

Second, I conduct a geographic regression discontinuity (RD) as an alternative identification strategy, exploiting the measure of distance between DHS villages and ancestrally matrilineal geographic areas that I computed. Restricting my attention to individuals living in DHS villages which are located close to an ancestral matrilineal ethnic boundary, I estimate the effect of living in a village located on the matrilineal side of the ancestral border, while controlling for geographic location/distance running variables. I find that RD estimates are qualitatively similar to OLS (though smaller in magnitudes, due to presumably blurry ancestral ethnic boundaries and potential spillovers at the border), and highlight again ancestral matrilineality as a driver of contemporaneous female HIV rates.

Finally, I estimate the average treatment over treated effect of being located on an ancestrally matrilineal geographic area on DHS villages' proportion of HIV positive females, matching villages with their nearest neighbor in an ancestrally patrilineal geographic area based on a large array of geographic observables. Finding that villages located in an ancestrally matrilineal areas are characterized by significantly higher females HIV rates than their nearest neighbor in patrilineal areas, I provide additional support to my main OLS findings

To explain such long-term effect of ancestral matrilineality on contemporaneous female HIV in Sub-Saharan Africa, and consistent with my conceptual framework, I provide evidence that, benefiting from a higher social status, bargaining power and subsequent higher sexual autonomy, matrilineal females adopt riskier sexual behaviour which are more conducive to HIV, *as compared to their patrilineal counterparts*. In line with this findings, I also show that women originating from ancestrally matrilineal ethnic groups are, relative to women originating from ancestrally patrilineal ethnic groups, more likely to be HIV positive while having a HIV negative husband, a result consistent with my conceptual framework and which suggests extramarital channels of transmission of the virus. [Anderson \(2018\)](#) having highlighted intramarital routes of infection for less empowered females, this result calls for complementary policies. Additionally, I underline that matrilineal females' higher contraception-related decision-making translate into them substituting short-term contraceptive method (condom in particular) by long-term contraceptive methods. However, by doing so, matrilineal females incidentally substitute contraceptive methods that are more protective against HIV by methods that are less protective. Consequently, I find that differences in contraceptive behaviours also explain the relatively higher rates of HIV found among matrilineal female populations.

Nevertheless, I also find evidences that when they have internalized the risk of infection, matrilineal individuals are then more likely to adopt condom as a contraceptive method. This last result is promising and call for policies aimed at raising awareness about the risk associated with promiscuous behaviours, to fight against the spread of HIV for this population at high risk. Finally, I provide additional evidences that discard differences in access to condom and in sexual debuts as other explaining mechanisms. Indeed, I find that matrilineal females have in fact an easier access to condom and begin their sexual life later than their patrilineal counterparts.

Last but not least, I build on the epidemiological literature ([Tassier \(2013\)](#); [Worden and Porco \(2017\)](#)) to perform a numerical simulation exercise, aimed at illustrating how the differences in gender-specific sexual and contraceptive behaviours found between matrilineal and patrilineal populations translate into differences in gender-specific HIV rate dynamics between these populations. Under credible parameter values (most of them being determined by my empirical findings), my simulations indeed shows dynamics that are consistent with my conceptual framework as well as my empirical findings.

This paper contributes to several strands of literature. It contributes first to a recently growing literature exploring the long-term determinants of female HIV, with a specific focus on Sub-Saharan Africa. [Anderson \(2018\)](#) explores the legal origins of female HIV in Sub-Saharan Africa and show that current HIV prevalence rates are higher for women living in common law countries, as compared to women living in civil law countries. She argues that female's property rights being weaker in common law countries, women would suffer from less intrahousehold bargaining power and related ability to impose safe sexual practices to their husbands in such countries. [Bertocchi and Dimico \(2018\)](#) highlight historical slave trade as an other long-term determinant of female HIV in the continent. More precisely, they expose that historical slave trade would have fostered current polygynous practices, which are associated with more female's marital dissatisfaction. Therefore, females in polygynous union would be more likely to adopt riskier sexual behaviour increasing their likelihood to contract and transmit the virus, through the husbands, to their faithful co-wives, with a multiplicative effect among women. [Cagé and Rueda \(2018\)](#) show that contemporaneous geographical variation accross Sub-Saharan Africa is influenced by Protestant and Catholic missions in the early 20th century, as well as their health investments. According to their result, such missions have conflicting effects on HIV today: regions close to historical mission stations exhibit higher HIV prevalence, likely because of a lower knowledge about condom use due to the persistent effect of conversion in those regions. On the contrary, among regions historically close to missionary settlements, proximity to a mission with a health investment is associated with lower HIV prevalence nowadays. They propose safer sexual behaviours as well as the persistence of health infrastructures around these missions as possible explanatory channel.

Contemporaneous determinants of the spread of HIV in Africa have also been emphasized. [Oster \(2005\)](#) argue that Africa very high HIV transmission rates are likely due to high rates of other untreated sexually transmitted infections, and that within continent differences in HIV transmission rates can be attributed to differences in sexual behaviour and epidemic timing. [Oster \(2012b\)](#) provides evidences of a fairly consistent positive relationship between exports and new HIV infections, suggesting that increased exports increase the movement of people (trucking), which increases sexual contacts. [Corno and De Walque \(2012\)](#) underline mine workers' migration as a driver if HIV infection of both mine workers and their wives since both of them are less likely to adopt safe sexual behaviours. More precisely, they find that miners are less likely to abstain and to use condoms, in particular during occasional

sexual intercourse; while women with a miner as a partner are less likely to abstain, to be faithful or to use condom with their miner husbands. [Case and Paxson \(2013\)](#) explore the adverse role of girl's education on female HIV in Africa and find that, by delaying teen marriage, increase in girls' schooling in some regions triggered risky adolescent sexual behaviour, more conducive to HIV. Consequently, they find that regions that had higher rates of female education in 1980's have higher HIV rates today.

This paper also contributes to the literature aimed at understanding the influence of HIV risk perception on sexual and contraceptive behaviour. [Oster \(2012a\)](#) investigates the lack of behavioural change despite the high prevalence of HIV in Sub-Saharan Africa, and finds reduction in risky sexual behaviour only in areas with higher life expectancy, since, consistent with optimizing behaviour, high rates of non-HIV mortality suppress behavioural response; but she does not find evidence of greater behavioural change in areas with higher knowledge of the epidemic. [Bjorkman Nyqvist et al. \(Forthcoming\)](#) provides experimental evidences that behavioural response from lottery incentives on HIV prevention are higher for risk-lover individuals, who are also individuals with ex-ante riskier sexual behaviour. [Thornton \(2008\)](#) provides experimental evidences showing that learning HIV positive status induces an increase in condom use, while learning HIV negative status does not affect condom use behaviour. [Paula et al. \(2014\)](#) show that downward revisions in the belief assigned to being HIV positive increase risky behaviour, while upward revisions decrease it. In the same vein, [Delavande and Kohler \(2015\)](#) show that risky sexual behaviour is influenced by individuals' expectations about survival, and future HIV status, which in turn depend on the perceived impact of HIV/AIDS on survival, expectations about own and partner's current HIV status, and expectations about HIV transmission rates. Interestingly, they find that subjective expectations, in particular about mortality risk but not the risk of living with HIV, play an important role in determining the decision to have multiple sexual partners.

This paper also relates to the literature emphasizing the cultural factors shaping contraception use, crucial for limiting the spread of HIV. In the context of Bangladesh, [Islam et al. \(2009\)](#) underly that matrilineal *Garos* women contraceptive behaviour differs from that of their *patrilineal* Bengali counterparts, in that their current use of contraceptives is higher than at national level, but the prevalence of modern male methods is low. In Sub-Saharan African context, [Cordero-Coma and Breen \(2012\)](#) emphasize fidelity norm and reproduction norm as two of the fundamental elements that guide spouses' condom use behaviour in Sub-Saharan African context. Focusing on rural Malawi, [Chimbiri \(2007\)](#) documents notably that condom use is negligible inside marriage, and that initiating a discussion of condom use for preventing infection in marriage is like bringing an "intruder" into the domestic space.

More broadly, this papers speaks to the literature exploring the long-term determinants of contemporaneous gender outcomes, which can be found summarized in [Giuliano \(2017\)](#). [Alesina et al. \(2013\)](#) find that descendants of societies that traditionally practiced plough agriculture, characterized by gender division of labor, today have less equal gender norms. [Alesina et al. \(2016\)](#) provide evidences that pre-colonial customs about marriage patterns, living arrangements and the productive role of women are associated with contemporary violence against women. [Becker \(2018\)](#) highlights that pre-industrial economic production relying on pastoralism favored the adoption of customs aimed at constraining female sexuality, such as a particularly invasive form of female genital cutting, restrictions on women's mobility, and norms about female sexual behaviour, which are still at play. [Teso \(2019\)](#) shows that women whose ancestors were more exposed to the transatlantic slave trade, leading to female-biased

sex ratio during centuries, in Sub-Saharan Africa, are today more likely to be in the labor force, have lower levels of fertility, and are more likely to participate in household decisions.

Finally, this paper also links more directly to the burgeoning literature assessing the influence of ancestral norms on women’s well-being. [Lowes \(2018a\)](#) provide evidences from Democratic Republic of Congo that men and women from matrilineal ethnic groups cooperate less with their spouses in a lab experiment, while is not the case when paired with a stranger of the opposite sex. She also finds that children of matrilineal women are healthier and better educated, and matrilineal women experience less domestic violence. [Ashraf et al. \(2019\)](#) highlight the positive role of the bride price custom on female education in Indonesia and Zambia; while [Bargain et al. \(2019\)](#) show the complementarity between access to justice and ancestral matrilocality on women’s empowerment in Indonesia.

The remainder of the paper is organized as follows. In [section 2](#), I provide an overview of ancestral matrilineality as well as contemporaneous female HIV in Sub-Saharan Africa, and I expose my conceptual framework. Then, I provide a description of the data and I discuss my empirical strategy in [section 3](#). My main results, robustness checks and alternative identification strategies are presented in [section 4](#). I then explore the mechanisms in [section 5](#), and I conduct an epidemiological simulation exercise in [section 6](#). Finally, I provide concluding thoughts in [section 7](#).

2 Context and Conceptual Framework

2.1 Ancestral Matrilineality in Sub-Saharan Africa

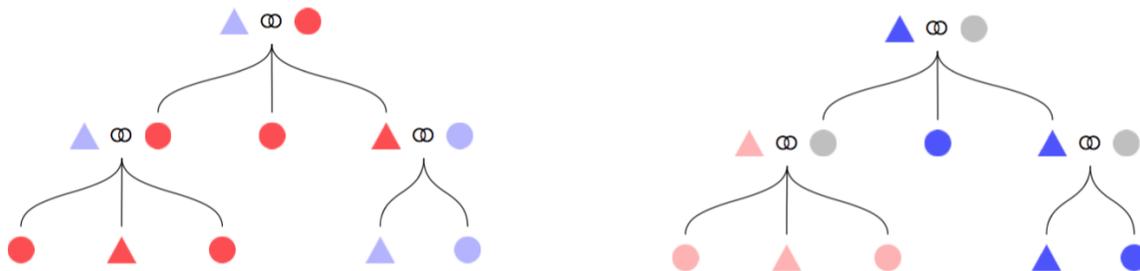
In matrilineal kinship system, individuals trace lineage and descent through women. Under this kinship system, while biologically related to family of both their mother’s side and their father’s side, individuals are considered kin only if they share a common female ancestor. Such kinship relations are important in the context of Sub-Saharan Africa where kinship groups form a basic political unit in which members recognize each other as kin and often have certain obligations toward each other, such as land sharing, contribution to bride price payments for lineage members, provision of financial support (school fees, burial payments, etc.) ([Fox \(1934\)](#)).

[Figure 1 \(a\)](#) is from [Lowes \(2018a\)](#) and illustrates the structure of matrilineal kinship systems. In the diagram, men are represented by triangles and women are represented by circles. Membership in the same matrilineal group is denoted with red. Children are in the same matrilineal group as their mothers. Likewise, a mother is in the same matrilineal group as her male and female siblings. In many matrilineal societies, the mother’s brother has an important role relative to his sister’s children. His inheritance and lineage will be traced through his sister’s children, and he has obligations to financially support her children. Importantly, husband and wife do not share the same lineage (i.e. even once married, the wife remains in her lineage of origin) - for all married couples one spouse is red and the other spouse is blue. Consequently, as noted by anthropologists, husbands are less able to mistreat their spouses in matrilineal systems since these latter have greater support from their kin groups.

[Figure 1 \(b\)](#) is also from [Lowes \(2018a\)](#) and presents the structure of patrilineal kinship. Now, children are in the same group as their father, as denoted in blue. In a patrilineal society, rather than maintaining strong ties with her own lineage, a woman is effectively incorporated into the lineage of her husband upon marriage. This is because once she is married, she is not relevant for determining descent and inheritance for her lineage. This is illustrated in the patrilineal kinship diagram by the married women denoted in grey, while the unmarried daughter shares the

same color as her father.

Figure 1: Diagrams of kinship systems (source: Lowes (2018a))



Legend: ▲ = Males, ● = Females, ●/▲ = Same Matriliney

Legend: ▲ = Males, ● = Females, ●/▲ = Same Patriliney

(a) Matrilineal Kinship

(b) Patrilineal Kinship

Historically, matrilineal kinship systems are correlated with other cultural traits, which have been shown to have long-term impact on gender roles and economic development⁵. For example, Lowes (2018a) shows that, in Africa⁶, matrilineality is negatively correlated with the practice of bride price, the use of the plough as well as animal husbandry. As reported in Table 10 in appendix, I find similar evidences looking at my main sample’s descriptive statistics. Therefore, as described in subsection 3.3, it will be crucial to control for all these historical correlates.

15 percent of the 527 Sub-Saharan societies recorded in the Ethnographic Atlas are matrilineal; while 70 percent are patrilineal. Furthermore, the vast majority of these matrilineal societies are distributed across the center of Africa in the so called “*Matrilineal Belt*”; which intersects present day Angola, Republic of Congo, DRC, Gabon, Malawi, Mozambique, Namibia, Tanzania and Zambia. Figure 2 illustrates the matrilineal belt across Africa, with matrilineal groups denoted in blue, patrilineal groups denoted in green, and bilateral and other groups in beige⁷. In my final sample, about 67% of individuals originating from an ancestrally matrilineal ethnic groups are indeed living in one of the countries crossed by the matrilineal belt. Interestingly, the highest female HIV rates in Sub-Saharan Africa are also found in the countries crossed by the matrilineal belt. In fact, ones talk about an “*AIDS Belt*”. I describe it in more detail in the next subsection.

2.2 Female HIV in Sub-Saharan Africa

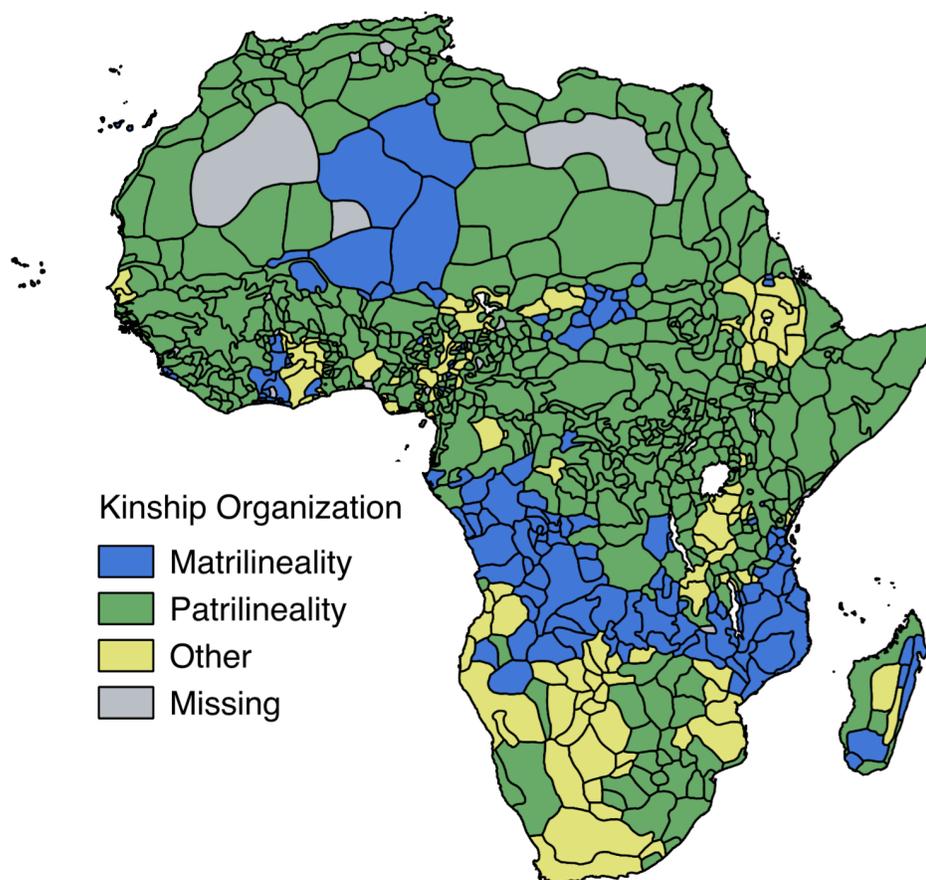
In 2018, 14.5 millions of women lived with HIV in Sub-Saharan Africa, representing about 80% of the world-wide female HIV positive population (UNAIDS (2018b)). Uniquely, it is the only place in the world where more women than men live with HIV: adult female HIV positive population is about 1.5 larger than male HIV positive

⁵Interested reader may find an exhaustive overview of the origins of matrilineal kinship systems and related women’s empowerment in Lowes (2018a) appendix.

⁶Using data from the *Ethnographic Atlas*.

⁷This map is based on Murdock’s map of ethnic group’s ancestral boundaries. Note that Murdock’s classification of ethnic groups slightly differs from the Ethnographic Atlas (EA)’s classification. I therefore use Michalopoulos et al. (2019) and Teso (2019) mappings of Murdock-EA ethnic groups.

Figure 2: Ancestral Ethnic Group Boundaries and Matrilineal Belt



population⁸, and adult females represented about 57% of new adult infections in 2017⁹. This phenomenon has been referred to “Feminization of HIV” in Sub-Saharan Africa¹⁰.

Interestingly, these numbers hide substantial within Sub-Saharan Africa variation in female HIV rates. Indeed, among the 14.5 millions of women living with HIV in Sub-Saharan Africa in 2018, 11.2 millions lived in “Eastern and Southern Africa”, against 3.3 millions in “West and Central Africa”. The geographical variation in female HIV rates in Sub-Saharan Africa, computed from observations in my final sample using DHS data and linking individuals to their ancestral ethnicity, is represented in Figure 3¹¹ and shows high prevalence rate in Eastern and Southern Africa, the so-called “*AIDS Belt*”¹². Zambia and Malawi, two of the countries with highest female HIV rates are also two of the countries with the highest proportions of individuals originating from matrilineal ethnic groups¹³.

⁸9.5 millions males live with HIV in Sub-Saharan Africa

⁹Among the 1.02 millions newly infected adult individuals in Sub-Saharan Africa in 2017, 580,000 were females.

¹⁰These numbers are from <http://aidsinfo.unaids.org/>

¹¹This map is based on Murdock’s map of ethnic group’s ancestral boundaries. Note that Murdock’s classification of ethnic groups slightly differs from the Ethnographic Atlas (EA)’s classification. I therefore use Michalopoulos et al. (2019) and Teso (2019) mappings of Murdock-EA ethnic groups.

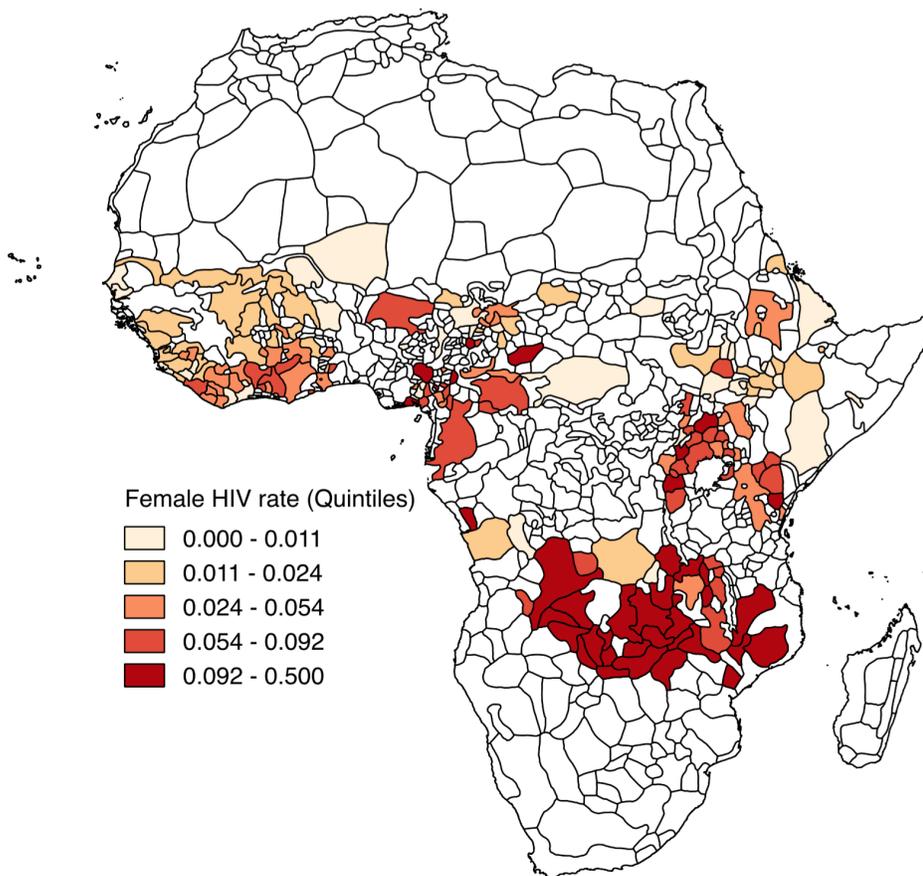
¹²Interested reader might have a look at <https://www.hsph.harvard.edu/news/magazine/spr08circumcisionmap/>; and/or <https://www.prb.org/thestatusofthehivaidsinsubsaharanafrica/>

¹³In my final sample of matrilineal versus patrilineal females, proportions of matrilineal females are about 87% in Malawi and about 60% in Zambia; female HIV rates are about 12% in Malawi and about 14% in Zambia. These numbers are computed using the following DHS surveys: Malawi (2004, 2010, 2014) and Zambia (2007, 2013). See subsection 3.1 for more details. This compares to the proportion of about 15% of matrilineal females, and the average female HIV rate of about 5% in my full final sample of matrilineal versus patrilineal females.

As detailed in [section 1](#), several factors have been emphasized to explain variation in female HIV rates within Sub-Saharan Africa. Institutions have highlighted gender inequality and disempowerment as key barriers to progress against the HIV epidemic ([UNAIDS \(2018a\)](#)), and urged for effort to address these issues. A widespread conjecture is that strengthening women’s property and inheritance rights will prevent the spread of HIV/AIDS by promoting women’s economic security and empowerment. [Anderson \(2018\)](#) provides the first empirical evidence of a causal relationship between female bargaining power and female HIV infection rates in Sub-Saharan Africa. Exploiting variation in legal origins of Sub-Saharan countries, she finds that HIV prevalence is today higher for women living in common law countries, where code of law is associated with weaker female property rights, as compared to women living in civil law countries, due to lower intrahousehold bargaining power and ability to impose safe sexual practice to their husband.

However, this mechanism does not explain the geographical correlation between the so-called “*Matrilineal Belt*” and the so-called “*AIDS Belt*” in Africa, according to which the highest rates of female HIV prevalence can be found in ancestrally matrilineal geographical areas, correlated with higher women’s status and sexual autonomy. I explore this counterintuitive pattern in this paper. To do so, I begin with a description of my conceptual framework in the next subsection.

Figure 3: Ancestral Ethnic Group Boundaries and Contemporaneous Female HIV Rates (Final sample)



2.3 Conceptual Framework: Gender Differences in Sexual Strategies

Humans, like other sexually reproducing species, do not choose mate randomly. According to the latest theories from evolved psychologists, our mating is strategic, and the sexual strategies we developed were shaped in the (very) long run by natural selection through our ability to survive and reproduce successfully. As [Buss \(2016\)](#) states: *“Those in our evolutionary past who failed to mate successfully failed to become our ancestors. All of us descend from a long and unbroken line of ancestors who competed successfully for desirable mates, attracted mates who were reproductively valuable, retained mates long enough to reproduce, fended off interested rivals, and solved the problems that could have impeded reproductive success. We carry in us the sexual legacy of those success stories.”* Further, underlying each sexual strategy¹⁴ are psychological adaptations (e.g. preferences for a mate, feelings of love, desire for sex, sexual jealousy, etc.), that are sensitive to the information or cues from the external world.

These theories provide an interesting framework to understand why males and females exhibit on the average marked different sexual preferences according to numerous psychological studies¹⁵, with males exhibiting a stronger (relative to females) preference for casual relationships and sexual variety; and females tending to be relatively more choosy in their mate as well as relatively more looking for long-term committed relationships. Indeed, producing million of sperms, which are replenished at a rate of roughly 12 millions per hour ([Buss \(2016\)](#)), males can expect at most to reproduce as many times as there are available fertile females willing to have sex. Additionally, males only bear minimal initial parental investment in case of fecundation. Consequently, preference for sexual variety and casual sexual relationships is a psychological trait that has been relatively more developed in male evolved sexual psychology. On the contrary, producing only a fixed limited amount of reproductive cells (i.e. only 400 ova mature to the point where they are capable of being fertilized), females face a limited number of reproductive opportunities. In addition, bearing the larger share of the initial parental investment (i.e. gestating, bearing, nursing, nurturing and protecting a child), females provide extraordinarily valuable, but limited, reproductive resources. Therefore, because females in our evolutionary past risked enormous investment from having sex, evolution favored females who were highly selective about their mates. In particular, evolution favored females capable of reproducing with long-term committed and reliable provider mates, ensuring the survival and reproduction success of their offsprings ([Diamond \(1998\)](#)).

Nevertheless, while the reproductive benefits of casual relationships for males are large and direct, women may also reap benefits from short-term mating, according to the latest evolutionary psychology theories. As a matter of fact, even though having access to more sperm would not increase a woman’s reproductive success, it is possible for a woman to gain superior genes through casual sex that are passed on to their children. This is the so-called “Better Genes Theory”¹⁶ ([Buss \(2016\)](#); [Greiling and Buss \(2000\)](#)). As such, a woman might try to secure the investment of a lower-ranking man by marrying him, for example, while simultaneously securing the genes of a higher-ranking man by mating with him casually (the mating marketplace rendering far easier for a woman to get a man with better genes to have sex with her than to get him to marry her) ([Buss \(2016\)](#)). One version of this “Better Genes Theory” has been called the “Sexy Son Hypothesis”, according to which women prefer to have casual sex with men who are

¹⁴“Sexual strategies do not require conscious planning or awareness. [...] Indeed, just as pianist’s sudden awareness of her hands may impede performance, most human sexual strategies are most successfully carried out without the awareness of the actor. ([Buss \(2016\)](#))”

¹⁵See [Buss \(2016\)](#) for an extensive review.

¹⁶I refer reader interested in alternative hypothesis on women’s extra-pair mating function to [Buss \(2016\)](#) and [Greiling and Buss \(2000\)](#), which provide extensive discussions.

attractive to other women because they will have sons who possess the same charming characteristics and therefore will enjoy greater mating success in the next generation.

However, adopting such extra-pair mating strategies may come at a cost. In particular, an unfaithful married woman risks the withdrawal of resources by her husband, reputational damages as well as getting pregnant and bearing a child without the benefits of an investing partner. To sum up, short-term mating poses hazards, but has powerful benefits as well, and women have evolved psychological mechanisms to select circumstances in which the cost of short-term mating are minimized and the benefits increased.

Matrilineal kinship organizations, relative to patrilineal ones, constitute environments in which extra-pair mating sexual relationships should be beneficial for women. As a matter of fact, as detailed in [subsection 2.1](#), in matrilineal kinship organizations males' biological children do not integrate their lineage but the lineage of their mother, in addition of males not transmitting their inheritance to their biological children but to their sisters' children. In other words, females in matrilineal societies should have lower expectations regarding their mate's propensity to engage in long-term committed relationships to maximize the reproductive success of their offsprings. Consequently, they should also expect relatively greater benefits from gaining access to better genes, through extra-pair mating, for their reproductive success. Furthermore, contrary to patrilineal societies where females integrate their husband lineage once married, matrilineal females remain in their lineage of origin after marriage and keep strong bounds with their family of origin. Said differently, matrilineal females benefit from higher marriage outside option, relative to their patrilineal counterparts, but also from relatively lower opportunity cost associated with extra-pair mating sexual strategy. Finally, being relatively inherently valued more through their position in their societies, matrilineal females should also benefit from a greater sexual autonomy than their patrilineal counterparts, and a related greater ability to pursue their preferred sexual strategy.

If, according to my conceptual framework, ancestral matrilineality fostered females to seek for better genes through sexual variety, matrilineality should be correlated with greater genetic diversity as well as father diversification. This is exactly what I find in [Table 11](#) in appendix. As reported in the two first columns, I find a positive country-level association between the estimated proportion of countries' citizens with ancestors that had matrilineal inheritance rule (using data from [Giuliano and Nunn \(2018\)](#)) and countries' indexes of genetic diversity (computed by [Ashraf and Galor \(2013a\)](#)), robust when including continent fixed effects as well as a large array of controls, including distance to Addis Ababa, the cradle of humankind. Further, as reported in the two last columns, running individual-level regressions following my main identification strategy on my final sample from DHS (described in more detail in [subsection 3.1](#)), I find that women originating from ancestrally matrilineal ethnic groups are significantly more likely to diversify the fathers of their biological children.

All in all, my conceptual framework points toward a main prediction, that I empirically test in the paper: while natural selection has shaped, through past reproductive success, male's evolved sexual preferences towards short-term casual sexual relationships and female's evolved sexual preferences towards long-term committed sexual relationships, ancestral matrilineal kinship organizations may have constituted environments in which substituting long-term committed sexual relationships by short term casual sexual relationships could have been a dominant sexual strategy for female's reproductive success. Consequently, I expect to find that, as of today, matrilineal females adopt sexual behaviours that are more promiscuous relative to their patrilineal counterparts, and therefore

suffer from more HIV.

3 Data and Empirical Strategy

To study the long-term impact of matrilineality on female HIV, I match contemporaneous individual-level data from the Demographic and Health Surveys (DHS) with ethnic group-level data from the *Ethnographic Atlas (EA)*. This Section describes the data and the empirical strategy.

3.1 Contemporaneous Data

Data on HIV infection come from the DHS, which have been conducted in sub-Saharan African countries since the 1990's. The DHS household surveys typically interview a nationally representative sample of between 10,000 to 20,000 women (aged 15-49) and men (aged 15-59). By collecting blood for HIV testing from representative samples of the population, the DHS Program provides nationally representative estimates of HIV prevalence rates. The testing is simple: the interviewer collects dried blood spots (DBS) on filter paper from a finger prick and the filter paper is transported to a laboratory for testing. The testing is anonymous, voluntary, and non-informative to respondents. The average response rate is extremely high; 93 percent for women (slightly lower for men).

I restrict my sample to DHS surveys containing both HIV testing and individual ethnicity information, as well as GPS data. I further restrict my samples to individuals originating from an ethnic group which is either ancestrally matrilineal, or patrilineal. This leaves me with a sample of 159,560 women across 18 Sub-Saharan African countries (i.e. 32 country-surveys¹⁷). The proportion of men tested for HIV is lower and I have a sample of 120,580 men. As a main outcome, I build an individual-level indicator variable, *HIV*, that takes value one if the respondent is seropositive according to the DHS HIV Test.

On average, 4.8 percent of women in my initial sample are HIV positive (this compares to 3.1 percent of men). The average HIV infection rate of women originating from an ancestrally matrilineal ethnic group is approximately 11.4 percent. It is close to one-fourth, at 3.7 percent, for women originating from an ancestrally patrilineal ethnic group. This compares to 7.8 percent for matrilineal males vs. 2.1 percent for patrilineal males. Though only correlational, these numbers provide a first evidence of the higher prevalence of HIV among individuals originating from ancestrally matrilineal group, women in particular.

Further, DHS surveys present questions that are useful measures of female sexual autonomy, actual sexual behaviour, contraception use, acknowledgment about HIV risks and protective methods, and sexual debuts, which I investigate in [section 6](#).

Finally, I also exploit information on geographic covariates, computed at the village level from numerous data sources, and provided by the DHS. Village-level geographic controls are described in [subsection 3.3](#).

¹⁷Burkina-Faso (2003, 2010); Cameroon (2004, 2011); Chad (2014); Congo Democratic Republic (2007, 2013); Ethiopia (2005, 2011, 2016); Gabon (2012); Ghana (2003, 2014); Guinea (2005, 2012); Ivory Coast (2011); Kenya (2003, 2008); Liberia (2013); Malawi (2004, 2010, 2014); Mali (2006, 2012); Senegal (2005, 2010); Sierra Leone (2008, 2013); Togo (2013); Uganda (2011); Zambia (2007, 2013)

3.2 Historical Data

Data on pre-colonial ethnic groups traits come from the *Ethnographic Atlas (EA, 1967)*, a worldwide ethnicity-level database constructed by George Peter Murdock, covering 1,265 ethnic groups in the world¹⁸, and that contains detailed ethnographic information on cultural aspects and ways of life of the portrayed ethnic groups, *prior to industrialization and colonial contact*¹⁹, such as kinship and family organization, settlement patterns, political organization, institutional complexity, historical mode of subsistence, etc.

To match individual-level contemporaneous data with historical data, I use information provided in DHS on individual’s ethnicity. However, the classification of the respondents’ ethnic groups used in the DHS does not always coincide with the *Ethnographic Atlas*’ one, requiring a matching between the two datasets²⁰. Therefore, I first follow [Michalopoulos et al. \(2019\)](#) matching, which enables me to match most of the individuals in my sample. I then follow [Teso \(2019\)](#) matching²¹ to match some remaining unmatched individuals. Finally, I build on online sources²² to match ethnic groups not matched in previous procedures. I end up with a sample of 349,895 individuals in DHS matched with the *Ethnographic Atlas (EA)*.

I then discard individuals originating from ethnic groups with forms of inheritance rule and kinship organization which are neither matrilineal, nor patrilineal (such as bilaterality of ambilinearity), or with missing information on such inheritance rule, representing 69,755 out of 349,895 individuals in my sample. I therefore restrict my sample to individuals originating from an either ancestrally matrilineal or patrilineal ethnic group, and I end up with a sample of 280,140 individuals (159,560 women and 120,580 men).

Finally, I construct my main explanatory variable, *Matrilineality*, an indicator of ethnic-group’s ancestral matrilineality, from EA information on ethnic group’s inheritance rule. In my analysis, I additionally use a wide array of historical controls varying at the ethnic group level and computed from the EA: I describe these controls in the next subsection.

3.3 OLS Empirical Strategy

I explore the gender-specific long term effect of ancestral matrilineality on HIV by estimating the following equation:

$$y_{ievert} = \alpha + \beta_1 \text{Matrilineality}_e + \beta_2 \text{Female}_i + \beta_3 \text{Female}_i \times \text{Matrilineality}_e + \mathbf{X}'_{ievert} \Delta + \mathbf{X}'_{ert} \Omega + \mathbf{X}'_{vrt} \Pi + \lambda_{rt} + \varepsilon_{ievert} \quad (1)$$

with y_{ievert} denoting an individual-level outcome (for example an indicator for whether the individual is HIV

¹⁸The majority of the ethnicities sampled are in Africa.

¹⁹“For the parts of the world without a written history, the information is from the earliest observers of these cultures. For some cultures, the first recorded information is from the 20th century, but even for these cultures, the data capture as much as possible the characteristics of the ethnic group prior to European contact.” ([Alesina et al. \(2013\)](#))

²⁰For many of the groups, the matching is straightforward as the name used in the DHS is the same or very similar to the one used in *Ethnographic Atlas (EA)*. When the name of an ethnic group is not found in EA’s classification, this is typically because an alternative group’s name is used.

²¹[Teso \(2019\)](#) matching procedure largely builds on [Michalopoulos et al. \(2019\)](#) matching procedure.

²²One of the most useful sources of information on alternative ethnic groups’ names is the Joshua Project website (<http://www.joshuaproject.net/>). For most of the unmatched ethnicities, the respondent lists her nationality as ethnicity.

positive) for individual i from ethnic group e , living in village v in within-country DHS region r , and surveyed at year t . $Matrilineality_e$ is an indicator for whether an individual originates from an ancestrally matrilineal ethnic group (an ancestrally patrilineal ethnic group otherwise); $Female_i$ is an indicator for whether an individual is a female. β_1 is intended to capture a “Matrilineal effect”; and β_2 is intended to capture a “Gender effect” on HIV²³. β_3 is my main coefficient of interest and captures the effect of ancestral matrilineality on female HIV once the “Matrilineal effect” and the “Gender effect” have been controlled for. The inclusion of this interaction term is motivated by my conceptual framework, according to which originating from an ethnic group with an ancestral matrilineal kinship organization should significantly influence contemporaneous sexual behaviour for female individuals specifically. This hypothesis is further motivated by the first descriptive statistics reported in [subsection 3.1](#) on heterogeneity of contemporaneous HIV rates by gender and individual’s ethnic group’s ancestral kinship organization.

\mathbf{X}'_{ievert} represents a set of individual-level covariates which includes indicators of marital status; a dummy for whether an individual is in a polygynous union; number of children; age; age squared; a dummy for whether an individual lives in an urban location; education in number of years; a dummy for whether an individual is currently working; wealth index indicators and religion indicators.

\mathbf{X}'_{ert} represents a set of ethnic group-level ancestral covariates which includes intensity of women’s historical participation in agriculture; ancestral polygyny; ancestral bride price; ancestral plough use; ancestral pastoralism; ancestral presence of clans; indicators of ancestral settlement patterns; indicators of ancestral jurisdictional hierarchies beyond local communities; ancestral reliance on hunting; ancestral reliance on fishing; ancestral reliance on gathering; ancestral reliance on animal husbandry; ancestral reliance on agriculture; ancestral presence of large domesticated animals; indicators of intensity of ancestral agriculture; and year of observation of the ethnic group in the *Ethnographic Atlas*.

\mathbf{X}'_{vrt} represents a set of village-level geographic covariates which includes latitude; longitude; altitude; nightlight composite; population density (2010); distance to lake or coastline; distance to nearest international border; average time (minutes) required to reach a high-density urban center (2015); malaria incidence (2010); vegetation index; indicators for the length of the growing season; distance to the nearest active mine; a dummy indicating the presence of an active mine within 1000 km; index of ethnic fractionalization; and index of ethnic polarization. Summary statistics of all my control variables are reported in [Table 10](#) in appendix.

Crucially, I also include (within-country) DHS region-survey (year) fixed effects, λ_{rt} , to take into account time trends, as well as unobserved country-level and within-country level institutional, economic and geographic factors that could potentially affect contemporaneous HIV prevalence and also be correlated with the geographical distribution of ancestral matrilineality. Doing so allows me to assess the influence of matrilineal versus patrilineal ancestral kinship organizations on individuals located in the same institutional, economic, and geographic environment.

Finally, since variation in the main explanatory variable occurs at the ethnic group level, observations of outcomes of individuals of the same ethnic group may not be independent. Further, I exploit within DHS region-survey variation in ancestral matrilineality vs. patrilineality. Consequently, in order to account for potential within-group correlation of the residuals (ε_{ievert}), throughout the analysis standard errors are clustered at the ethnic group-(within-country) DHS region-survey (year) level.

²³According to the medical literature, females are more likely than males to be infected when exposed to HIV, due to physiological and anatomical reasons, such as larger surface area of mucosal HIV exposure ([Yi et al. \(2013\)](#)).

A crucial concern for the causal interpretation of the OLS estimates is the possible presence of omitted variable(s) that are correlated with both contemporaneous HIV prevalence and with ancestral matrilineality of ethnic groups. For instance, if ancestrally matrilineal ethnic groups were more likely to have social organizations and institutions²⁴, as well as modes of production more conducive of equal gender norm²⁵ or of the spread of virus, this would translate in an estimate of β_3 that is biased upward. The ethnicity-level controls are meant to alleviate these concerns. Additionally, I include ethnic group's year of observation in *Ethnographic Atlas* to alleviate the concern that some groups were portrayed later than others and might therefore have been more developed, and hence potentially more gender equal, for example.

Likewise, β_3 might be biased upward if ancestral matrilineality of ethnic groups was correlated with geographic factors that are also conducive of HIV. For example, geographic characteristics such as the type of vegetation, or the altitude may be correlated with the spread of the virus. I directly control for that, and I also control for malaria incidence, a proxy measure of virus spread. Further, following [Corno and De Walque \(2012\)](#), mine workers' international migration is an other driver of the spread of HIV. I therefore control for the village's distance to the nearest active mine, as well as for the presence of an active mine within 1000km. In the same vein, [Oster \(2012b\)](#) highlights exports and road networks, and subsequent increase in movements of people and sexual contacts, as an other factor of HIV infection. Controlling notably for distance to international borders as well as average time to reach a high-density urban center as a proxy allows me to alleviate these concerns. An other long-term determinant of female HIV in Sub-Saharan Africa put forward by [Cagé and Rueda \(2018\)](#) is the geography of Protestant and Catholic missions in the early 20th century, as well as their health investments. In the same vein, [Teso \(2019\)](#) shows the long term effect of the slave trade in Sub-Saharan Africa on contemporaneous gender norms. Again, my OLS estimates would be biased if ancestral matrilineality was correlated with such factors. The inclusion of numerous geographic covariates, such as latitude, longitude as well as distance to lake or coastline and average time to reach a high-density urban center, is meant to alleviate these concerns.

Along the same lines, β_3 might also be biased upward if ancestral matrilineality was correlated with ethnic fractionalization and/or ethnic polarization. Indeed, [Tequame \(2012\)](#) provides evidences that these latter are associated with higher information asymmetry and lower social sanctions within communities, which are more conducive to risky sexual behaviour and therefore HIV. Computing indexes of ethnic fractionalization and polarization at the village level and including them in my regressions is exactly aimed at controlling for such otherwise potential omitted bias.

Ethnic groups' economic prosperity may be an other possible omitted variable, potentially correlated with both ancestral matrilineality and contemporaneous HIV prevalence. To account for this, I control for village's population density (2010) and nightlight composite, two proxies of economic prosperity. The inclusion of historical ethnic groups' modes of production and institutional controls are also meant to capture historical economic prosperity. Again, including an ethnic group's year of observation as a control allow me to alleviate the concern that some groups were portrayed later than others and might therefore have been more developed. Related to this concern, ethnic group's access to contemporaneous health infrastructure could be an other possible omitted variable: controlling for average time (minutes) required to reach a high-density urban center (2015) is meant to control for this.

²⁴e.g. [Ashraf et al. \(2019\)](#) on the role of bride price

²⁵e.g. [Alesina et al. \(2013\)](#) on the role of plough use

Finally, institutional factors may also drive female HIV. As an illustration, [Anderson \(2018\)](#) underlines that legal origins of Sub-Saharan African countries significantly determine current-day female HIV rates: she namely finds that female HIV rates are significantly higher in common law sub-Saharan African countries, compared to civil law ones. The inclusion of (within-country) DHS region x survey (year) fixed effects is notably meant to capture such country (and within-country) institutional factors, in addition of time trends.

4 The Long-Term Effect of Ancestral Matrilineality on Female HIV

4.1 Main Results - OLS Estimates

[Table 11](#) presents the OLS estimates of the effect of ancestral matrilineality on contemporaneous female HIV, once gender and ancestral matrilineality have been controlled for. While in column 1 I do not include any control, I include (within-country) region-survey (year) fixed effects in column 2. The coefficient of the interacted variable *Female* × *Matrilineality* is positive and statistically significant (row 3), and is not affected by the inclusion of region-survey fixed effects. Further, following the inclusion, in column 3, of individual controls to account for differences in socio-demographic composition of matrilineal and patrilineal individuals, the main estimate of interest remains unchanged. To alleviate omitted variable concerns detailed in [subsection 3.3](#), I also subsequently include ethnic group’s controls in my regressions (column 4), and village level geographic controls (column 5). Again, the coefficient of the interacted variable *Female* × *Matrilineality* remains very consistent and of large magnitude across the specifications: I find that women originating from an ancestrally matrilineal ethnic group are 1.6 to 2.2 percentage points more likely to be HIV positive than their patrilineal counterparts; an effect that is large in magnitude, corresponding to 70% to 104% of the average HIV prevalence (2.1% to 2.3%) among patrilineal males (control group) in my full regression sample. This result supports my main prediction that, because ancestral matrilineal kinship organizations, marked by lower males’ investment in their biological offsprings, constituted specific environments in which adoption of more promiscuous sexual behaviours may have been a successful reproductive strategy for females, contemporaneous matrilineal female populations suffer more from HIV than their patrilineal counterparts.

Interestingly, I also find across my specifications a consistent and statistically significant positive estimate of being a female on the likelihood of being HIV positive (row 2). This is an additional evidence of the well-documented “Feminization of HIV in Sub-Saharan Africa” discussed in [section 1](#). As highlighted in [Greenwood et al. \(2019\)](#) and in the medical literature, for physiological and anatomical reasons, women are at higher risk than men to contract the virus when exposed to it ([Yi et al. \(2013\)](#)).

Importantly, the “Matrilineal effect” (row 1) becomes non-significant and very close to zero, once region-survey fixed effects are included. Ancestral matrilineality in Sub-Saharan Africa being essentially geographically located in countries of the so-called “Matrilineal Belt”, this suggests that the unconditional positive significant effect of *matrilineality* found in column 1 may in fact capture the effect of other country level and within-country level factors. Among them, legal systems and codes of law could be a plausible candidate. Indeed, [Anderson \(2018\)](#) highlights that female HIV rates are significantly higher in common law sub-Saharan African countries, compared to civil law ones. Since most of countries of the “Matrilineal Belt” are common law countries, the “matrilineal effect” found in column 1 might in fact be a “Common Law Countries effect” (I discuss it in more details in the following

subsection). However, this effect disappears once (within-country) region-survey fixed effects are included, and my main coefficient of interest remains consistent, positive and statistically significant at the 1% level.

All in all, the absence of significance of the “matrilineal effect” as well as the positive significant effect of my main variable of interest (*Female × Matrilineality*) is in line with my conceptual framework according to which ancestral matrilineality has a long-lasting effect specific on female’s contemporaneous sexual behaviour and consequent HIV. This result is confirmed by the analysis made by gender subsamples, and presented in the columns 6 and 7 of [Table 11](#): I find that women originating from an ancestral matrilineal ethnic group are 1.2 percentage points (column 6) significantly more likely to be HIV positive (representing 31% of the average HIV prevalence of patrilineal females in my regression sample); while I do not find any significant effect for their male counterparts (column 7).

Table 1: **The Effect of Ancestral Matrilineality on Female HIV (OLS)**

	HIV						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Matrilineality	0.058*** (0.006)	0.000 (0.005)	0.001 (0.004)	-0.000 (0.005)	-0.001 (0.006)	0.012* (0.007)	-0.004 (0.006)
Female	0.016*** (0.001)	0.012*** (0.001)	0.009*** (0.001)	0.009*** (0.001)	0.008*** (0.001)		
Female × Matrilineality	0.020*** (0.004)	0.022*** (0.004)	0.019*** (0.004)	0.016*** (0.004)	0.017*** (0.004)		
Ind. Controls			Yes	Yes	Yes	Yes	Yes
Ethnic Group Controls				Yes	Yes	Yes	Yes
Village-Geographic Controls					Yes	Yes	Yes
Region-survey FE		Yes	Yes	Yes	Yes	Yes	Yes
Gender	Both	Both	Both	Both	Both	Female	Male
Observations	280,140	280,140	273,417	193,991	182,312	105,964	76,348
Adj. R-squared	0.018	0.050	0.078	0.076	0.078	0.087	0.065
Clusters	2,537	2,537	2,515	1,711	1,688	1,550	1,402
Mean Dep. Var. (Patri. Males)	0.021	0.021	0.021	0.023	0.023	0.039	0.023

Notes: OLS estimates are reported with standard errors clustered at the ethnicity-DHS (within-country) region-survey (year) level in brackets. The unit of observation is an individual originating from either a traditionally matrilineal or a traditionally patrilineal ethnic group (it therefore excludes individuals originating from ethnic groups with alternate inheritance rules (ambilineality, bilinearity, duolinearity, etc.)). “Matrilineality” indicates (dummy) whether an individual belongs to a traditionally matrilineal ethnic group. “Female” indicates (dummy) whether an individual is a female. “**Female × Matrilineality**” indicates (dummy) whether an individual is a female belonging to a traditionally matrilineal ethnic group. “**HIV**” is a dummy indicating whether an individual is HIV positive (from DHS HIV Tests). *The Individual Controls* are computed from DHS and include: indicators of marital status; a dummy for polygynous union; number of children; age; age squared; a dummy for living in an urban location; education in number of years; a dummy for currently working; wealth index indicators and religion indicators. *The (Ancestral) Ethnic Group Controls* are computed from the Ethnographic Atlas (EA) and include: intensity of women’s historical participation in agriculture; ancestral polygyny; ancestral bride price; ancestral plough use; ancestral pastoralism; ancestral presence of clans; indicators of ancestral settlement patterns; indicators of ancestral jurisdictional hierarchies beyond local communities; ancestral reliance on hunting; ancestral reliance on fishing; ancestral reliance on gathering; ancestral reliance on animal husbandry; ancestral reliance on agriculture; ancestral presence of large domesticated animals; indicators of intensity of ancestral agriculture; and year of observation of the ethnic group. *The Village-Geographic Controls* are computed at the village level and include: latitude; longitude; altitude; nightlight composite; population density (2010); distance to lake or coastline; distance to nearest international border; average time (minutes) required to reach a high-density urban center (2015); distance (in km) to nearest active mine; a dummy indicating whether an active mine is located within a distance of 1000 km max. of the village; malaria incidence (2010); vegetation index; indicators for the length of the growing season; index of ethnic fractionalization; and index of ethnic polarization. *Region-survey* is a subnational region defined in DHS, interacted with its survey-year. Sample in column 6 consists of female individuals. Sample in column 7 consists of male individuals. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

4.2 Robustness Checks

4.2.1 Selection Analysis

The blood test being not compulsory, selection might arise in the sample. However, the DHS program reports that the average response rate, for those who are eligible for the test, is extremely high and that a comparison between the characteristics of those who agreed to be tested and those who refused testing shows minimal bias²⁶. Moreover, it is reasonable to expect that any selection will cause a downward bias since infected individuals should be less keen to be tested²⁷. As a result, if female HIV infection is positively affected by ancestral matrilineality, more underreporting among females originating from ancestrally matrilineal ethnic groups should be expected. In order to test this hypothesis, I test the potential effect of ancestral matrilineality on the probability of underreporting HIV infection, which in practice amounts to a refusal to consent to the blood test and non take-up of it.

Further, given the persistence of mistrust in medicine in regions where colonial medical campaigns were established (Lowes and Montero (2018)), lack of consent in regions close to missions could pose a potential threat to the estimation. The inclusion of (within-country)-survey fixed effects as well as numerous village-level geographical controls allow me to rule out such potential threat.

Column 1 and column 2 of Table 1 report estimates on the probability that DHS respondent consent to take HIV test, and on the probability that she actually takes it. The estimate of *Female* \times *Matrilineal* is non significant and very close to zero in both cases, suggesting that potential selection can be ruled out (in fact, descriptive statistics show that about 95% of patrilineal females against 93% of matrilineal females consent to HIV test in my final sample; and these numbers are similar for actual take-up.)

4.2.2 Other Health Outcomes as a Falsification Test

HIV is a highly infectious, largely sexually transmitted disease. Using DHS data on level of anemia in respondents, a not sexually transmitted disease²⁸, as well as BMI and Rohrer index as other objective measures of health status²⁹, I perform falsification tests by estimating the long-term effect of ancestral matrilineality on such health outcomes³⁰. Doing so allow me to rule out the possibility that differences in HIV rates found previously hide more general differences in economic development and/or health infrastructures, and therefore differences in overall health status. Results are reported in column 3 to 5 in Table 1, and I find that, unlike with HIV, ancestral matrilineality is not associated with an increase in the prevalence of anemia, neither with Body Mass Index (BMI)³¹ or Rohrer index³² of female individuals, corroborating the hypothesis that sexual behaviour is the actual driver of my results on HIV.

²⁶See <https://dhsprogram.com/topics/HIV-Corner/hiv-prev/index.cfm>

²⁷This 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 those not tested.

²⁸Bertocchi and Dimico (2018) select anemia for their falsification test because of its relevance, given the association between anemia and malaria, another vast-scale health problem in Africa

²⁹Only available for female respondents

³⁰I measure the severity of anemia with a dummy variable taking value one when an individual is diagnosed with either mild, moderate, or severe anemia, and zero otherwise.

³¹ $mass/height^2$

³² $mass/height^3$

4.2.3 Robustness to Alternative Channels

The previous results demonstrate that ancestral matrilineality is a long-term determinant of contemporaneous prevalence of HIV among female populations in Sub-Saharan Africa. However, other long-term determinants of female HIV in Sub-Saharan Africa have also been recently highlighted in the literature. I discuss here the robustness of my main findings to these competing channels.

Common Law vs. Civil Law Countries. [Anderson \(2018\)](#) have recently highlighted the legal origins of female HIV in Sub-Saharan Africa. In particular, exploiting variation in legal origins of formerly colonized countries and the fact that common law is associated with weaker female marital property laws as compared to civil law countries, she finds higher prevalence of female HIV in common law countries, due to a lower bargaining power of women in these countries and therefore lower ability to impose safe sexual practices to their husbands. Interestingly, in her identification strategy she exploits geographical variation in common law vs. civil law countries *within ethnic group*, by including ethnic group fixed effects, and therefore rule out any ethnic effect. In some sense, my identification strategy is symmetric since I exploit ethnic norm (and related kinship organization) variation within (*within-country*) *region - survey (year)*, therefore allowing me to rule out any legal/institutional effect.

As a further robustness check, I also perform an heterogeneous analysis, estimating [Equation 1](#) on two subsamples, namely on individuals residing in common law countries versus individuals residing in civil law countries³³. The two first columns of [Table 13](#) reports the result of these estimations and provide evidence of an heterogeneity: the effect of ancestral matrilineality on female HIV holds true within common law countries only. In fact, looking at the magnitude, it seems that the average effect found in [subsection 4.1](#) is mostly driven by individuals residing in common law countries. However, this is not a surprising result since most of the variation in ancestral inheritance norm is found within common law countries³⁴.

Polygyny. [Bertocchi and Dimico \(2018\)](#) have recently highlighted contemporaneous polygyny as an other driver of female HIV in Sub-Saharan Africa, due to both females' riskier sexual behaviour triggered by the absence of the husband, and subsequent multiplicative virus transmission by the husband to his other wives. The inclusion of actual polygynous union in my individual-level controls, as well as ancestral polygyny norm in my ethnic-group controls in my regressions are meant to capture this alternative transmission channel.

Again, I also perform an heterogeneous analysis as an additional robustness check, by estimating [Equation 1](#) on the subsamples of individuals who are not currently in a polygynous union versus individuals who are currently in a polygynous union. The results are reported in column 3 and 4 of [Table 13](#) and indicate an heterogeneity in the long-term effect of ancestral matrilineality on female HIV in Sub-Saharan Africa: these effect holds true for non-polygynous individuals only. Similarly to the previous heterogeneous analysis, the magnitude of the effect found for non-polygynous individuals suggest that these latter are in fact mainly driving the average effect found in [subsection 4.1](#). However, here again, this is not a surprising result since most of the variation in ancestral inheritance norm is found among non-polygynous females³⁵.

³³Following [Anderson \(2018\)](#), I use the dataset in [La Porta et al. \(2008\)](#) in order to identify common law and civil law countries.

³⁴In my final sample, within common law countries about 29% of females originate from an ancestrally matrilineal ethnic group vs. about 5% within civil law countries.

³⁵In my final sample, about 17% of non-polygynous females originate from an ancestrally matrilineal ethnic group vs. about 7% of

Ancestral Matrilocality. Matrilocality, an ethnic-group traditional norm according to which a married couple is supposed to live with or close to the wife’s family, is highly correlated with matrilineality in Sub-Saharan Africa: in my main regression sample, about 92% of matrilocality individuals are matrilineal and about 94% of non-matrilocality individuals are non-matrilineal. Hence, ancestral matrilocality is hardly separable from ancestral matrilineality in my conceptual framework, and can be seen as an additional dimension of women’s position in ancestral matrilineal societies. For these reasons, I decide to not control for it in my preferred specifications. Nevertheless, as reported in [Table 14](#) in appendix, I perform a robustness check exercise to check that ancestral matrilocality is not driving my main results.

Several lessons can be drawn from this exercise. First, looking at column 1, my main findings are unchanged when I additionally control for ancestral matrilocality. Unsurprisingly, the adjusted r-squared remains also unchanged. Second, looking at column 2, my main findings remain valid for non-matrilocality individuals, who constitute the majority of my regression sample. All in all, these results suggest that it is very unlikely that my main findings are driven by differences between ancestrally matrilineal and patrilineal ethnic groups’ adoption of a matrilocality traditional norm.

Geographic Channels. As previously detailed, several other geographical factors have been highlighted in the literature. Among them, [Corno and De Walque \(2012\)](#) show that mine workers’ international migration is a driver of the spread of HIV. In the same vein, [Oster \(2012b\)](#) highlights exports and road networks, and subsequent increase in movements of people and sexual contacts, as an other factor of HIV infection. An other long-term determinant of female HIV in Sub-Saharan Africa put forward by [Cagé and Rueda \(2018\)](#) is the geography of Protestant and Catholic missions in the early 20th century, as well as their health investments. In the same vein, [Teso \(2019\)](#) shows that slave trade in Sub-Saharan Africa has long-term effects on contemporaneous gender norms, and [Bertocchi and Dimico \(2018\)](#) underline that slave trade was a driver of actual polygyny in Sub-Saharan Africa.

In my main specification, and in addition of (within-country) DHS survey region fixed effects which should capture most of these geographical variations, I also include a host of village-level geographical controls, including latitude, longitude, altitude, nightlight composite, population density (2010), distance to lake or coastline, distance to nearest international border, average time (minutes) required to reach a high-density urban center (2015), distance to nearest active mine, an indicator for the presence of an active mine within 1000km of the village, malaria incidence (2010), vegetation index, and indicators for the length of the growing season (see [subsection 3.3](#)). My main results being robust and almost unchanged following the inclusion of these controls (see [Table 1](#), column 5), I can reasonably rule out the possibility that such competing geographic channels could drive my main results.

Ethnic Fractionalization and Polarization. Ethno-linguistic diversity has been emphasized by [Tequame \(2012\)](#) as a driver of risky sexual behaviour and subsequent HIV. As a matter of fact, as she argues, one mechanism might be that social sanction due to risky behaviour is less costly in heterogeneous societies rather than homogeneous ones. A second mechanisms might be that information spread more easily in homogeneous than heterogeneous societies, because the former are more likely to have the same language, culture and networks. Since to be subject polygynous females.

to social sanction individuals should be detected as having risky sexual behaviour, individuals who want to keep risky sexual behaviour secret might find heterogeneous societies more favorable. In addition, it is worth noting that social interactions, which might differ by the degree of ethnical homogeneity, provide information about the level of HIV/AIDS at community level, including infectious status and risky behaviour of partners. Therefore, to account for the fact that ancestral matrilineality might be associated with different degrees of ethnical heterogeneity, which might spuriously drive my results on HIV and sexual behaviours, I control for indexes of ethnical fractionalization and polarization, which I compute at the village level using information on individual’s ethnicity, following [Montalvo and Reynal-Querol \(2005\)](#) formulas:

$$Ethnic_Fractionalization = 1 - \sum_{i=1}^N \pi_i^2 = \sum_{i=1}^N \pi_i(1 - \pi_i) \quad (2)$$

where π_i is the proportion of people who belong to the ethnic group i , and N is the number of groups. The index of ethnic fractionalization, which is an Herfindahl index, has a simple interpretation as the probability that two randomly selected individuals from a given geographic area will not belong to the same ethnic group.

$$Ethnic_Polarization = 1 - \sum_{i=1}^N \left(\frac{1/2 - \pi_i}{1/2} \right)^2 \pi_i^2 = 4 \sum_{i=1}^N \pi_i^2(1 - \pi_i) \quad (3)$$

The original purpose of this Ethnic Polarization (EP) index is to capture how far the distribution of the ethnic groups is from the $(1/2, 0, 0, \dots, 0, 1/2)$ distribution (bipolar), which represents the highest level of polarization. Ranging between 0 and 1, a higher value of the EP index indicates a higher ethnical polarization, with EP equal to 0 indicating an ethnical homogeneity, and and EP equal to 1 for two ethnic groups of the same size.

Economic Development. An other possibility is that differences in economic development may trigger differences in HIV rates. For example, legal systems and institutions are well-know driver of economic development (see [La Porta et al. \(2008\)](#) for a review). Such differences are controlled both at the country and within-country level with the inclusion of (within-country) region fixed effects in my regressions. I also control for potential differences at the village-level by including, as proxies of village-level economic development, village’s nightlight composite as well as village’s population density. I finally control for such differences at the ethnic group level by controlling for ancestral settlement patterns; ancestral jurisdictional hierarchies beyond local communities; ancestral reliance on hunting; ancestral reliance on fishing; ancestral reliance on gathering; ancestral reliance on animal husbandry; ancestral reliance on agriculture; ancestral presence of large domesticated animals; intensity of ancestral agriculture; and year of observation of the ethnic group (to alleviate the concern that some groups were portrayed later than others and might therefore have been more developed).

I again find robust and almost unchanged estimates of the long-term effect of ancestral matrilineality on female HIV once these controls are included (see column 5 in [Table 1](#)), thus alleviating the concern that differences in economic development may drive my results.

Controlling for additional observables. Despite the great emphasis put so far in controlling for numerous

alternative channels, based on observables computed at either the individual, Ethnographic Atlas (EA) ethnic groups or DHS village level, I intend here to control more directly for alternative channels discussed above, adding in my regressions covariates computed at the Murdock’s ethnic group level, and based on [Nunn \(2010\)](#), [Nunn and Wantchekon \(2011\)](#) and [Teso \(2019\)](#) datasets. The limitation of this exercise is that I cannot match DHS ethnic groups with Murdock’s ethnic groups as extensively as I did when I matched DHS ethnic groups with EA ethnic groups. Therefore, adding these new covariates will restrict my sample size. However, the value of this exercise is to assess the robustness and the stability of my estimates when explicitly controlling for additional alternative channels. In this way, I more explicitly control for the slave trade alternative channel by controlling for the logarithm of 1 plus the number of slaves taken from the individual’s Murdock ancestral ethnic group in the transatlantic and in the indian slave trade, divided by the area of land historically inhabited by the ethnic group. Then, I aim to control for individual’s ethnic group’s contact with colonizers during colonization by computing a dummy for whether a European explorer’s route traveler crossed the historical land of the ethnic group, and a dummy for whether part of the railway network built by the Europeans was on the land of the ethnic group. Further, I also control for the differential effects of the different types of religious missions ([Cagé and Rueda \(2018\)](#)) by including in my regression the number of catholic, protestant and British and Foreign Bible Society (BFBS) missions per square kilometer of an ethnic group’s land during the colonial period. Additionally, I directly control for the minimum distance of the centroid of the land historically inhabited by the ethnic group from the routes of the Saharan trade, and from the closest city of the Saharan trade. Finally, I also estimate my regressions by controlling for the number of cities with more than 20,000 inhabitants that were present in 1400 on the land inhabited by the ethnic group, and for the number of conflicts between 1400 and 1700 in the area inhabited by the ethnic group, based on [Besley and Reynal-Querol \(2014\)](#) original dataset.

In the same spirit, I also control for additional covariates, computed at the DHS village level, which likewise reduce my sample size but allow me to assess the robustness of my estimates. More specifically, using answers from male respondents in DHS³⁶, I compute village’s proportion of male circumcised as well as village’s proportion of males who report having paid for sex in the last 12 months. These two covariates are intended to control for circumcision and prostitution³⁷, which have both been emphasized in the literature as an impediment and a driver of HIV, respectively.

[Table 2](#) reports the OLS estimates of this additional robustness exercise and underlines that my OLS estimates are remarkably consistent when controlling for all of these covariates both separately and simultaneously. Thus, according to my fully-controlled OLS regression reported in column 9, I still find that females are more likely to suffer from HIV than their male counterpart. Further, I also still find that female originating from an ancestrally matrilineal ethnic group are 1.6 percentage point more likely to suffer from HIV than their patrilineal counterparts, an effect of very similar magnitude than the effect I found in my main specification, in column 5 of [Table 1](#).

³⁶The reduction in sample size when adding these controls is notably explained by the fact that only females are interviewed in some DHS villages.

³⁷Unfortunately for this study, in DHS female respondents have not been interrogated about prostitution.

Table 2: Robustness of OLS Estimates to Additional Controls

	HIV								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Matrilineality	-0.000 (0.006)	-0.001 (0.006)	0.000 (0.006)	-0.000 (0.006)	0.001 (0.006)	0.001 (0.006)	-0.007 (0.006)	-0.007 (0.006)	-0.006 (0.008)
Female	0.008*** (0.001)	0.008*** (0.001)	0.008*** (0.001)	0.008*** (0.001)	0.008*** (0.001)	0.008*** (0.001)	0.009*** (0.002)	0.008*** (0.001)	0.010*** (0.002)
Female × Matrilineality	0.016*** (0.004)	0.016*** (0.004)	0.016*** (0.004)	0.016*** (0.004)	0.016*** (0.004)	0.016*** (0.004)	0.017*** (0.005)	0.017*** (0.004)	0.016*** (0.005)
Ind. Controls	Yes								
Ethnic Group Controls	Yes								
Village-Geographic Controls	Yes								
Region-survey FE	Yes								
<i>Murdock Ethnic Groups Historical Controls:</i>									
ln(1 + Total slave exports/area)	Yes								Yes
Contact explorer route		Yes							Yes
Colonial railway		Yes							Yes
Catholic missions/area			Yes						Yes
Protestant missions/area			Yes						Yes
BFBS missions/area			Yes						Yes
Distance Saharan route				Yes					Yes
Distance Saharan node				Yes					Yes
Cities in 1400					Yes				Yes
Precolonial conflicts (1400-1700)						Yes			Yes
<i>DHS Villages Contemporaneous Controls:</i>									
Village's prop. of males circumcised							Yes		Yes
Village's prop. of males who paid for sex								Yes	Yes
Observations	168,333	168,333	168,333	168,897	168,897	168,897	115,620	136,586	108,009
Adj. R-squared	0.080	0.080	0.080	0.080	0.080	0.080	0.090	0.086	0.092
Clusters	1,498	1,498	1,498	1,508	1,508	1,508	1,088	1,190	970
Mean Dep. Var. (Patri. Males)	0.024	0.024	0.024	0.024	0.024	0.024	0.026	0.023	0.027

Notes: OLS estimates are reported with standard errors clustered at the ethnicity-DHS (within-country) region-survey (year) level in brackets. The unit of observation is an individual originating from either a traditionally matrilineal or a traditionally patrilineal ethnic group (it therefore excludes individuals originating from ethnic groups with alternate inheritance rules (ambilineality, bilinearity, duolinearity, etc.)). “HIV”, “Matrilineality”, “Female”, “Female × Matrilineality”, “Ind. Controls”, “Ethnic Group Controls”, “Village-Geographic Controls”, and “Region-survey FE” are defined in Table 1. “ln(1 + Total slave exports/area)” is the logarithm of 1 plus the number of slaves taken from the respondent’s ethnic group in the transatlantic and/or the indian slave trade divided by the area of land historically inhabited by the group. “Contact explorer route” is a dummy taking value one if an European explorer route crossed the land of the ethnic group. “Colonial railway” is a dummy taking value one if a part of the railway network built by the Europeans was on the land of the ethnic group. “Catholic missions”, “Protestant missions” and “BFBS missions” are the number of catholic, protestant, British and Foreign Bible Society (BFBS) missions per square kilometer of an ethnic group’s land during the colonial period, respectively. “Distance Saharan route” and “Distance Saharan node” are the minimum distance of the centroid of the land historically inhabited by the ethnic group from the routes of the Saharan trade and from the closest city of the Saharan trade, respectively. “Cities in 1400” is the number of cities with more than 20,000 inhabitants that were present in 1400 on the land inhabited by the ethnic group. “Precolonial conflicts (1400-1700)” is the number of conflicts between 1400 and 1700 in the area inhabited by the ethnic group. “Village’s prop. of males circumcised” is the within DHS village’s proportion of male respondents circumcised. “Village’s prop. of males who paid for sex” is the within DHS village’s proportion of male respondents who report having paid for sex in the last 12 months. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

4.2.4 Assessing Selection on Unobservables

Despite my attempts to control for a large set of observable factors, at both the individual, ethnic group and village level, the estimates reported in [Table 11](#) may still be biased by unobservable factors correlated with both ancestral matrilineality and females contemporaneous sexual and contraceptive behaviours, and subsequently HIV. A priori, the direction of the potential omitted variable bias is not clear. Consider for instance the possibility that ethnic groups where women historically benefited from more sexual freedom adopted matrilineality as a resulting adaptative kinship organization. As such, some evolutionary anthropologists explain the existence of matrilineal societies as the result of an evolutionary process that created institutions suitable for the ecological and social environment³⁸. In particular, matrilineality may be advantageous in environments with low paternal certainty since while it is difficult to confirm paternity, maternity is easily observable. Thus, an inheritance system in which property passes from the mother's brother to her sons may be optimal since the brother knows he is related to his sister, but cannot verify that he is related to his children ([Fortunato \(2012\)](#)). This could drive the OLS estimates away from zero. On the other side, mating process might be inherently more assortative in matrilineal societies, relative to patrilineal ones, yielding matrilineal females suffering from relatively lower marital dissatisfaction. If that was the case, this might lead to matrilineal females being relatively less likely to adopt risky sexual behaviours³⁹, and therefore drive the OLS estimates towards zero. In this subsection I consequently assess the likelihood that the OLS estimates might be biased by unobservables.

Coefficients Ratio Tests ([Altonji et al. \(2005\)](#)). I start by assessing the sensitivity of the OLS estimates to controlling for observable characteristics. To do so, I first employ the strategy adapted by [Nunn and Wantchekon \(2011\)](#) from [Altonji et al. \(2005\)](#) that allows to determine how much stronger selection on unobservables would have to be compared to selection on observables in order to fully explain away my results. To perform this test, I calculate the ratio $\hat{\beta}_F / (\hat{\beta}_R - \hat{\beta}_F)$, where $\hat{\beta}_F$ is my coefficient of interest from a regression that includes my full set of controls, while $\hat{\beta}_R$ is my coefficient of interest from a regression that includes a restricted set of controls. The intuition behind the formula is straightforward. First, consider why the ratio is decreasing in $(\hat{\beta}_R - \hat{\beta}_F)$. The smaller is the difference between $\hat{\beta}_R$ and $\hat{\beta}_F$, the less the estimate is affected by selection on observables, and the stronger selection on unobservables needs to be (relative to observables) to explain away the entire effect. Next, consider the intuition behind $\hat{\beta}_F$ in the numerator. The larger $\hat{\beta}_F$, the greater is the effect that needs to be explained away by selection on unobservables, and therefore the higher is the ratio.

The results are reported in columns 1 of [Table 15](#), where each row reports result for different set of restricted covariates. This yields four ratios that range from -509.27 to 167.42. In some cases, the coefficient in the fully-controlled model is larger than that on the uncontrolled model, giving a negative ratio. In general, these ratios are far greater than 1 in absolute value, and therefore suggest that the influence of unobservable characteristics would have to be far greater than the influence of observable characteristics to fully account for my OLS findings.

Minimum Coefficient Lower Bound ([Oster \(2017\)](#)). Further, I also use the method from [Oster \(2017\)](#) to

³⁸[Lowe \(2018a\)](#) appendix provides an extensive overview of the hypothesized origins of matrilineal kinship systems.

³⁹In fact [Bertocchi and Dimico \(2018\)](#) show that, due to marital dissatisfaction, women in polygamous union in Sub-Saharan Africa adopt riskier sexual behaviours which are more conducive to HIV.

calculate a bias-adjusted lower bound of my coefficient of interest. Oster shows that if one assumes that observables and unobservables have the same explanatory power in the outcome variable, then the following is a consistent estimator: $\beta^* = \hat{\beta}_F - (\hat{\beta}_R - \hat{\beta}_F) \times ((R_{Max}^2 - \hat{R}_F^2)/(\hat{R}_F^2 - \hat{R}_R^2))$, where $\hat{\beta}_R$ and $\hat{\beta}_F$ are defined as above; \hat{R}_F^2 is the R-squared from the fully-controlled regression; \hat{R}_R^2 is the R-squared from the restricted regression; and R_{Max}^2 is the R-squared from a regression that includes all observable and unobservable controls. Although in theory the maximum possible value of R_{Max}^2 is one, as underlined by [González and Miguel \(2015\)](#), in the real world, where there is significant measurement error, the value of R_{Max}^2 should be much lower than one. In fact, by definition, $R_{Max}^2 \in [\hat{R}_F^2; 1]$. [Oster \(2017\)](#) provides some insights on how R_{Max}^2 should be chosen, showing that $R_{Max}^2 = 1$ may lead to over-adjustment in many cases. I follow her procedure and present bias-adjusted lower bound coefficients for $R_{Max}^2 = \min(\mathbf{1.3}\hat{R}_F^2; 1)$ ⁴⁰, $R_{Max}^2 = \min(\mathbf{1.5}\hat{R}_F^2; 1)$, and $R_{Max}^2 = \min(\mathbf{2}\hat{R}_F^2; 1)$ in column 2, 3 and 4 respectively (I also report bias-adjusted lower bound coefficients for $R_{Max}^2 = 1$ in column 5 for informational purpose).

All bias-adjusted lower bound estimates from this exercise remain positive and, taken at face value, still imply a sizeable estimated effect of ancestral matrilineality on female HIV, of same order of magnitude that previously found in my OLS regressions. Further, it is worth noting that the full set of these biased-adjusted lower bound estimates (column 2 to 5) falls within the 99.5% confidence interval of my fully-controlled OLS estimate (column 7), which suggests that the size of the estimate from the OLS regression with full controls is similar to the bias-adjusted estimates. All in all, these tests suggest that my fully-controlled OLS estimates are very unlikely to be affected by omitted variable bias, and therefore support a causal interpretation of my OLS findings.

4.3 Alternative Identification Strategies

Though my OLS estimates are very unlikely to be affected by omitted variable bias as discussed in [subsubsection 4.2.4](#), reverse causality may remain a concern if, for example, matrilineality had been an adaptative response in environments with low paternal certainty since while it is difficult to confirm paternity, maternity is easily observable. In order to provide further support to the causal interpretation of my main individual-level OLS findings, I implement three alternative identification strategies, all of them having the common feature of allowing me to estimate the effect of the variation in cultural trait within similar environments. To do so, I exploit data on the GPS location of DHS villages as well as the digitized Murdock’s map of ancestral ethnic groups in Africa.

4.3.1 Instrumenting for Ancestral Matrilineality

As a first alternative identification strategy, I perform an IV strategy, instrumenting individual’s ethnic group’s ancestral matrilineality (*Matrilineality* regressor) with a measure of the minimum distance (in km) between the individual’s DHS village to the nearest ancestral matrilineal ethnic boundary (based on digitized Murdock’s map of ancestral ethnic groups in Africa, see [Figure 2](#)); and I instrument my main regressor of interest $Female \times Matrilineality$ with a variable interacting *Female* dummy with this measure of distance⁴¹.

⁴⁰[Oster \(2017\)](#) suggests using $R_{Max}^2 = 1.3\hat{R}_F^2$ as a cutoff to test for the stability of a treatment effect consistent with randomized treatment.

⁴¹I assign positive values to this measure of distance for DHS villages located within the boundaries of a matrilineal ancestral ethnic groups, while I assign negative values for villages located within the boundaries of a non-matrilineal (i.e. patrilineal or other) ancestral ethnic group. Therefore, a distance of +100 means that a village is located within an ancestrally matrilineal area, 100 km away from the nearest matrilineal/non-matrilineal ancestral ethnic boundary; while a distance of -100 means that a village is located within an ancestrally non-matrilineal area (i.e. patrilineal or other), 100 km away from the nearest matrilineal/non-matrilineal ancestral ethnic

The relevance of this instrument is *a priori* straightforward. The critical issue is whether the distance between an individual’s location and ancestral matrilineal ethnic boundary is uncorrelated with factors, other than individual’s ethnic group’s ancestral matrilineality, that may have affected individual’s sexual and contraceptive behaviour and therefore HIV susceptibility. Therefore, there remain a number of other reasons why the exclusion restriction may not be satisfied. First, distance between contemporaneous individual’s location and ancestral matrilineal ethnic boundary may be correlated with geographic characteristics (e.g. vegetation, altitude, remoteness, etc.) which might affect the spread of HIV virus. Second, this contemporaneous measure of distance may also be correlated with the geographical distribution of economic activities, infrastructures, transportation networks, etc. and therefore reflect migration patterns that appeared in the meantime (i.e. since pre-colonial period) and which might affect contemporaneous HIV rates. Third, this measure of distance may also have been affected by european contact, colonization, religious missions and past conflicts, which may notably have led to population displacements and different health infrastructure investments. Therefore, in order to alleviate such concerns I include in my IV regressions region (within-country) \times survey (time) fixed effects, and I directly control for such alternative channels with a large array of ethnic group’s historical controls and village-level geographic controls, which are already detailed in [subsection 3.3](#) and [subsubsection 4.2.3](#).

[Table 3](#) presents estimates of my instrumental variable strategy. The OLS estimates of regressing my measure of distance on individual’s ethnic group’s ancestral matrilineality, as well as F-Stats of the test that the coefficient of the instrument is null, reported in column 3 and 4, suggest that my instrument is a strong predictor of individual’s ethnic group’s ancestral matrilineality. Similar conclusions can be drawn from the OLS estimates and F-Stats of regressing *Female \times Matrilineality* on the variable interacting the *Female* dummy with my instrument, reported in columns 5 and 6. More importantly, the IV-2SLS estimates reported in columns 1 and 2 confirm my OLS results and indicate that women originating from ancestrally matrilineal ethnic groups significantly suffer from relatively more HIV today. Remarkably, my IV estimates are close to my OLS estimates once Murdock’s ethnic groups historical controls are included (column 2). Noteworthy, instrumenting two potentially endogenous regressors with two instruments and assuming clustered standard errors, I follow [Andrews \(2018\)](#) and compute weak instrument-robust 95% confidence intervals of my IV estimates.⁴² These latter are reported in square brackets and also suggest non-null positive IV estimates of *Female \times Matrilineality*. Importantly, calculating 95% two-step weak instrument-robust CI based on [Andrews \(2018\)](#) and [Sun \(2018\)](#), I find a 5% distortion cutoff, which suggests that my instruments are strong instruments, in the sense that size distortions are below 5% for conventional 95% confidence intervals.

4.3.2 Accounting for Unobservables: Geographic RD Estimates

As a second alternative identification strategy, I undertake a geographic regression discontinuity analysis. More precisely, I examine and compare individuals living in villages geographically close, but where some villages are located within the ancestral boundaries of an ancestrally matrilineal ethnic group and the other within the ancestral boundaries of an ancestrally non-matrilineal (i.e. patrilineal or other) ethnic group. In this framework, the ancestral matrilineal ethnic boundary is the delineation created by the ancestral borders of ethnic groups that practiced

boundary.

⁴²I use “*twostepweakiv*”, a Stata package developed by and presented in [Sun \(2018\)](#). I compute weak instrument-robust 95% confidence intervals, projected on *Female \times Matrilineality* regressor.

Table 3: IV Estimates of the Effect of Ancestral Matrilineality on Female HIV

	(1)	(2)	(3)	(4)	(5)	(6)
	IV-2SLS		OLS (First Stages)			
	HIV		Matrilineality		Female \times Matri.	
Matrilineality	-0.010	0.058				
	(0.037)	(0.064)				
Female	0.006***	0.007***	Yes	Yes	Yes	Yes
	(0.001)	(0.001)				
Female \times Matrilineality	0.029***	0.023***				
	(0.007)	(0.006)				
	[0.018;0.040]	[0.011;0.037]				
Distance to matri. boundary			0.0003***	0.0002***		
			(0.0000)	(0.0000)		
Female \times Distance to matri. boundary					0.0008***	0.0010***
					(0.0001)	(0.0001)
Ind. Controls	Yes	Yes	Yes	Yes	Yes	Yes
Ethnic Group Controls	Yes	Yes	Yes	Yes	Yes	Yes
Village-Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes
Region-survey FE	Yes	Yes	Yes	Yes	Yes	Yes
Murdock Ethnic Group Historical Cont.		Yes		Yes		Yes
Observations	179,312	165,365	179,312	165,365	179,312	165,365
Adj. R-squared	0.032	0.032	0.835	0.885	0.617	0.672
Clusters	1,671	1,479	1,671	1,479	1,671	1,479
Mean Dep. Var.	0.023	0.023	0.158	0.169	0.085	0.091
F-Stat (test coeff. = 0)			19.24	13.96	63.57	77.54

Notes: IV-2SLS estimates are reported in column 1 and 2; first stage OLS estimates are reported in column 3 to 6; with standard errors clustered at the ethnicity-DHS (within-country) region-survey (year) level in brackets. Weak instrument-robust 95% confidence intervals (Andrews (2018), Sun (2018)) are reported in square brackets (calculating 95% two-step weak instrument-robust CI, the distortion cutoff is 5%, indicating strong instruments). The unit of observation is an individual originating from either a traditionally matrilineal or a traditionally patrilineal ethnic group (it therefore excludes individuals originating from ethnic groups with alternate inheritance rules (ambilineality, bilinearity, duolinearity, etc.)). “Distance to matri. boundary” is computed on QGIS as the minimum distance between DHS village and the nearest ancestral matrilineal ethnic boundary (based on Murdock’s map of ancestral ethnic group boundaries, see Figure 2). This takes negative values for DHS villages located within boundaries of non-matrilineal ancestral ethnic groups. “HIV”, “Matrilineality”, “Female”, “Female \times Matrilineality”, “Ind. Controls”, “Ethnic Group Controls”, “Village-Geographic Controls”, and “Region-survey FE” are defined in Table 1. Murdock Ethnic Groups Historical Controls are defined in Table 2. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

matrilineal descent alongside groups that practiced patrilineal or alternate descent (based on digitized Murdock’s map of ancestral ethnic groups in Africa, see Figure 2). The intuition behind this specification is that the ancestral matrilineal ethnic boundary is determined by the ancestral borders of multiple matrilineal and non-matrilineal ethnic groups. The boundaries between these multiple ethnic groups are arbitrary, and along the border these areas are quite similar⁴³.

Therefore, my strategy is to use a regression discontinuity (RD) estimation method that restricts the sample to individuals living in villages that are sufficiently close to the ancestral matrilineal ethnic boundary and estimate the causal effect of living in a village located on the matrilineal side of the ancestral ethnic boundary on female HIV, using the estimated difference in female HIV at the ancestral matrilineal ethnic boundary. Importantly, I do not estimate the effect of originating from an ancestrally matrilineal ethnic group anymore, but the effect of living in a village located in an ancestrally matrilineal geographic area. The benefit of this strategy is that it accounts for

⁴³Note that the ancestral matrilineal ethnic boundary does not coincide with any actual border.

unobservable factors that vary smoothly across space. Therefore, as long as the determinants of unobservable traits (e.g. geography, history, idiosyncratic shocks, state presence etc.) vary smoothly, the unobservable traits will be accounted for by the RD strategy. Further, in order to get the more conservative estimates, I directly control for the large array of individual, ethnic group’s historical and village-level geographic control, as well as (within-country) region \times survey (year) fixed effects described in the previous sections. More specifically, my RD specification takes the following form:

$$y_{ievert} = \alpha + \beta_1 \text{Matrilineality}_v + \beta_2 \text{Female}_i + \beta_3 \text{Female}_i \times \text{Matrilineality}_v + \mathbf{X}'_{ievert} \Delta + \mathbf{X}'_{ert} \Omega + \mathbf{X}'_{vrt} \Pi + f(\text{location}_v) + \lambda_{rt} + \varepsilon_{ievert} \quad (4)$$

with y_{ievert} denoting an individual-level outcome (indicator for whether the individual is HIV positive) for individual i from ethnic group e , living in village v in within-country DHS region r , and surveyed at year t . Matrilineality_v is an indicator for whether an individual lives in a village located on an ancestrally matrilineal geographic area; Female_i is an indicator for whether an individual is a female. β_1 is intended to capture a “Matrilineal effect”; and β_2 is intended to capture a “Gender effect” on HIV. The coefficient of interest β_3 captures the effect of living in a village located in an ancestrally matrilineal geographic area on female HIV once the “Matrilineal effect” and the “Gender effect” have been controlled for. $f(\text{location}_v)$ denotes a RD polynomial that controls for a smooth function of the geographic location of DHS villages. In my specifications I alternatively use the minimum distance to the nearest ancestral matrilineal ethnic boundary (in km) or the gps coordinates (latitude and longitude) of the village as running variables. Further, I use several functional forms of the polynomial, using polynomials of different orders, and alternatively estimating them separately on each side of the boundary (“flexible polynomials”). \mathbf{X}'_{ievert} ; \mathbf{X}'_{ert} ; and \mathbf{X}'_{vrt} represent a set of individual-level covariates, ethnic group-level ancestral covariates, and village-level geographic covariates respectively, which are defined in [subsection 3.3](#); λ_{rt} denotes (within-country) DHS region-survey (year) fixed effects. Standard errors ε_{ievert} are clustered at the ethnic group-(within-country) DHS region-survey (year) level, and the sample is restricted to individuals living in villages that are within a certain distance of the ancestral matrilineal ethnic boundary, either 100, 150, or 200 kilometers.

Validating the assignment of matrilineal individuals. The boundaries used for my RD estimates are from Murdock (1959, see [Figure 2](#)), a source that has been used previously in a number of studies that use a similar RD approach (see [Moscona et al. \(2019\)](#) for a recent and related example). However, an important assumption when using the ethnic boundaries is that they accurately reflect true discontinuities (i.e., boundaries) of ethnic affiliation today. This is particularly important since, in reality, one may not observe clear borders between ethnic groups, and instead only a gradual change of the mix of ethnicities over space. Further, an additional assumption that I make is that matching between individual’s self-reported ethnicity in DHS and ancestral ethnic group’s in *Ethnographic Atlas* is accurate⁴⁴. Therefore, I now check the validity of my use of Murdock’s ethnic boundaries by examining how individual’s ancestral matrilineality varies at ancestral matrilineal ethnic boundaries. This is shown in [Figure 8](#)

⁴⁴This is important since individual’s ethnic group’s ancestral matrilineality variable is based on kinship organization of ancestral ethnic groups reported in the *Ethnographic Atlas*.

in appendix, which reports the bivariate relationship between distance from the ancestral matrilineal ethnic border and individual’s ethnic group’s ancestral matrilineality. The y -axis displays the fraction of the population in a 5 km bin that reports that they are a member of an ancestrally matrilineal ethnic group, and the x -axis is distance in kilometers from the border, with a positive distance indicating a location within the ancestral territory of ancestrally matrilineal ethnic group and a negative distance indicating a location within the ancestral territory of ancestrally non-matrilineal (i.e. patrilineal or other) ethnic group. Reassuringly, I find that there is a discontinuous change in the fraction of the population that report that they are members of an ancestrally matrilineal ethnic group at the borders⁴⁵.

Geographic RD Estimates. Before turning to my estimates I first examine the raw data for the RD sample. Figure 9 in appendix shows a bin scatterplots of the predicted HIV rate for females living in villages located within 150 km of the ancestral matrilineal ethnic border, using a flexible third-order RD polynomial conditioned on region (within-country) \times survey (time) FE and estimated separately on each side of the border. Positive values, on the horizontal axis, reflect 5 km bins in ancestrally matrilineal geographic areas and negative values reflect 5 km bins in ancestrally non-matrilineal geographic areas. Even in the raw data, a discontinuity at the border is apparent: a discontinuous increase in female HIV rate on the matrilineal side of the border can be observed. I next turn to my more formal RD estimates.

Table 4 reports the geographic RD estimates for different bandwidths: 100 km (columns 1 and 2), 150 km (columns 3 and 4), and 200 km (columns 5 and 6); different running variables: minimum distance in km between DHS village and nearest ancestral matrilineal ethnic boundary (columns 1, 3 and 5), as well as village’s latitude and longitude (columns 2, 4 and 6); and different polynomial specifications: linear polynomial, flexible linear polynomial, quadratic polynomial, flexible quadratic polynomial, cubic polynomial and flexible cubic polynomial (with “flexible” standing for polynomials estimated separately at each side of the boundary). Several lessons can be drawn from these estimates. First, I find a remarkably consistent and significantly positive “Gender effect”, consistent with the “feminization of HIV” in Sub-Saharan Africa, already extensively discussed in the literature and in section 4. Second, and more importantly, I also find a remarkably consistent and significantly positive estimate of $Female \times Matrilineality$, which corroborate my findings so far that females originating from ancestrally matrilineal ethnic groups suffer from significantly more HIV today, relative to their patrilineal counterparts.

Noteworthy, the magnitudes of my geographic RD estimates of $Female \times Matrilineality$ are lower than the magnitudes of my OLS estimates. A first explanation stems from the fact that in this analysis $Matrilineality$ indicates whether an individual lives in a village located within the ancestral boundaries of an ancestrally matrilineal ethnic group, instead of an individual’s ethnic group’s ancestral matrilineality. In addition, although one other explanation for this is a potential bias from unobservables present in my OLS estimates (which is unlikely according to the tests I perform in subsection 4.2.3; and which would nevertheless go in opposite direction than suggested by my IV-2SLS estimates), the difference might also be explained by the fact that within an ancestral matrilineal ethnic group’s territory, and close to the border, only a fraction of the population is today likely to belong to ancestrally

⁴⁵It is however important to note that, while information on individual’s ethnic group’s ancestral matrilineality is based on the matching between individual’s self-reported ethnicity in DHS and ancestral ethnic groups in Ethnographic Atlas, these latter slightly differ from ethnic groups classification in Murdock’s Map of ancestral ethnic groups (1959). This might partially explain why the discontinuous increase very close to the boundary is of limited size.

matrilinal ethnic groups. As shown in [Figure 8](#), close to the border on the matrilineal side approximately 60% to 40% of the population does not belong an ancestrally matrilineal ethnic group. This suggests that the magnitude of the RD estimates could be biased downwards by this amount as well. In addition, ancestral ethnic groups boundaries are susceptibly blurry, and spillovers from matrilineal group to patrilineal group may arise close to the boundary, especially when it comes to sexual relationships and sexually transmitted diseases. Higher ethnic fractionalization and/or polarization at the boundary may help explain such spillovers since, as underlined by [Tequame \(2012\)](#), riskier sexual behaviours are more easy to conceal in fragmented societies. Alternatively, migration at the boundary could also increase sexual mixing between individuals originating from ancestrally matrilineal ethnic groups and those originating from ancestrally patrilineal ethnic groups. All in all, such spillovers will cause estimated effect at the border to be muted⁴⁶.

4.3.3 Nearest Neighbor Matching

As a final alternative identification strategy to individual-level OLS regressions, I use nearest neighbor matching to compare each DHS village located in an ancestrally matrilineal area to the DHS village located in an ancestrally non-matrilineal (i.e. patrilineal or other) area⁴⁷ that is the most similar in terms of geographic characteristics. The matching average treatment over treated effects are reported in [Table 5](#), with village pairs being matched⁴⁸ on the full range of village geographic controls of [Equation 1](#). As reported in column 1, villages located in ancestrally matrilineal areas are characterized by average female HIV rates that are higher by 2.9 percentage points to their nearest neighbor village located in an ancestrally non-matrilineal area, an effect significant at 1% and of large magnitude, representing about 78% of the average village’s female HIV rate of villages located in an ancestrally patrilineal area. Interestingly, no such effect is found on village’s proportion of HIV positives males⁴⁹, consistent with my hypothesis that the effect of ancestral matrilineality on contemporaneous HIV rates is specific to female populations.

5 Mechanisms

The previous section has uncovered a robust relationship between ancestral matrilineality and the contemporaneous spread of the HIV epidemic among female individuals in Sub-Saharan Africa. According to my conceptual framework, the legacy of ancestral matrilineality on women’s sexual preferences and ability to implement them may have led matrilineal women to adopt sexual and contraceptive behaviours that favor a higher rate of transmission of HIV today. In this section, I turn to a more direct investigation of these channels, by exploring the empirical relationship between ancestral matrilineality and female’s sexual autonomy, sexual behaviour and contraception use.

⁴⁶The estimated effect of *Female × Matrilineality* becomes non-significant when restricting the sample to individuals living within 50 km of the boundary (not reported). While the large reduction in sample size may explain why the effect is less precisely estimated, the spillover effects may also be strongly at play close to the boundary.

⁴⁷Based on Murdock’s map of ancestral ethnic groups (see [Figure 2](#)).

⁴⁸I use nearest neighbor matching based on Mahalanobis distance. Estimates are corrected for bias due to matching on multiple continuous variables, based on [Abadie and Imbens \(2006\)](#) and [Abadie and Imbens \(2011\)](#).

⁴⁹The smaller sample size is explained by the fact that in some DHS villages only females were interviewed.

Table 4: Geographic RD Estimates

	(1)	(2)	(3)	(4)	(5)	(6)
	Dep. Var.: HIV					
Distance from matrilineal boundary:	100km		150km		200km	
<i>Running variable:</i>	<i>Distance</i>	<i>Lat./Long.</i>	<i>Distance</i>	<i>Lat./Long.</i>	<i>Distance</i>	<i>Lat./Long.</i>
<i>Panel A: Linear Polynomial</i>						
Matrilineality	-0.000 (0.004)	0.000 (0.004)	-0.001 (0.004)	-0.001 (0.004)	-0.002 (0.004)	-0.001 (0.004)
Female	0.011*** (0.002)	0.011*** (0.002)	0.010*** (0.002)	0.010*** (0.002)	0.011*** (0.002)	0.011*** (0.002)
Female × Matrilineality	0.008* (0.004)	0.008* (0.004)	0.009** (0.003)	0.009** (0.003)	0.009** (0.003)	0.009** (0.003)
<i>Panel B: Flexible Linear Polynomial</i>						
Matrilineality	0.005 (0.005)	0.002 (0.006)	0.000 (0.004)	0.001 (0.006)	-0.001 (0.004)	0.000 (0.006)
Female	0.011*** (0.002)	0.011*** (0.002)	0.010*** (0.002)	0.010*** (0.002)	0.011*** (0.002)	0.011*** (0.002)
Female × Matrilineality	0.008* (0.004)	0.008* (0.004)	0.009** (0.003)	0.009** (0.003)	0.009** (0.003)	0.009** (0.003)
<i>Panel C: Quadratic Polynomial</i>						
Matrilineality	0.000 (0.004)	0.001 (0.004)	-0.001 (0.004)	0.000 (0.004)	-0.001 (0.004)	-0.000 (0.004)
Female	0.011*** (0.002)	0.011*** (0.002)	0.010*** (0.002)	0.010*** (0.002)	0.011*** (0.002)	0.011*** (0.002)
Female × Matrilineality	0.008* (0.004)	0.008* (0.004)	0.009** (0.003)	0.009** (0.003)	0.009** (0.003)	0.009** (0.003)
<i>Panel D: Flexible Quadratic Polynomial</i>						
Matrilineality	0.005 (0.008)	0.001 (0.008)	0.005 (0.007)	0.001 (0.008)	0.001 (0.006)	0.001 (0.008)
Female	0.011*** (0.002)	0.011*** (0.002)	0.010*** (0.002)	0.010*** (0.002)	0.011*** (0.002)	0.011*** (0.002)
Female × Matrilineality	0.008* (0.004)	0.008* (0.004)	0.009** (0.003)	0.009*** (0.003)	0.009** (0.003)	0.009** (0.003)
<i>Panel E: Cubic Polynomial</i>						
Matrilineality	-0.000 (0.004)	0.001 (0.004)	-0.001 (0.004)	0.000 (0.004)	-0.001 (0.004)	-0.000 (0.004)
Female	0.011*** (0.002)	0.011*** (0.002)	0.010*** (0.002)	0.010*** (0.002)	0.011*** (0.002)	0.011*** (0.002)
Female × Matrilineality	0.008* (0.004)	0.008* (0.004)	0.009** (0.003)	0.009** (0.003)	0.009** (0.003)	0.009** (0.003)
<i>Panel F: Flexible Cubic Polynomial</i>						
Matrilineality	0.007 (0.010)	-0.005 (0.010)	0.000 (0.008)	-0.007 (0.010)	0.005 (0.008)	-0.008 (0.010)
Female	0.011*** (0.002)	0.011*** (0.002)	0.010*** (0.002)	0.010*** (0.002)	0.011*** (0.002)	0.011*** (0.002)
Female × Matrilineality	0.008* (0.004)	0.008* (0.004)	0.009** (0.003)	0.009** (0.003)	0.009** (0.003)	0.009** (0.003)
Baseline Controls + FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	48,947	48,947	71,333	71,333	89,982	89,982
Clusters	631	631	840	840	1,015	1,015
Mean Dep. Var.	0.034	0.034	0.033	0.033	0.032	0.032

Notes: Geographic RD estimates based on Equation 4 are reported with standard errors clustered at the ethnicity-DHS (within-country) region-survey (year) level in brackets, for different bandwidths, running variables and polynomial specifications. The unit of observation is an individual originating from either a traditionally matrilineal or a traditionally patrilineal ethnic group (it therefore excludes individuals originating from ethnic groups with alternate inheritance rules (ambilineality, bilinearity, duolinearity, etc.)). “Matrilineality” is an indicator variable for whether the individual lives in a village located within the ancestral boundaries of a matrilineal ethnic group (based on Murdock’s map of ancestral ethnic groups, see Figure 2). “HIV” and “Female” are defined as in Table 1. “Baseline Controls + FE” are the “Ind. Controls”, “Ethnic Group Controls”, “Village-Geographic Controls”, and “Region-survey FE” defined in Table 1. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Nearest Neighbor Matching (ATT)

	(1)	(2)
	Village's proportion of HIV positive females	Village's proportion of HIV positive males
Ancestrally matrilineal area	0.029*** (0.011)	0.007 (0.011)
Observations	13,176	12,710
Mean Dep. Var. (Villages in patri. area)	0.037	0.023

Notes: This table reports the average treatment effect on the treated (ATT) on the proportion of village's HIV positive females in column 1, and males in column 2, between DHS villages located on an ancestrally matrilineal area and DHS villages located on an ancestrally non-matrilineal area (patrilineal or other), using nearest neighbor matching, where villages are matched using the Mahalanobis distance function based on all *Village-Geographic Controls* detailed in Table 1. Estimates are corrected for bias due to matching on multiple continuous variables (Abadie and Imbens (2006); Abadie and Imbens (2011)). Abadie and Imbens robust standard errors are reported in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

5.1 Female Sexual Autonomy

First, I explore whether women originating from ancestrally matrilineal ethnic groups benefit from a higher social status, and consequent higher bargaining power, which would allow them to benefit from a higher sexual autonomy. First, building on the household economics literature highlighting the role of marriage outside option on intrahousehold bargaining power (Baland and Ziparo (2017); Bargain et al. (2019)), I create a dummy equals to one if an individual is currently divorced. Then, following Anderson (2018), I use information provided in DHS on land and house ownership, as a measure of female property rights in case of divorce, (restricting my sample to divorced women), and I create a dummy which is equal to one if a female reports owning a house and/or a land when divorced. Finally, I also construct a dummy equals to one if there is at least one wife's parent in the household⁵⁰. Working with these outcomes allow me to check whether women originating from ancestrally matrilineal ethnic groups benefits from better marriage outside options and, consequently, higher bargaining power within marriage.

To further explore how this translates into female sexual autonomy, I emphasize a second set of outcomes which pertain to women's decision-making regarding contraception. In particular, I use answer to the question asked to women in union currently using contraception, on who decide on the use of contraception, and I create a dummy equals to one if a woman reports being decision-maker. Other questions in DHS directly cover sex negotiations and autonomy and asks whether a woman could get a male condom herself, whether a wife is justified to ask husband to use condom if he has a STI (Sexually Transmitted Infection); and whether it is justified for a women to refuse sex with her husband if he has another women. For each of these questions, I create a dummy equals to 1 if a woman answers affirmatively.

For the analysis in this subsection, I focus on the subsample of female individuals and estimate the following equation:

⁵⁰Bargain et al. (2019) show the influence of the presence of wife's family in the household on wife's intrahousehold decision-making, in the Indonesian context.

$$y_{ieVRT} = \alpha + \beta_1 \text{Matrilineality}_e + \mathbf{X}'_{ieVRT} \Delta + \mathbf{X}'_{ert} \Omega + \mathbf{X}'_{VRT} \Pi + \lambda_{rt} + \varepsilon_{ieVRT} \quad (5)$$

where \mathbf{X}'_{ieVRT} , \mathbf{X}'_{ert} , \mathbf{X}'_{VRT} , λ_{rt} and ε_{ieVRT} are individual-level, ethnic group-level and village-level controls and (within-country) DHS region-survey (year) fixed effects, defined as in [subsection 3.3](#). Standard errors are clustered at the ethnic group-(within-country) DHS region-survey (year) level. My coefficient of interest is now β_1 and captures the long-term effect of ancestral matrilineality on female populations⁵¹.

In [Table 6](#), I report results from estimating [Equation 5](#) on these outcomes. I find that ancestral matrilineality is indeed significantly positively associated with all these dimensions of female’s social status, bargaining power and sexual autonomy. More specifically, according to the estimates reported in column 1, matrilineal women are 1.5 percentage points significantly more likely to be divorced, meaning twice more likely to be divorced than patrilineal females⁵². Also, according to the estimate reported in column 2, I find that divorced women originating from ancestrally matrilineal ethnic group are 23.5 percentage points (representing about 65% of the mean probability of the patrilineal females in my sample) more likely to own a house and/or a land, as compared to their divorced patrilineal counterparts. I additionally find in column 3 that matrilineal females are 0.08% more likely to have at least one parent present in their household, an effect of large magnitude given that only 2% of patrilineal females benefit from the presence of at least one of their parent in their household. This result is consistent with the high correlation underlined in [subsection 4.2.3](#) between ancestral matrilineality and ancestral matrilocality. Further, estimates reported in column 3 to 6 indicates, in the same order, that matrilineal women are 4.3 percentage points more likely to be decision-maker regarding the use of contraception; 2.6 percentage points more likely to answer that they could get (male) condom themselves; 3 percentage points more likely to find justified for a woman to ask her husband to use a condom if he has a STI; and 2.8 percentage points more likely to find justified for a woman to refuse sex with her husband if he has another women.

All in all, in line with my conceptual framework, when exploring the effect of ancestral matrilineality on female HIV, I can rule out [Anderson \(2018\)](#) mechanism, according to which HIV prevalence should be higher for less empowered females since these latter should be less able to impose safe sexual practices to their husbands. I propose two main other mechanisms in the two next subsections. More precisely, I explore whether the higher sexual autonomy of women originating from ancestrally matrilineal ethnic groups translates into them implementing their preferred sexual strategy which have been shaped in the long-term by past reproductive success, and thereby translates into them adopting riskier sexual and contraception use behaviours.

⁵¹Remember that I restrict my sample to individuals originating from an either ancestrally matrilineal or an ancestrally patrilineal ethnic group. Therefore, $\text{Matrilineality} = 0$ means that the individual originates from an ancestrally patrilineal ethnic group (comparison group).

⁵²The mean proportion of divorced patrilineal females in my regression sample is 1.5%.

Table 6: **Ancestral Matrilineality and Female Sexual Autonomy (OLS)**

	Divorced	Own house and/or land (divorced)	Wife's Parents in HH	Decide Contraception	Can get condom	Wife justified ask condom if husband has STI	Wife justified refuse sex if husband has other women
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Matrilineality	0.015*** (0.003)	0.235* (0.129)	0.008** (0.004)	0.043** (0.018)	0.026** (0.013)	0.030*** (0.011)	0.028** (0.013)
Ind. Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ethnic Group Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Village-Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region-survey FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	105,964	1,191	66,902	13,709	56,270	91,420	95,604
Adj. R-squared	0.041	0.316	0.026	0.054	0.123	0.157	0.139
Clusters	1,550	238	1,407	981	1,392	1,392	1,392
Mean Dep. Var. (Patri.)	0.015	0.359	0.020	0.871	0.632	0.772	0.622

Notes: OLS estimates are reported with standard errors clustered at the ethnicity-DHS (within-country) region-survey (year) level in brackets. The unit of observation is a woman originating from either a traditionally matrilineal or a traditionally patrilineal ethnic group. “Matrilineality” indicates (dummy) whether an individual belongs to a traditionally matrilineal ethnic group. “Divorced” is a dummy indicating whether an individual is currently divorced. “Own house and/or land (divorced)” is a dummy indicating whether an individual owns a house and/or a land (divorced females sample). “Decide contraception” is a dummy indicating whether an individual is decision-maker regarding contraception. “Can get condom” is a dummy indicating whether an individual can get herself a male condom. “Wife justified ask condom is husband has STI” is a dummy indicating whether an individual find justified for a wife to ask her husband to use a condom if he has STI. “Wife justified refuse sex if husband has other women” is a dummy indicating whether a respondent find justified for a wife to refuse sex with her husband if he has other women. Controls are defined in [Table 1](#). *Region-survey* is a subnational region defined in DHS, interacted with its survey-year. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

5.2 Sexual Behaviour

Reproductive Evolved Psychology and Sexual Behaviour. In order to capture sexual behaviours that are prominent risk factors of HIV contagion, I follow the well-established medical and economic literature (e.g. Bertocchi and Dimico (2018)), and I create a dummy equals to one if an individual had any sexual activity in the last 4 weeks, a dummy for whether an individual reports having used a condom in her last sexual intercourse, a dummy equals to one if an individual reports having any extramarital affair in the last 12 months, and the number of extramarital affairs in the last 12 months that she reports (I focus on formally married individuals for these two latter outcomes).

Several lessons can be drawn from Table 7, which presents results from estimating Equation 1 on these sexual behaviour outcomes. To begin with, according to the estimates reported in second row, sub-saharan african women are significantly and sizably less likely to report adopting risky sexual behaviours than their male counterparts. This result is consistent with the evolutionary psychology theory I build on in my conceptual framework, as well as with numerous psychological studies (Buss (2016)) according to which, relative to females, males exhibit a stronger preference for casual relationships and sexual variety. However, I also find that among female individuals, women originating from ancestrally matrilineal ethnic groups have significantly riskier sexual behaviours that are more conducive to the spread of HIV, as compared to their patrilineal counterparts. More precisely, estimates reported in the third row indicate that matrilineal married women are 4.1 percentage points more likely to have had a sexual activity in the last 4 weeks; and 1.6 percentage points less likely to have used a condom during their last sexual intercourse, an effect representing about (23%) of the mean probability that a patrilineal male in my regression sample reported condom use. Interestingly, I find qualitatively similar results when working on individuals being in any marital status, while controlling for this latter (column 1 and 2). Further, I find that matrilineal married females are 2.6 percentage points (representing about 20% of the patrilineal married males' mean probability) more likely to report any extramarital affairs in the last 12 months. In addition, estimate from column 4 indicates that married women originating from an ancestrally matrilineal ethnic group had 0.068 more extramarital partners in the last 12 months than their patrilineal counterparts, an effect of large magnitude representing about 33% of the mean of patrilineal married males. Finally, it is worth noting that according to the estimates reported in the first row, and contrary to their female counterparts, matrilineal males do not have an overall significantly different propensity to adopt risky sexual behaviour than their patrilineal counterparts. All these results are fully consistent with my conceptual framework and highlight mechanisms at play in explaining the long-term effect of ancestral matrilineality of contemporaneous female HIV in Sub-Saharan Africa.

Routes of infection. Building on my conceptual framework and the results on sexual behaviours highlighted above, I expect matrilineal females to be mainly infected through extramarital routes of infection. This is exactly what I find in the analysis performed in Table 18 on couple's serodiscordancy status.

Restricting my attention to non-polygynous formally married couples with both wife and husbands tested for HIV in DHS and both of them originating from either a traditionally matrilineal or a traditionally patrilineal ethnic group, I find that when the wife originates from an ancestrally matrilineal ethnic group, she is about twice more likely (i.e. 1.3 percentage points) to be HIV positive while having a HIV negative husband, relative to wives

originating from ancestrally patrilineal ethnic groups. Importantly, this result remains marginally significant when I also include individual’s, (ancestral) ethnic group’s and village-geographic controls computed for the husband, in addition of controlling for whether the wife and the husband originate from different ancestral ethnic groups (“*Mixed ethnicity*”), as well as controlling for whether the wife and the husband originate from ethnic groups with different ancestral kinship organizations (matrilineal vs. patrilineal) (“*Mixed matrilineality*”). All in all, these results indicate that, in line with their riskier (relative to patrilineal women) sexual behaviour outside the domestic sphere, matrilineal women are significantly more likely to be infected by HIV through extramarital channels.

Moreover, results in the two last columns indicate that matrilineal couples are relatively less likely to be HIV+ seroconcordant. This finding may be interpreted as an evidence that the mechanisms highlighted by [Anderson \(2018\)](#) is also at play, namely that patrilineal females are more likely to be infected by their husband due to their lower sexual autonomy. Alternatively, this finding may also be explained by differential couple formation/dissolution dynamics between matrilineal and patrilineal groups (e.g. a matrilineal female may be relatively more likely to divorce⁵³ if her husband is HIV positive. In this case, the couple is dropped from my sample of couples, leading to non-random attrition).

While DHS data do not allow me to directly empirically test for the effect of matrilineality on gender-related social behaviour, it is worth noting that a recent literature has highlighted the influence of matrilineality on such gender related behaviour. In particular, [Lowes \(2018a\)](#) provides experimental evidences from Democratic Republic of Congo (DRC) that spouses from matrilineal ethnic groups cooperate less than their patrilineal counterparts. Further, [Lowes \(2018b\)](#) provides, in a similar context, experimental evidences that matrilineality closes the gender gap in risk-preference, with matrilineal women having a higher preference for risk than their patrilineal counterparts. These differences in contemporaneous behaviours could be additional factors shaping riskier sexual behaviour found here, and within marriage contraceptive behaviour explored in the next subsection.

5.3 Contraceptive Behaviour

Long-Term versus Short-Term Contraceptive Method. The result highlighted in the last subsection on matrilineal female’s use of condom during their last sexual intercourse suggests that matrilineal females may also differ in their contraceptive behaviour. This is what I investigate in this subsection.

To do so, I exploit information in DHS on current contraception method used by respondent to create (1) a dummy for whether the respondent report using condom as her current contraceptive method, (2) a dummy for whether the respondent report using a long-term contraception method (i.e. “male condom”; “(periodic) abstinence”; “withdrawal” ; or “prolonged abstinence”); (3) a dummy for whether the respondent report using a long-term contraceptive method (i.e. “pill”, “IUD”, “injection”, “female sterilization”, “implants/norplant”, or “lactational amenorrhea (LAM)”); (4) and a dummy for whether the respondent uses *any* contraceptive method. Incidentally, short-term contraceptive methods are also more *protective* methods against HIV, relative to short-term ones.

[Table 8](#) presents results from estimating [Equation 1](#) on these contraceptive use outcomes and reveals that, while not having a significantly different contraceptive behaviour at the extensive margin (i.e. the estimate of

⁵³See my result on divorce in [Table 6](#).

Table 7: **Ancestral Matrilineality and Sexual Behaviour (OLS)**

	All Sample		Formally Married Individuals			
	Sexual activity	Condom during last sex	Sexual activity	Condom during last sex	Infidelity	Nb. of extramarital partners
	(1)	(2)	(3)	(4)	(5)	(6)
Matrilineality	-0.005 (0.009)	-0.008 (0.008)	-0.017 (0.012)	-0.006 (0.007)	-0.006 (0.008)	-0.038* (0.021)
Female	-0.079*** (0.005)	-0.064*** (0.005)	-0.094*** (0.005)	-0.045*** (0.004)	-0.127*** (0.007)	-0.196*** (0.014)
Female × Matrilineality	0.032*** (0.010)	-0.019** (0.008)	0.041*** (0.011)	-0.016** (0.007)	0.026*** (0.010)	0.068*** (0.016)
Ind. Controls	Yes	Yes	Yes	Yes	Yes	Yes
Ethnic Group Controls	Yes	Yes	Yes	Yes	Yes	Yes
Village-Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes
Region-survey FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	181,108	134,888	104,483	96,448	104,489	104,489
Adj. R-squared	0.310	0.250	0.087	0.057	0.109	0.022
Clusters	1,688	1,630	1,558	1,541	1,558	1,558
Mean Dep. Var. (Patri. Males)	0.537	0.175	0.768	0.069	0.136	0.204

Notes: OLS estimates are reported with standard errors clustered at the ethnicity-DHS (within-country) region-survey (year) level in brackets. The unit of observation is an individual originating from either a traditionally matrilineal or a traditionally patrilineal ethnic group. “Matrilineality” indicates (dummy) whether an individual belongs to a traditionally matrilineal ethnic group. “Female” indicates (dummy) whether an individual is a female. “**Female × Matrilineality**” indicates (dummy) whether an individual is a female belonging to a traditionally matrilineal ethnic group. “Sexual activity” is a dummy indicating whether an individual reports having had any sexual activity in the last 4 weeks. “Condom during last sex” is a dummy indicating whether an individual reports having used a condom in her last sexual intercourse. “Infidelity” is a dummy indicating whether an individual reports having had any extramarital partner in the last 12 months. “Nb. of extramarital partners” is the number of extramarital partners in the last 12 months reported by an individual. Controls are defined in Table 1. *Region-survey* is a subnational region defined in DHS, interacted with its survey-year. Samples in columns 3 to 6 consist of formally married individuals. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Female × Matrilineality on the probability of using any contraception method is not significant (column 4)), matrilineal women significantly differ from their patrilineal counterparts in the contraceptive methods they use. More specifically, I find that matrilineal women are 3.1 percentage points less likely to report using condom as a current contraceptive method, as compared to patrilineal woman, an effect representing about 22% of the mean probability that patrilineal males report using condom as a contraceptive method (column 1)⁵⁴. This very low level of condom use within marriage in Sub-Saharan Africa is documented in the literature⁵⁵. More generally, I also find that matrilineal women are also 3.1 percentage points less likely to report using any short-term protective contraception methods, representing about 20% of the mean probability of patrilineal males (column 2). These results are in line with the effect of ancestral matrilineality on condom use in last sexual intercourse, that I found in the previous subsection. Nevertheless, I also find in column 3 that matrilineal women are 4.7 percentage points more likely than their patrilineal counterparts to report using a long-term contraceptive method, representing about 60% of the mean of patrilineal males.

To wrap up, it seems that, having a higher status and being more likely to be decision-maker regarding con-

⁵⁴Note that this mean probability is even lower for females, with only 3.1% (not reported) of females in my final sample reporting condom as a current contraceptive method.

⁵⁵Chimbiri (2007) calls condom an “intruder in marriage”.

traception (see [subsection 5.1](#)), matrilineal women are more likely to bear the responsibility of contraception, and substitute short-term (which are incidentally more protective) contraceptive methods by long-term contraceptive methods. This is in line with [Islam et al. \(2009\)](#), who find similar patterns in the context of Bangladesh, underlying that matrilineal *Garó* women’s contraceptive behaviour differs from that of their *patrilineal* Bengali counterparts, in that their current use of contraceptives is higher than at national level, but the prevalence of modern male methods is low.

An other driver of such contraceptive behaviour may be the desired fertility of matrilineal females⁵⁶. Indeed, as reported in [Table 16](#) in appendix, I find a higher gender discrepancy in desired fertility for the matrilineal group. In other words, I find that matrilineal females are relatively more likely to desire more children than their patrilineal counterparts, while the opposite is true for male individuals. Along this line, I find a similar result regarding actual fertility. These results may be fully rationalized in my conceptual framework: matrilineal male’s biological children will never integrate his lineage in addition of not inheriting from him. On the contrary, as highlighted in [Lowes \(2018a\)](#), by integrating her lineage and allowing her to benefit from support from her brothers, children constitute an asset for a matrilineal woman, which may drive her fertility preferences.

All in all, differences in contraceptive behaviours highlighted in this subsection, and in particular matrilineal females’ lower propensity to use protective contraceptive methods (particularly condom), constitute an additional mechanism that help explaining the highest contemporaneous rates of HIV found within matrilineal female populations.

Condom use and internalized risk. My previous finding that matrilineal female have a lower use of condom may hide substantial heterogeneity. I investigate in [Table 17](#) in appendix whether matrilineal individuals adopt different condom use behaviour when they have properly internalized the risk of HIV transmission, and I find that, indeed, it seems to be the case. More precisely, I first investigate heterogeneity in perception of the condom as a mean of reducing HIV transmission, and I find that the negative effect of matrilineality on condom use that I previously found is null for individuals who have a correct belief about the role of condom. Moreover, I find that when they are seropositive, matrilineal individuals are in fact more likely to report using condom as a contraceptive method. Even more strikingly, I find that this heterogeneity is mainly driven by individuals who have ever been tested for HIV before DHS, and who are therefore presumably aware of their serostatus. In fact, it seems that, when they have correctly internalized the risk of transmitting the virus, matrilineal females take advantage of their higher ability to impose safe sexual practices. This is an encouraging result which calls for complementary policies aimed at raising awareness about the actual riskiness of promiscuous sexual behaviour of empowered populations, namely matrilineal females.

5.4 Discarded Mechanisms

Acknowledgment of HIV Risks and Access to Condom. To explore whether matrilineal females’ adoption of riskier sexual and contraceptive behaviours may result from a lower acknowledgment of HIV risks and access to condom, I create several outcomes from information in DHS on acknowledgment of HIV risks and role of condom

⁵⁶Note that I do control for the number of children in my regressions.

Table 8: **Ancestral Matrilineality and Contraception (OLS)**

	Condom	Short-term protective contraceptive method	Long-term contraceptive method	Any contraceptive method
	(1)	(2)	(3)	(4)
Matrilineality	0.018* (0.010)	0.020* (0.010)	-0.017* (0.009)	0.005 (0.011)
Female	-0.106*** (0.009)	-0.105*** (0.009)	0.046*** (0.004)	-0.058*** (0.009)
Female × Matrilineality	-0.031** (0.014)	-0.031** (0.015)	0.047*** (0.009)	0.017 (0.013)
Ind. Controls	Yes	Yes	Yes	Yes
Ethnic Group Controls	Yes	Yes	Yes	Yes
Village-Geographic Controls	Yes	Yes	Yes	Yes
Region-survey FE	Yes	Yes	Yes	Yes
Observations	147,175	146,185	146,185	147,175
Adj. R-squared	0.155	0.152	0.155	0.156
Clusters	1,507	1,504	1,504	1,507
Mean Dep. Var. (Patri. Males)	0.138	0.163	0.078	0.245

Notes: OLS estimates are reported with standard errors clustered at the ethnicity-DHS (within-country) region-survey (year) level in brackets. The unit of observation is an individual originating from either a traditionally matrilineal or a traditionally patrilineal ethnic group. “Matrilineality” indicates (dummy) whether an individual belongs to a traditionally matrilineal ethnic group. “Female” indicates (dummy) whether an individual is a female. “**Female × Matrilineality**” indicates (dummy) whether an individual is a female belonging to a traditionally matrilineal ethnic group. “Condom” is a dummy indicating whether an individual reports “condom (male)” as her current contraceptive method. “Short-term protective contraceptive method” is a dummy indicating whether an individual reports a short-term contraceptive method (i.e. “male condom”; “(periodic) abstinence”; or “withdrawal”), and which is incidentally a protective method against HIV, as her current contraceptive method. “Long-term contraceptive method” is a dummy indicating whether an individual reports a long-term contraceptive method (i.e. “pill”, “IUD”, “injection”, “female sterilization”, “implants/norplant”, or “lactational amenorrhea (LAM)”), and which is incidentally a protective method against HIV, as her current contraceptive method. “Any contraceptive method” is a dummy indicating whether an individual reports currently using any contraception method. Controls are defined in [Table 1](#). *Region-survey* is a subnational region defined in DHS, interacted with its survey-year. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

in reducing these risks. I create a dummy indicating whether an individual has ever heard about AIDS (“*Heard of AIDS*”); a dummy indicating whether an individual has ever heard about any STI (“*Heard of STI*”); a dummy indicating whether an individual thinks that always using condoms during sex reduces chance of getting HIV (“*Think condom reduces HIV*”); and a dummy indicating whether an individual thinks that having only one sexual partner reduces chance of getting HIV (“*Think having one partner reduces HIV*”). Further, I also create outcomes related to access to condom, namely a dummy indicating whether an individual knows a source to get male condoms (“*Know a source to get condom*”); as well as a dummy indicating whether an individual can get herself a male condom (“*Can get condom*”).

Results from regressing [Equation 5](#) on these outcomes are reported in [Table 20](#). Two main conclusions can be drawn from this analysis: first, I do not find any statistically significant effect of originating from an ancestrally matrilineal ethnic group on female’s acknowledgment of HIV/STI risks (in fact, the vast majority of both matrilineal and patrilineal females acknowledges such risks⁵⁷). Consequently, I can rule out differential acknowledgment of HIV risks and role of condom against HIV as factors explaining matrilineal’s females adoption of more promiscuous sex-

⁵⁷In my sample, about 98% of matrilineal females and about 94% of patrilineal females have ever heard of HIV.

ual behaviour. Second, matrilineal females’ lower use of condom cannot be explained by a lack of access to condom. If any, I conversely find that, in line with their higher sexual autonomy, matrilineal women have a significantly easier access to condom.

Sexual Debuts. For biological reasons, young women constitute a population at high risk of contracting HIV when exposed to it (Yi et al. (2013)). As such, increase in early and/or premarital sexual activity is a driver of the spread of the virus (Oster (2005), Case and Paxson (2013)). To test whether this could be a driver of matrilineal female’s highest rate of HIV, I exploit information in DHS on reported age at first sexual intercourse, as well as age at first marriage. Interestingly and as reported in Table 19 in appendix, I find that matrilineal females start their sexual life later than their patrilineal counterparts, suggesting that my main result is in fact mitigated by this mechanism. Further, my results on age at first marriage also suggest a relatively lower age gap between matrilineal spouses, which additionally mitigates my main findings on female sexual behaviour and HIV. As a matter of fact, a higher age gap with her husband may constitute a source of dissatisfaction for a married wife, potentially encouraging her to engage into extramarital relationships (Bertocchi and Dimico (2018)). However, if one assumes that an older husband has a higher likelihood of being HIV positive (because of a longer sexual life) and is consequently more likely to infect his wife, this result may partially explain why, contrary to Anderson (2018), I mainly capture a female’s *extramarital* route of infection.

6 An Epidemiological Approach

In this section, I adopt an epidemiological approach to conduct a simulation exercise. The aim of this simulation exercise is not to embrace the full complexity of the world, but rather to provide some insights on how the differences in sexual behaviours found in the previous section translate into different gender-specific HIV rates dynamics, between ancestrally matrilineal and ancestrally patrilineal societies.

6.1 The model

To illustrate the differential effects of being matrilineal or patrilineal on female and male HIV propagation rates, through the two main routes of infection highlighted in the empirical results, namely extra-couple casual sexual relationships vs. intra-couple long-term sexual relationships, I adopt a compartmental SI (“Susceptible to Infected”) epidemic model, that I adapt from Worden and Porco (2017) and where I add heterogeneity by gender.

I consider a closed population composed of female and male individuals (denoted with subscript f and m respectively). Further, each of these gendered populations are composed of a fraction of “susceptible” (i.e. not infected) individuals, and of a fraction of “infected” individuals, denoted $s_{f;m}$ and $i_{f;m}$ respectively.

Sexual behaviours are represented by the following parameters: $\mathbb{1}$ is an indicator of a long-term committed relationship; $\alpha_{f;m}$ is the number of extramarital partners; $\mu_{f;m}$ is the number of sexual intercourses with the long-term committed partner; $\tau_{f;m}$ is the number of sexual intercourses with each extramarital partner. Further, the probability that a male is infected per each sexual intercourse with a HIV positive female is $\delta_m = \gamma_m \times \rho_m$, namely the product of the probability that he does not use condom during sexual intercourse with the transmission risk for one-

time male unprotected sex (assuming for simplicity that condom is totally efficient against HIV transmission). The probability that a female is infected per each sexual intercourse with a HIV positive male is $\delta_f = \gamma_f \times (\rho_m \times \omega_f)$, with ω_f capturing gender difference in biological susceptibility of contracting the virus when exposed to it⁵⁸. Finally, I assume random mixing, namely that the probability that an individual faces a HIV positive sexual encounter (should it be her long-term committed partner or an extramarital casual partner) is equal to the proportion of infected individuals in the population of the other sex (i.e. i_m for females' probability and i_f for males' probability).

Female propagation rates (λ_f) can thereby be modeled as follows:

$$\lambda_f = \underbrace{\mathbb{1}\mu_f\delta_f i_m}_{\text{Intra-couple infection}} + \underbrace{\alpha_f\tau_f\delta_f i_m}_{\text{Extra-couple infection}} \quad (6)$$

\Leftrightarrow

$$\lambda_f = (\underbrace{\mathbb{1}\mu_f}_{\text{Intra-couple intercourses}} + \underbrace{\alpha_f\tau_f}_{\text{Extra-couple intercourses}}) \delta_f i_m \quad (7)$$

Symmetrically, male's propagation rate (λ_m) can be modeled as follows:

$$\lambda_m = (\mathbb{1}\mu_m + \alpha_m\tau_m)\delta_m i_f \quad (8)$$

From [Equation 7](#) and [Equation 8](#), gender difference ($\lambda_f - \lambda_m$) in propagation rates can be modeled as follows:

$$\lambda_f - \lambda_m = (\mathbb{1}\mu_f + \alpha_f\tau_f)\delta_f i_m - (\mathbb{1}\mu_m + \alpha_m\tau_m)\delta_m i_f \quad (9)$$

\Leftrightarrow

$$\lambda_f - \lambda_m = \underbrace{\mathbb{1}(\mu_f\delta_f i_m - \mu_m\delta_m i_f)}_{\text{Gender difference in intra-couple infection}} + \underbrace{\alpha_f\tau_f\delta_f i_m}_{\text{Female extra-couple infection}} - \underbrace{\alpha_m\tau_m\delta_m i_f}_{\text{Male extra-couple infection}} \quad (10)$$

Interestingly, if one assumes that $\mu_f = \mu_m$, gender difference in intra-couple infection reduces simply to the gender difference in biological susceptibility of contracting the virus when exposed to it, and the difference in the probability that the husband versus the wife is infected.

It is worth noting that, for simplicity, I assume exogenous and time invariant sexual behavioural parameters ($\mathbb{1}; \alpha_{f;m}; \mu_{f;m}; \tau_{f;m}$ and $\lambda_{f;m}$). Further, I assume random mixing. In other words, I assume that the probability that an individual face a HIV positive partner is equal to the proportion of HIV positive individuals in the pool of the opposite sex ($i_{f;m}$). Additionally, I also assume that this probability is the same for a long-term committed partner

⁵⁸The medical literature highlights that females have a relatively higher susceptibility of contracting HIV when exposed to it, for biological reasons (i.e. larger surface area of mucosal HIV exposure; increased mucosal expression of the HIV co-receptor CCR5; and a greater probability of virus exposure on the rectal mucosa among other factors ([Yi et al. \(2013\)](#))).

or an extramarital partner. Moreover, I assume time-invariant contagiousness. Those simplifying assumptions might be relaxed if one is interested in embracing with greater care the full complexity of the world, which is beyond the scope of this simulation exercise.

6.2 The effect of Matrilineality: Comparative Statics

Females propagation rates (λ_f). From Equation 7, it is straightforward to see that λ_f is increasing in all females' sexual behavioural parameters. Further, according to my conceptual framework, as well as my empirical results on sexual behaviour and contraceptive use, the value of these parameters should be higher for matrilineal females, relative to their patrilineal counterparts, leading to a higher HIV propagation rate for matrilineal female populations. Further, λ_f is also increasing in i_m . Therefore, assuming that i_f and i_m coevolve in the same direction, I expect a self-reinforcing effect for matrilineal females: since matrilineal females are characterized by higher infection rates, more matrilineal males should be contaminated to, leading in turn to even more contaminated matrilineal females. All in all, this leads to the following prediction:

Prediction 1: $\lambda_f^M > \lambda_f^P$,⁵⁹ the HIV propagation rate is higher for matrilineal females, relative to their patrilineal counterparts.

Males propagation rates (λ_m). From Equation 8, it is straightforward to see that λ_m is increasing in all males' sexual behavioural parameters. However, since ancestral matrilineality does not yield significant sexual behavioural changes among males populations according to my conceptual framework, as well as my empirical results, the value of these parameters should be the same for both matrilineal and patrilineal males. Nevertheless, λ_m is also increasing in i_f , which should be higher for matrilineal populations (i.e. infection rate is higher for matrilineal females than their patrilineal counterparts). Consequently, relative to patrilineal male populations, I expect the HIV propagation rate of matrilineal male populations to be only weakly higher. This leads to the following prediction:

Prediction 2: $\lambda_m^M \geq \lambda_m^P$, the HIV propagation rate is weakly higher for matrilineal males, relative to their patrilineal counterparts.

Gender differences in propagation rates ($\lambda_f - \lambda_m$). From Equation 10, it is straightforward to see that $\lambda_f - \lambda_m$ is increasing in females extramarital promiscuous sexual behaviour, and decreasing in males extramarital promiscuous sexual behaviour. As discussed earlier, and in line with my conceptual framework as well as my empirical findings, matrilineal females adopt more promiscuous extramarital sexual behaviours, while matrilineal males adopt as promiscuous extramarital sexual behaviours as their patrilineal counterparts. This leads to the following prediction:

Prediction 3: $\lambda_f^M - \lambda_m^M > \lambda_f^P - \lambda_m^P$, the gender difference in HIV propagation rates is higher for matrilineal populations.

⁵⁹Superscripts M and P denote Matrilineality and Patrilineality, respectively.

6.3 Simulation

Building on my SI compartmental model of infection discussed earlier, I perform here a numerical simulation of the gender-specific HIV rates dynamics in matrilineal versus patrilineal contexts. These HIV rates dynamics can be represented by the following system of equations:

$$\begin{cases} s_{f;t+1} &= s_{f;t} - (\mathbb{1}\mu_f + \alpha_f\tau_f)\delta_f i_{m;t} s_{f;t} \\ i_{f;t+1} &= i_{f;t} + (\mathbb{1}\mu_f + \alpha_f\tau_f)\delta_f i_{m;t} s_{f;t} \\ s_{m;t+1} &= s_{m;t} - (\mathbb{1}\mu_m + \alpha_m\tau_m)\delta_m i_{f;t} s_{m;t} \\ i_{m;t+1} &= i_{m;t} + (\mathbb{1}\mu_m + \alpha_m\tau_m)\delta_m i_{f;t} s_{m;t} \end{cases}$$

with $s_{f;t}$ and $s_{m;t}$ denoting the proportion of “susceptible” (i.e. not infected) individuals among female and male populations respectively; and $i_{f;t}$ and $i_{m;t}$ denoting the proportion of “infected” individuals among female and male populations respectively.

According to my conceptual framework as well as my empirical results, the highest rates of HIV among matrilineal females are mainly driven by matrilineal females adopting more promiscuous extramarital sexual behaviours, and I find that matrilineal females are mainly contaminated through extramarital routes of infection. To illustrate how this leads to higher gender differential in HIV rates among matrilineal populations, I assume a hypothetical closed population (i.e. no births, no deaths, no migration) of 100 females and 100 males who are in a long-term committed sexual relationship (e.g. formal marriage), with 1 infected female (1%) and 1 infected male (1%) at the initial period, and I simulate HIV rates dynamics on a monthly basis, over a 10 years period.

Parameters values. $\mathbb{1}$ is an indicator of a long-term committed sexual relationship (e.g. formal marriage). I assign to the parameters α_f and α_m the values I have estimated in column 6 of [Table 7](#). More precisely, I assume no difference between matrilineal and patrilineal males and assign $\alpha_m^M = \alpha_m^P = \alpha_m = 0.204$ for both matrilineal and patrilineal males, based on the mean of my regression sample patrilineal males; and I assign $\alpha_m^M = 0.076$ for matrilineal females; and $\alpha_m^P = 0.008$ for patrilineal females.⁶⁰

The probability that a male is infected per each sexual intercourse with a HIV positive female is $\delta_m = \gamma_m \times \rho_m$, namely the product of the probability that he does not use condom during sexual intercourse with the transmission risk for one-time male unprotected sex (assuming for simplicity that condom is totally efficient against HIV transmission). Therefore, the value of γ_m is simply 1 minus the probability that males use condom. Consequently, based on the mean probability that a condom was used during last intercourse by patrilineal males in my regression sample of column 4 of [Table 7](#), $\gamma_m = 1 - 0.069 = 0.931$. Again, I assign the same value for both matrilineal and patrilineal males. ρ_m is taken from [Greenwood et al. \(2019\)](#) and is set equal to 0.0045.⁶¹ As a result, $\delta_m = 0.931 \times 0.0045 = 0.00419$.

The probability that a female is infected per each sexual intercourse with a HIV positive male is $\delta_f = \gamma_f \times (\rho_m \times \omega_f)$, with ω_f capturing gender difference in biological susceptibility of contracting the virus when exposed

⁶⁰As for the other females parameter values based on my estimations, I assign values as follows: mean patrilineal males + “Female” estimate for patrilineal females; mean of patrilineal males + “Female” estimate + “Female x Matrilineal” estimate for matrilineal females. Here, it gives $0.204 - 0.196 = 0.008$ for patrilineal females; and $0.204 - 0.196 + 0.068 = 0.076$ for matrilineal females, based on my estimates in column 6 in [Table 7](#).

⁶¹This number falls in the range of estimates reported by a variety of studies, according to [Greenwood et al. \(2019\)](#).

to it. Following my estimates in column 4 of [Table 7](#), $\gamma_f^M = 1 - (0.069 - 0.045 - 0.016) = 0.992$, and $\gamma_f^P = 1 - (0.069 - 0.045) = 0.976$, for matrilineal and patrilineal females respectively. Further, building on several medical studies, [Greenwood et al. \(2019\)](#) assume that women are 75% more likely to get infected than males for physiological and anatomical reasons. I follow them and assign $\omega_f = 1.75$. Therefore, $\delta_f^M = 0.992 \times 0.0045 \times 1.75 = 0.007812$ for matrilineal females; and $\delta_f^P = 0.976 \times 0.0045 \times 1.75 = 0.007686$ for patrilineal females

Unfortunately, the DHS does not directly provide data on the frequency of sexual intercourse within couples. Therefore, I perform the simulation exercise for different values of μ_m , τ_m , μ_f , and τ_f . For consistency purpose, I assume that $\mu_m = \mu_f$ in both matrilineal and patrilineal context. In other words, wife and husband have the same number of intra-couple sexual intercourse for each time period. Further, for simplicity, I assume that in matrilineal context $\tau_m = \tau_f$ (i.e. males and females have the same number of sexual intercourses per time period with *each* of their extramarital partner⁶²). Finally, given that according to my empirical findings patrilineal females report a lower sexual activity than their matrilineal counterparts, to determine τ_f^P I lower τ_f^M in the same proportion than patrilineal females reported sexual activity. More formally, according to my estimate in column 3 of [Table 7](#), patrilineal females are 4.1 percentage points less likely to report any sexual activity in the last month than matrilineal females. Accordingly, I set $\tau_f^P = 0.959 \times \tau_f^M$. As detailed below, I perform the exercise for different values of μ_m and τ_m . The values and definition of parameters are summarized in [Table 21](#) in appendix.

Simulation results. The main results of this simulation exercise⁶³, computed under a scenario where $\mu_m = \tau_m = 4$, are presented numerically in [Table 9](#), and graphically in [Figure 4](#). Several lessons can be drawn. First, at any time period matrilineal female HIV rates are higher than patrilineal female HIV rates, and this gap is further increasing through time (comparing column 1 and 4). This is consistent with the first prediction of my model. Second, the gap in HIV rates between matrilineal and patrilineal males is of weaker magnitude than for females. This is consistent with the second prediction of my model. Finally, in line with these two previous results and as indicated in the last column, at any time period the gender gap in HIV rates is relatively higher in the matrilineal context, and is further increasing through time. This result is consistent with the third prediction of my model.

As a robustness check, in [Table 22](#) as well as [Figure 10](#) and [Figure 11](#) in appendix, I perform the same simulation exercise under alternative scenarios, namely $\mu_m = 4$ and $\tau_m = 2$ to allow for the possibility that the frequency of sexual intercourse per extramarital partner is lower than the frequency with long-term committed partner; and $\mu_m = 2$ and $\tau_m = 2$ to emphasize an overall lower frequency of sexual intercourses within both long-term comitted and extramarital couples. Reassuringly, the conclusions drawn from these alternative exercises are qualitatively similar to my main simulation exercise.

To conclude, in addition of providing an illustration of the different HIV propagation dynamics at play, the results of this simulation exercise confirm the main finding of my paper: in Sub-Saharan Africa, ancestral matrilineality is a driver of contemporaneous female HIV.

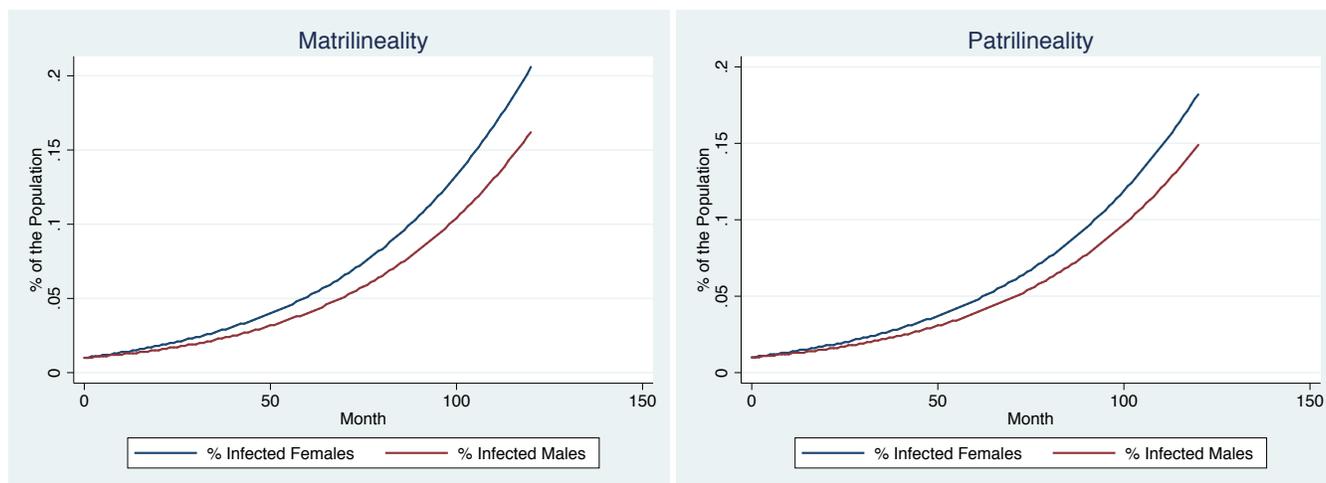
⁶²This does not mean that females and males have the same number of extramarital partners. This latter is captured by α_f and α_m respectively.

⁶³I only report proportions of infected individuals. According to my SI model, the proportion of susceptible individuals is simply the proportion of non-infected individuals.

Table 9: **Simulated HIV Rates**

Months	Scenario 1 ($\mu_m = \tau_m = 4$)						Raw Diff-in-Diff
	Matrilineality			Patrilineality			
	Females	Males	Difference	Females	Males	Difference	
1	0.010	0.010	0.000	0.010	0.010	0.000	0.000
25	0.021	0.017	0.004	0.020	0.017	0.003	0.001
50	0.040	0.032	0.008	0.037	0.031	0.006	0.002
75	0.074	0.058	0.016	0.067	0.055	0.012	0.004
100	0.133	0.104	0.029	0.119	0.097	0.022	0.007
120	0.206	0.162	0.044	0.182	0.149	0.033	0.011

Notes: This table reports the simulated HIV rates, based on the compartmental SI epidemic model detailed in [section 6](#), and assuming that $\mu_m = \tau_m = 4$.

Figure 4: Simulation - Scenario 1 ($\mu_m = \tau_m = 4$)

7 Conclusions

In this paper I build on the latest evolutionary psychological theories as well as on the anthropological literature to test the hypothesis that females originating from ethnic groups that ancestrally adopted matrilineal kinship organizations are more likely to be infected by HIV today than their patrilineal counterparts. As a matter of fact, ancestral matrilineal kinship organizations constituted environments in which substituting long-term committed sexual relationships by sexual variety may have been a relatively more beneficial strategy for females in terms of reproductive success, allowing them to substitute matrilineal males' relatively lower propensity to commit in the long-term by an increased access to better genes, in order to ensure the survival and subsequent reproductive success of their offsprings. Consequently, benefiting from a higher sexual autonomy through their social position, relative to patrilineal females, matrilineal females should display relatively more promiscuous sexual behaviours.

Using data from more than 280,000 individuals from 18 countries, including more than 190,000 women, and exploiting within-country variation in ethnic groups' ancestral kinship organizations, I find that women originating from ancestrally matrilineal ethnic groups are significantly more likely to be HIV positive today, as compared to women originating from ancestrally patrilineal ethnic groups. Further, I show that this result is robust to a large

set of cultural, historical, geographical and environmental controls, as well as the inclusion of (within-country) region-survey (year) fixed effects, aimed at controlling for alternative channels. Additionally, I show that this long-lasting “matrilineal effect” on female HIV is not driven by any differential selection into DHS HIV testing, nor by differences in overall women’s health status, but is specific to sexually transmitted diseases.

I go one step further by formally testing for omitted bias, using [Altonji et al. \(2005\)](#) ratios and estimating [Oster \(2017\)](#) bias-adjusted lower bound coefficients. These latter provide very little support for the presence of an omitted bias in my OLS estimates.

Going beyond in identifying causal relationships and correcting for potential reverse causality, I exploit GPS location of DHS villages as well as digitized Murdock’s map of ancestral ethnic groups boundaries in Africa to compute the distance between villages and nearest ancestral matrilineal ethnic boundary, a measure that I use to implement an instrumental variable (IV) strategy as well as a geographic regression discontinuity design (RDD). I show that my main estimates are remarkably qualitatively robust across all these specifications. Finally, estimating the average treatment over treated effect of being located on an ancestrally matrilineal geographic area on DHS villages’ proportion of HIV positive females, matching villages with their nearest neighbor in an ancestrally patrilineal geographic area based on a large array of geographic observables, I find that villages located in an ancestrally matrilineal areas are characterized by significantly higher females HIV rates than their nearest neighbor in patrilineal areas, providing additional support to my main OLS findings

Consistent with my conceptual framework, I find that the relatively higher HIV rates found among matrilineal female populations can be explained by these latter being more able than patrilineal females to follow their sexual preferences, and therefore adopting more promiscuous sexual behaviours. Along these lines, I find that matrilineal females are mainly infected through extramarital routes of infection.

Further, I highlight differences in contraceptive behaviours as a second main mechanism. More precisely, I find evidences that matrilineal females relatively higher contraception-related decision-making translates into them substituting short-term contraceptive methods (condom in particular) by long-term ones. By doing so, matrilineal women incidentally substitute more protective contraceptive methods against HIV by less protective ones. Promisingly, I nevertheless find evidences that when they have internalized the risk of HIV transmission, matrilineal individuals are then more likely to adopt condom as a contraceptive method. Finally, I discard differences in access to condom as well as in sexual debuts as additional mechanisms.

In the end, I build on the epidemiological literature to perform a numerical simulation exercise, aimed at illustrating how the differences in gender-specific sexual and contraceptive behaviours found between matrilineal and patrilineal populations translate into differences in gender-specific HIV rate dynamics between these populations. Under credible parameter values (most of them being determined by my empirical findings), my simulations indeed show dynamics that are consistent with my conceptual framework as well as my empirical findings.

While recent research literature and policy recommendations aimed at reducing the spread of HIV among women in Sub-Saharan Africa, a population at high risk, have put the emphasis on the need to empower women in order to enable them to impose safe sexual practices to their male counterparts, this paper highlights complement mechanisms which call for complementary policies. As a matter of fact, my results have shown that matrilineal women, despite their higher sexual autonomy and ability to impose safe sexual practices to their husbands, are relatively more

likely to adopt sexual and contraceptive behaviours that put them at risk of contracting HIV. Nevertheless, I have also highlighted that, once beliefs about actual risks of HIV transmission are correctly internalized, matrilineal women actually make use of their higher decision-making to increase their protective contraception use (condom in particular), as compared to their patrilineal counterparts. This promising result calls for complementary policies targeting empowered women, matrilineal women specifically, and aimed at raising their awareness about the actual riskiness of adopting promiscuous sexual behaviours. It is hoped that such policies will induce behavioural changes that will allow to restrain the spread of HIV in Sub-Saharan Africa.

References

- Abadie, Alberto and Guido W Imbens**, “Large sample properties of matching estimators for average treatment effects,” *Econometrica*, 2006, *74* (1), 235–267.
- **and –**, “Bias-corrected matching estimators for average treatment effects,” *Journal of Business & Economic Statistics*, 2011, *29* (1), 1–11.
- Alesina, Alberto, Benedetta Brioschi, and Eliana La Ferrara**, “Violence against women: a cross-cultural analysis for Africa,” *NBER Working Paper*, 2016.
- **, Paula Giuliano, and Nathan Nunn**, “On the origins of gender roles: Women and the plough,” *The Quarterly Journal of Economics*, 2013, *128* (2), 469–530.
- Altonji, Joseph G, Todd E Elder, and Christopher R Taber**, “Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools,” *Journal of Political Economy*, 2005, *113* (1), 151–184.
- Anderson, Siwan**, “Legal Origins and Female HIV,” *The American Economic Review*, 2018, *108* (6), 1407–39.
- Andrews, Isaiah**, “Valid two-step identification-robust confidence sets for GMM,” *Review of Economics and Statistics*, 2018, *100* (2), 337–348.
- Ashraf, Nava, Natalie Bau, Nathan Nunn, and Alessandra Voena**, “Bride Price and Female Education,” *Journal of Political Economy*, 2019.
- Ashraf, Quamrul and Oded Galor**, “The “Out of Africa” hypothesis, human genetic diversity, and comparative economic development,” *The American Economic Review*, 2013, *103* (1), 1–46.
- **and –**, “Genetic diversity and the origins of cultural fragmentation,” *The American Economic Review: Papers and Proceedings*, 2013, *103* (3), 528–33.
- Baland, Jean-Marie and Roberta Ziparo**, “Intra-household bargaining in poor countries,” *WIDER Working Paper*, 2017.
- Bargain, Olivier, Jordan Loper, and Roberta Ziparo**, “Traditional Norm, Access to Divorce and Women’s Empowerment: Evidence from Indonesia,” *Working Paper*, 2019.
- Becker, Anke**, “On the Economic Origins of Constraints on Women’s Sexuality,” *Working Paper*, 2018.
- Bertocchi, Graziella and Arcangelo Dimico**, “The Long-Term Determinants of Female HIV Infection in Africa: The Slave Trade, Polygyny, and Sexual Behavior,” *Working Paper*, 2018.
- Besley, Timothy and Marta Reynal-Querol**, “The legacy of historical conflict: Evidence from Africa,” *American Political Science Review*, 2014, *108* (2), 319–336.
- Buss, David M**, *The Evolution of Desire - Strategies of human mating*, Basic Books, 2016.
- Cagé, Julia and Valeria Rueda**, “Sex and the Mission: The Conflicting Effects of Early Christian Investments on the HIV Epidemic in sub-Saharan Africa,” *Working Paper*, 2018.

- Case, Anne and Christina Paxson**, “HIV risk and adolescent behaviors in Africa,” *The American Economic Review: Papers & Proceedings*, 2013, 103 (3), 433–38.
- Chimbiri, Agnes M**, “The condom is an “intruder” in marriage: evidence from rural Malawi,” *Social Science & Medicine*, 2007, 64 (5), 1102–1115.
- Cordero-Coma, Julia and Richard Breen**, “HIV prevention and social desirability: husband–wife discrepancies in reports of condom use,” *Journal of Marriage and Family*, 2012, 74 (3), 601–613.
- Corno, Lucia and Damien De Walque**, “Mines, migration and HIV/AIDS in Southern Africa,” *Journal of African Economies*, 2012, 21 (3), 465–498.
- Delavande, Adeline and Hans-Peter Kohler**, “HIV/AIDS-related expectations and risky sexual behaviour in Malawi,” *The Review of Economic Studies*, 2015, 83 (1), 118–164.
- Diamond, Jared M**, *Why is sex fun?: the evolution of human sexuality*, Basic Books, 1998.
- Fortunato, Laura**, “The Evolution of Matrilineal Kinship Organization,” *Proceedings of the Royal Society London*, 2012.
- Fox, Robin**, *Kinship and Marriage: An Anthropological Perspective*, Cambridge: Cambridge University Press, 1934.
- Giuliano, Paola**, “Gender: An historical perspective,” *The Oxford Handbook of Women and the Economy*, 2017.
- **and Nathan Nunn**, “Understanding cultural persistence and Change,” *NBER Working Paper*, 2017.
- **and –**, “Ancestral characteristics of modern populations,” *Economic History of Developing Regions*, 2018, 33 (1), 1–17.
- González, Felipe and Edward Miguel**, “War and local collective action in Sierra Leone: A comment on the use of coefficient stability approaches,” *Journal of Public Economics*, 2015, 128, 30–33.
- Greenwood, Jeremy, Philipp Kircher, Cezar Santos, and Michèle Tertilt**, “An equilibrium model of the African HIV/AIDS epidemic,” *Econometrica*, 2019, 87 (4), 1081–1113.
- Greiling, Heidi and David M Buss**, “Women’s sexual strategies: The hidden dimension of extra-pair mating,” *Personality and Individual Differences*, 2000, 28 (5), 929–963.
- Islam, Mohammad Amirul, M Rakibul Islam, and Banya Banowary**, “Sex preference as a determinant of contraceptive use in matrilineal societies: a study on the Garo of Bangladesh,” *The European Journal of Contraception & Reproductive Health Care*, 2009, 14 (4), 301–306.
- Lowes, Sara**, “Kinship systems, gender norms, and household bargaining: Evidence from the Matrilineal Belt,” *Working Paper*, 2018.
- , “Matrilineal Kinship and Gender Differences in Competition,” *Working Paper*, 2018.

- Lowes, Sara Rachel and Eduardo Montero**, “The Legacy of Colonial Medicine in Central Africa,” *Working Paper*, 2018.
- Michalopoulos, Stelios, Louis Putterman, and David N Weil**, “The influence of ancestral lifeways on individual economic outcomes in Sub-Saharan Africa,” *Journal of the European Economic Association*, 2019.
- Mishra, Vinod, Martin Vaessen, J Boerma, Fred Arnold, Ann Way, Bernard Barrere, Anne Cross, Rathavuth Hong, and Jasbir Sangha**, “HIV testing in national population-based surveys: experience from the Demographic and Health Surveys,” *Bulletin of the World Health Organization*, 2006, *84*, 537–545.
- Montalvo, José G and Marta Reynal-Querol**, “Ethnic polarization, potential conflict, and civil wars,” *The American Economic Review*, 2005, *95* (3), 796–816.
- Moscona, Jacob, Nathan Nunn, and James A Robinson**, “Social Structure and Conflict: Evidence from Sub-Saharan Africa,” *Working Paper*, 2019.
- Nunn, Nathan**, “Religious conversion in colonial Africa,” *The American Economic Review: Papers & Proceedings*, 2010, *100* (2), 147–52.
- **and Leonard Wantchekon**, “The slave trade and the origins of mistrust in Africa,” *The American Economic Review*, 2011, *101* (7), 3221–52.
- Nyqvist, Martina Bjorkman, Lucia Corno, Damien De Walque, and Jakob Svensson**, “Incentivizing Safer Sexual Behavior: Evidence from a Lottery Experiment on HIV Prevention,” *American Economic Journal: Applied*, Forthcoming.
- Oster, Emily**, “Sexually transmitted infections, sexual behavior, and the HIV/AIDS epidemic,” *The Quarterly Journal of Economics*, 2005, *120* (2), 467–515.
- , “HIV and sexual behavior change: Why not Africa?,” *Journal of Health Economics*, 2012, *31* (1), 35–49.
- , “Routes of infection: Exports and HIV incidence in Sub-Saharan Africa,” *Journal of the European Economic Association*, 2012, *10* (5), 1025–1058.
- , “Unobservable selection and coefficient stability: Theory and evidence,” *Journal of Business & Economic Statistics*, 2017, pp. 1–18.
- Paula, Áureo De, Gil Shapira, and Petra E Todd**, “How beliefs about HIV status affect risky behaviors: Evidence from Malawi,” *Journal of Applied Econometrics*, 2014, *29* (6), 944–964.
- Porta, Rafael La, Florencio Lopez de Silanes, and Andrei Shleifer**, “The Economic Consequences of Legal Origins,” *Journal of Economic Literature*, 2008, *46* (2).
- Sun, Liyang**, “Implementing valid two-step identification-robust confidence sets for linear instrumental-variables models,” *The Stata Journal*, 2018, *18* (4), 803–825.
- Tassier, Troy**, *The Economics of Epidemiology*, Springer, 2013.

- Tequame, Miron**, “HIV, risky behavior and ethno-linguistic heterogeneity,” *Journal of Economics and Statistics*, 2012, *232* (6), 606–632.
- Teso, Edoardo**, “The Long-Term Effect of Demographic Shocks on the Evolution of Gender Roles: Evidence from the Transatlantic Slave Trade,” *Journal of the European Economic Association*, 2019.
- Thornton, Rebecca L**, “The demand for, and impact of, learning HIV status,” *The American Economic Review*, 2008, *98* (5), 1829–63.
- UNAIDS**, “Miles to go - Closing gaps. Breaking barriers. Righting injustices.,” *Joint United Nations Programme on HIV/AIDS*, 2018.
- , “Women and girls and HIV,” *Joint United Nations Programme on HIV/AIDS*, 2018.
- Worden, Lee and Travis C Porco**, “Products of Compartmental Models in Epidemiology,” *Computational and mathematical methods in medicine*, 2017, *vol. 2017*.
- Yi, Tae Joon, Brett Shannon, Jessica Prodger, Lyle McKinnon, and Rupert Kaul**, “Genital immunology and HIV susceptibility in young women,” *American Journal of Reproductive Immunology*, 2013, *69*, 74–79.

Women's Position in Ancestral Societies and Female HIV: The Long-Term Effect of Matrilineality in Sub-Saharan Africa

Online Appendix

Jordan Loper*

August 28, 2019

*Aix-Marseille Univ., CNRS, EHESS, Centrale Marseille, IRD, AMSE, Marseille, France (email: jordan.loper@univ-amu.fr)

Appendix A Data Description

A.1 List of the countries included in the analysis

All Demographic Health Surveys (DHS) from Sub-Saharan African countries which data on both HIV, individual's ethnicity as well as village's GPS information are included in the analysis. Based on this selection criterion, the country surveys used include 32 surveys from 18 countries as follow:

- Burkina-Faso (2003, 2010)
- Cameroon (2004, 2011)
- Chad (2014)
- Congo Democratic Republic (2007, 2013)
- Ethiopia (2005, 2011, 2016)
- Gabon (2012)
- Ghana (2003, 2014)
- Guinea (2005, 2012)
- Ivory Coast (2011)
- Kenya (2003, 2008)
- Liberia (2013)
- Malawi (2004, 2010, 2014)
- Mali (2006, 2012)
- Senegal (2005, 2010)
- Sierra Leone (2008, 2013)
- Togo (2013)
- Uganda (2011)
- Zambia (2007, 2013)

A.2 Description of the main controls

A.2.1 Individual-level Data

- **Marital status:** Categorical variable of the current marital status of the respondent (0 “*Never married*”; 1 “*Married*”; 2 “*Living together*”; 3 “*Widowed*”; 4 “*Divorced*”; and 5 “*Not living together*”). **Source:** DHS (variable v501)
- **Actual polygyny:** Dummy variable indicating whether an individual is in a polygynous union. **Source:** DHS (computed from variable v505)
- **Age:** Individual’s Age in completed years. **Source:** DHS (variable v012)
- **Age squared:** Square of the individual’s age in completed years. **Source:** DHS (computed from variable v012)
- **Number of children:** Individual’s total number of children ever born. **Source:** DHS (variable v201)
- **Urban:** Dummy variable indicating whether an individual resides in an urban location. **Source:** DHS (variable v025)
- **Education:** Individual’s education in single years **Source:** DHS (variable v133)
- **Working:** Dummy variable indicating whether an individual is currently working. **Source:** DHS (variable v714)
- **Wealth:** Categorical variable of wealth index (a composite measure of a household’s cumulative living standard) (1 “*Poorest*”; 2 “*Poorer*”; 3 “*Middle*”; 4 “*Richer*”; and 5 “*Richest*”). **Source:** DHS (variable v190)
- **Religion:** Categorical variable of religion (1 “*Christian*”; 2 “*Muslim*”; 3 “*Other religion*”; and 4 “*No religion*”). **Source:** DHS (computed and harmonized between countries by the author, from variable v130)

A.2.2 Ethnicity-level Data

These variables are computed from the *Ethnographic Atlas (E.A.)*, a worldwide anthropological database containing ethnographic information on cultural aspects and ways of life of ethnic groups prior to industrialization and colonial contact

- **Women’s historical participation in agriculture:** Variable increasing in women’s historical participation in agriculture, and ranging from 1 “*Males only or almost alone*” to 6 “*Females only or almost alone*”. **Source:** Ethnographic Atlas (variable v54)
- **Polygyny:** Dummy variable indicating whether an ethnic group practiced polygyny (v8 = 2; 4 or 5 in the E.A.). **Source:** Ethnographic Atlas (variable v8)
- **Bride Price:** Dummy variable indicating whether an ethnic group practiced bride price (v6 not equal to 6 or 7 in the E.A.). **Source:** Ethnographic Atlas (variable v6)

- **Plough:** Categorical variable of plough use (1 “*No plough*”; 2 “*Not aboriginal but well established plough use at period of observation*”; and 3 “*Aboriginal plough use prior to contact*”). **Source:** Ethnographic Atlas (variable v39)
- **Pastoralism:** Index of ethnic group’s dependence on pastoralism, computed following [Becker \(2018\)](#) by interacting the index of ethnic group dependence on animal husbandry (v4 in the E.A.) with a dummy indicating whether herd animals were the predominant ethnic group’s type of animal husbandry (i.e v40 = 3, 4, 5, 6 or 7 in the E.A.). **Source:** Ethnographic Atlas (computed from variables v4 and v40)
- **Clans:** Dummy variable indicating whether an ethnic group was organized in clan communities (v15 = 6 in the E.A.). **Source:** Ethnographic Atlas (computed from variable v15)
- **Settlement patterns:** Variable increasing in the complexity of ethnic group’s settlement patterns, ranging from 1 “*Nomadic or fully migratory*” to 8 “*Complex settlements*”. **Source:** Ethnographic Atlas (variable v30).
- **Jurisdictional Hierarchies:** Categorical variable of the number of levels in the ethnic group’s jurisdictional hierarchy beyond local community (= v33 - 1 in the E.A.). **Source:** Ethnographic Atlas (variable v33)
- **Reliance on hunting:** Categorical variable of ethnic group’s reliance on hunting, ranging from 0 “*0 - 5% dependence*” to 9 “*86 - 100% dependence*”. **Source:** Ethnographic Atlas (variable v2)
- **Reliance on fishing:** Categorical variable of ethnic group’s reliance on fishing, ranging from 0 “*0 - 5% dependence*” to 9 “*86 - 100% dependence*”. **Source:** Ethnographic Atlas (variable v3)
- **Reliance on gathering:** Categorical variable of ethnic group’s reliance on gathering, ranging from 0 “*0 - 5% dependence*” to 9 “*86 - 100% dependence*”. **Source:** Ethnographic Atlas (variable v1)
- **Reliance on animal husbandry:** Categorical variable of ethnic group’s reliance on animal husbandry, ranging from 0 “*0 - 5% dependence*” to 9 “*86 - 100% dependence*”. **Source:** Ethnographic Atlas (variable v4)
- **Reliance on agriculture:** Categorical variable of ethnic group’s reliance on agriculture, ranging from 0 “*0 - 5% dependence*” to 9 “*86 - 100% dependence*”. **Source:** Ethnographic Atlas (variable v5)
- **Large domesticated animals:** Dummy variable indicating whether an ethnic group practiced animal husbandry (v40 > 1 in the E.A.). **Source:** Ethnographic Atlas (computed from variable v40)
- **Intensity of agriculture:** Categorical variable increasing in ethnic group’s intensity of agriculture, and ranging from 1 “*No agriculture*” to 6 “*Intensive irrigated agriculture*”. **Source:** Ethnographic Atlas (variable v28)
- **Year of observation in the E.A.:** Year of observation of the ethnic group in the Ethnographic Atlas. **Source:** Ethnographic Atlas (variable v102)

A.2.3 Village-level Data

Except when otherwise indicated, most of village-level data are computed from DHS geospatial covariates. Documentation related on DHS geospatial covariates can be found here: <https://spatialdata.dhsprogram.com/covariates/>.

- **Latitude:** Latitude of DHS village (in decimal degrees). **Source:** DHS Geospatial Covariates
- **Longitude:** Longitude of DHS village (in decimal degrees). **Source:** DHS Geospatial Covariates
- **Altitude:** Altitude of DHS village (in meters). **Source:** DHS Geospatial Covariates
- **Nightlight composite:** The average nighttime luminosity, of the area within the 2 km (urban) or 10 km (rural) buffer surrounding the DHS village location (2015). **Source:** DHS Geospatial Covariates
- **Population density (2010):** The average UN-adjusted population density of the area within the 2 km (urban) or 10 km (rural) buffer surrounding the DHS village location (2010). **Source:** DHS Geospatial Covariates
- **Distance to lake or coastline:** The geodesic distance to either a lake or the coastline. **Source:** DHS Geospatial Covariates
- **Distance to international border:** The geodesic distance to the nearest international borders. **Source:** DHS Geospatial Covariates
- **Average time urban center: (2015):** The average time (minutes) required to reach a high-density urban center, from the area within the 2 km (urban) or 10 km (rural) buffer surrounding the DHS village location, based on year 2015 infrastructure data. **Source:** DHS Geospatial Covariates
- **Malaria incidence (2010):** The average number of people per year who show clinical symptoms of *Plasmodium falciparum* malaria within the 2 km (urban) or 10 km (rural) buffer surrounding the DHS village location (2010). **Source:** DHS Geospatial Covariates
- **Vegetation index (2010):** The average vegetation index value within the 2 km (urban) or 10 km (rural) buffer surrounding the DHS village (2010). **Source:** DHS Geospatial Covariates
- **Length of the growing season:** Categorical variable of the length of the growing season in days for the area within the 2 km (urban) or 10 km (rural) buffer surrounding the DHS village location; ranging from 1 “0 days” to 16 “ > 365 days”. **Source:** DHS Geospatial Covariates
- **Distance to nearest active mine:** Distance (in km) to the nearest active mine (active the same year than the village is surveyed in DHS). **Source:** Calculation from the author
- **Active mine within 1000 km:** Dummy indicating whether the distance to the nearest active mine is lower or equal to 1000 km. **Source:** Calculation from the author

- **Ethnic fractionalization:** Index of ethnic fractionalization computed at the DHS village level following the formula of [Montalvo and Reynal-Querol \(2005\)](#) detailed in [Equation 2](#) in the main text, using information on individual's (who reside in the village) ethnicity. **Source:** Calculation from the author
- **Ethnic polarization:** Index of ethnic polarization computed at the DHS village level, following the formula of [Montalvo and Reynal-Querol \(2005\)](#) detailed in [Equation 3](#) in the main text, using information on individuals' (who reside in the village) ethnicity. **Source:** Calculation from the author

Appendix B Additional Figures

Figure 5: Ancestral Ethnic Group Boundaries and Contemporaneous Gender Differences in HIV Rates

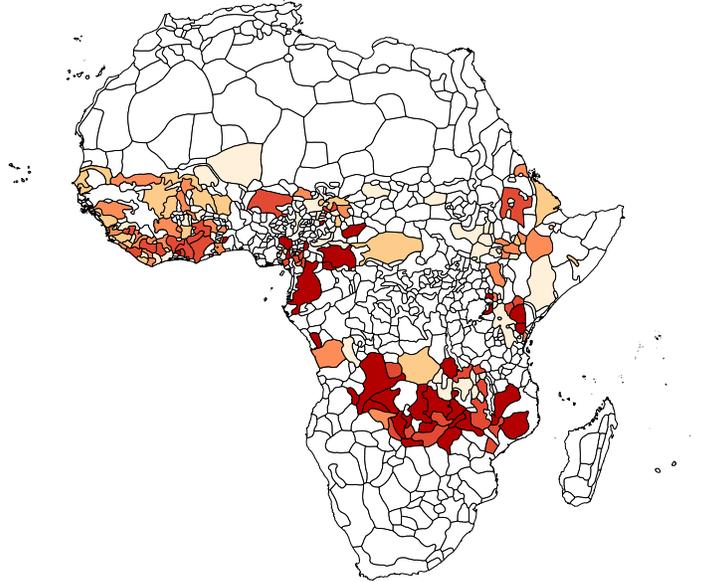


Figure 6: Ancestral Ethnic Group Boundaries and Contemporaneous Male HIV Rates

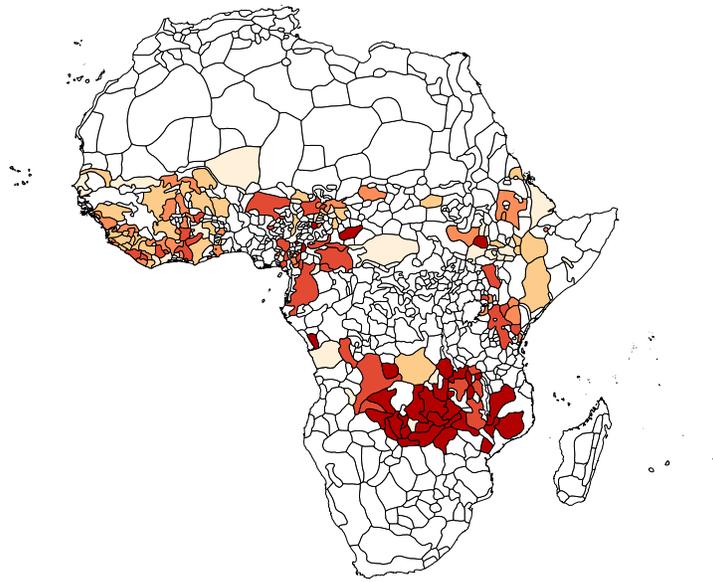
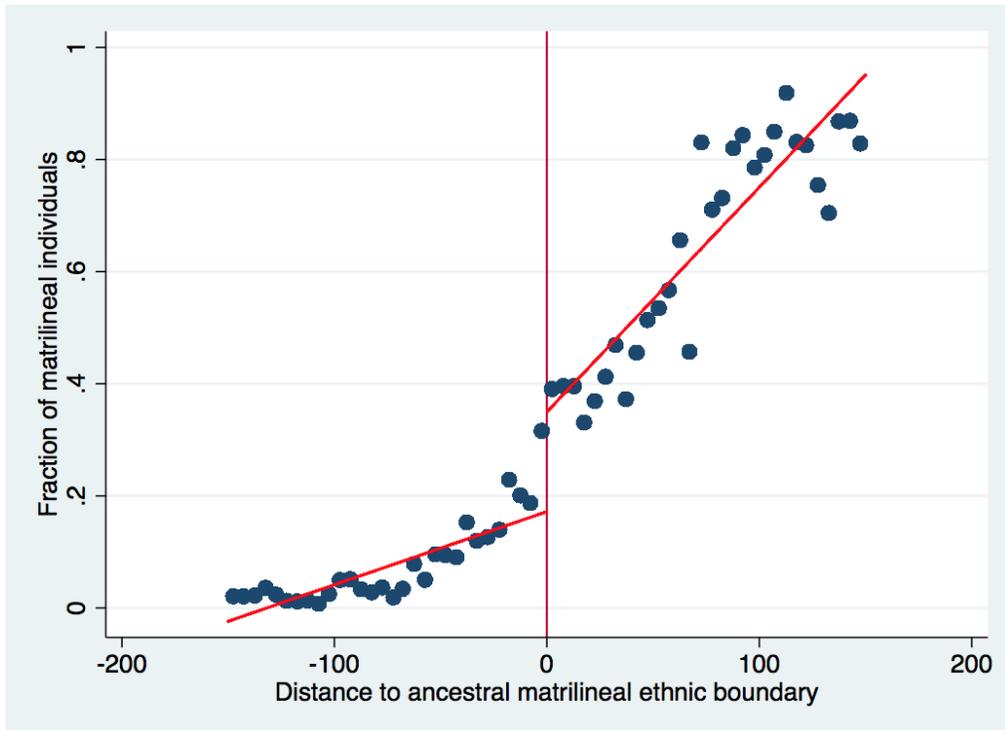


Figure 7: Location of DHS Villages

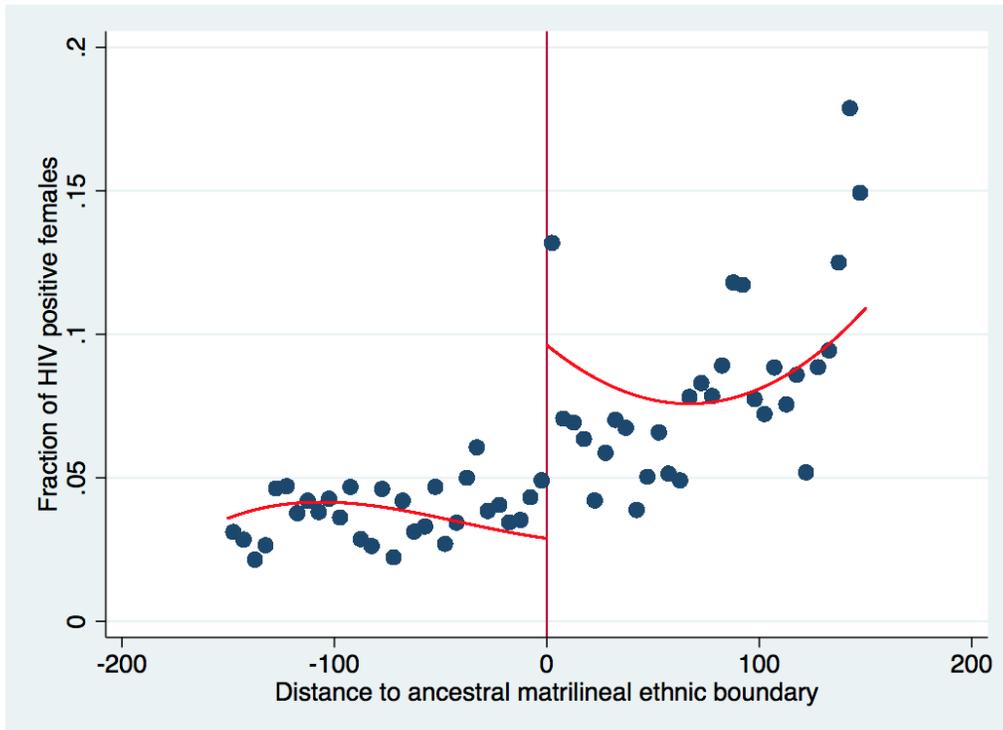


Figure 8: Individual's Matrilineality and Distance to Nearest Ancestral Matrilineal Ethnic Boundary (RDD)



This graph presents the unconditional relationship between individual's ethnic group's ancestral matrilineality and individual's DHS village geographic location, for which a linear polynomial is estimated separately at each side of the boundary. The sample is limited to individuals living in villages located within 150 km of an ancestral matrilineal ethnic boundary. The x -axis reports geographic distance. Positive values are kilometers into the territory of an ancestrally matrilineal ethnic group and negative values are kilometers into the territory of an ancestrally non-matrilineal (i.e. patrilineal or other) ethnic group. The y -axis measures the fraction of the population at each distance that originates from an ancestrally matrilineal ethnic group.

Figure 9: Female HIV Rate and Distance to Nearest Ancestral Matrilineal Ethnic Boundary (RDD)



This graph presents the relationship between female HIV rate and individual's DHS village geographic location, for a specification that conditions on region (within-country) \times survey (time) FE, and for which a cubic polynomial is estimated separately at each side of the boundary. The sample is limited to females living in villages located within 150 km of an ancestral matrilineal ethnic boundary. The x -axis reports geographic distance. Positive values are kilometers into the territory of an ancestrally matrilineal ethnic group and negative values are kilometers into the territory of an ancestrally non-matrilineal (i.e. patrilineal or other) ethnic group. The y -axis measures the fraction of the female population at each distance that is HIV positive.

Figure 10: Simulation - Scenario 2 ($\mu_m = 4$ and $\tau_m = 2$)

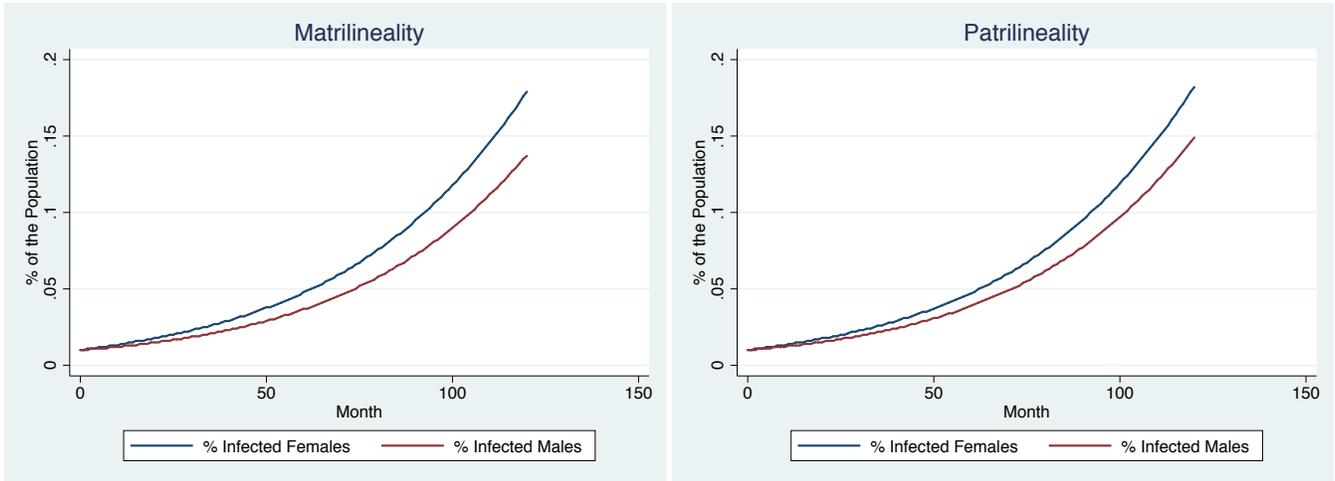
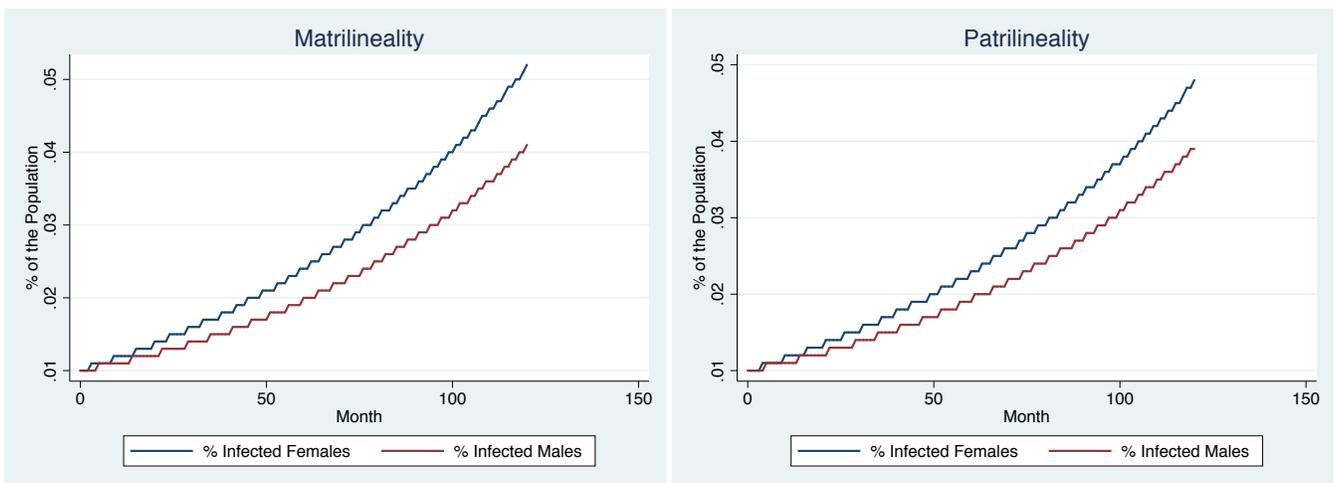


Figure 11: Simulation - Scenario 3 ($\mu_m = \tau_m = 2$)



Appendix C Additional Tables

Table 10: Summary Statistics (Final Regression Sample)

	Females			Males			All Sample		
	Matrilineal	Patrilineal	Difference	Matrilineal	Patrilineal	Difference	Matrilineal	Patrilineal	Difference
HIV	0.098 (0.194)	0.039 (0.298)	0.059*** (0.002)	0.066 (0.150)	0.023 (0.249)	0.043*** (0.002)	0.083 (0.277)	0.033 (0.177)	0.051*** (0.001)
<i>Individual-level variables</i>									
Marital status	1.204 (1.228)	1.082 (1.037)	0.122*** (0.009)	0.810 (0.934)	0.786 (0.888)	0.025** (0.008)	1.021 (1.119)	0.960 (0.989)	0.061*** (0.006)
Actual polygyny	0.106 (0.308)	0.272 (0.445)	-0.166*** (0.004)	0.055 (0.227)	0.121 (0.326)	-0.066*** (0.003)	0.082 (0.275)	0.210 (0.407)	-0.128*** (0.002)
Age	28.236 (9.410)	28.937 (9.970)	-0.701*** (0.085)	30.205 (11.498)	31.277 (12.096)	-1.073*** (0.113)	29.151 (10.478)	29.896 (10.952)	-0.745*** (0.069)
Age squared	885.820 (578.551)	936.766 (635.385)	-50.946*** (5.416)	1044.514 (776.859)	1124.572 (835.740)	-80.058*** (7.790)	959.575 (682.563)	10103.752 (730.113)	-54.177*** (4.599)
Nb. children	3.062 (2.790)	3.283 (3.126)	-0.221*** (0.027)	2.960 (3.560)	3.155 (4.114)	-0.195*** (0.038)	3.014 (3.172)	3.230 (3.565)	-0.216*** (0.022)
Urban	0.274 (0.446)	0.310 (0.462)	-0.035*** (0.004)	0.262 (0.474)	0.341 (0.440)	-0.078 (0.004)	0.269 (0.003)	0.322 (0.001)	-0.054*** (0.003)
Education	5.338 (3.956)	4.185 (4.349)	1.153*** (0.037)	6.419 (3.907)	5.516 (4.757)	0.903*** (0.044)	5.841 (3.970)	4.731 (4.567)	1.110 (0.029)
Working	0.575 (0.494)	0.636 (0.481)	-0.061*** (0.004)	0.782 (0.412)	0.788 (0.409)	-0.006 (0.004)	0.671 (0.470)	0.698 (0.459)	-0.027*** (0.003)
Wealth	3.013 (1.416)	2.997 (1.430)	0.016 (0.012)	3.038 (1.385)	2.993 (1.424)	0.045*** (0.013)	3.025 (1.402)	2.995 (1.428)	0.029** (0.009)
Religion	1.237 (0.565)	1.520 (0.651)	-0.283*** (0.006)	1.314 (0.715)	1.615 (0.689)	-0.300*** (0.007)	1.273 (0.640)	1.559 (0.668)	-0.286*** (0.004)
<i>Ethnicity-level variables</i>									
Women's participation in agri.	4.624 (1.023)	3.758 (1.236)	0.866*** (0.010)	4.647 (1.000)	3.600 (1.224)	1.047*** (0.011)	4.635 (1.012)	3.693 (1.233)	0.942*** (0.008)
Polygyny	0.770 (0.421)	0.524 (0.499)	0.247*** (0.004)	0.788 (0.409)	0.484 (0.500)	0.304*** (0.005)	0.779 (0.415)	0.508 (0.500)	0.271*** (0.003)
Bride Price	0.867 (0.339)	1.000 (0.022)	-0.132*** (0.001)	0.880 (0.325)	1.000 (0.022)	-0.119*** (0.001)	0.873 (0.332)	1.000 (0.022)	-0.126*** (0.001)
Plough	1.000 (0.000)	1.019 (0.137)	-0.019*** (0.001)	1.000 (0.000)	1.000 (0.022)	0.000 (0.000)	1.000 (0.000)	1.011 (0.106)	-0.011*** (0.001)
Pastoralism	0.087 (0.045)	0.208 (0.131)	-0.121*** (0.122)	0.089 (0.043)	0.195 (0.131)	-0.107*** (0.001)	0.088 (0.044)	0.203 (0.131)	-0.115*** (0.001)
Clans	0.234 (0.423)	0.262 (0.440)	-0.028*** (0.004)	0.223 (0.417)	0.259 (0.438)	-0.035*** (0.004)	0.229 (0.002)	0.261 (0.001)	-0.032*** (0.003)
Settlement patterns	6.679 (0.610)	6.628 (1.129)	0.051*** (0.009)	6.699 (0.599)	6.828 (1.044)	-0.129*** (0.009)	6.688 (0.605)	6.710 (1.100)	-0.022*** (0.007)
Jurisdictional hierarchies	1.470 (0.637)	1.574 (0.973)	-0.103*** (0.008)	1.509 (0.631)	1.525 (0.935)	-0.016* (0.008)	1.488 (0.635)	1.554 (0.958)	-0.066*** (0.006)
Reliance on hunting	15.451 (6.505)	8.723 (5.432)	6.729*** (0.048)	15.809 (6.585)	8.777 (5.604)	6.932*** (0.055)	15.618 (6.545)	8.786 (5.504)	6.831*** (0.036)
Reliance on fishing	14.762 (7.758)	9.229 (7.429)	5.533*** (0.065)	14.629 (7.732)	9.312 (7.517)	5.317*** (0.071)	14.700 (7.746)	9.263 (7.466)	5.437*** (0.048)
Reliance on gathering	7.198 (4.414)	5.863 (4.779)	1.336*** (0.041)	7.237 (4.357)	6.416 (4.979)	0.821*** (0.046)	7.216 (4.388)	6.089 (4.870)	1.127*** (0.031)
Reliance on animal husbandry	9.666 (2.655)	20.651 (12.684)	-10.985*** (0.101)	9.720 (2.524)	19.472 (12.571)	-9.752*** (0.108)	9.991 (2.595)	20.168 (12.651)	-10.477*** (0.074)
Reliance on agriculture	54.339 (5.659)	58.996 (9.489)	-4.657*** (0.078)	59.235 (5.552)	59.235 (10.006)	-5.266*** (0.088)	54.167 (5.613)	59.094 (9.705)	-4.927*** (0.058)
Large domesticated animals	0.839 (0.368)	0.960 (0.195)	-0.122*** (0.002)	0.856 (0.351)	0.946 (0.227)	-0.090*** (0.002)	0.847 (0.360)	0.954 (0.209)	-0.108*** (0.002)
Intensity of agriculture	3.016 (0.180)	3.449 (0.889)	-0.432*** (0.007)	3.015 (0.175)	3.348 (0.830)	-0.332*** (0.007)	3.016 (0.178)	3.407 (0.867)	-0.391*** (0.005)
Year of observation in E.A.	1918.87 (11.451)	1916.973 (18.575)	1.896*** (0.153)	1919.177 (11.297)	1917.302 (15.702)	1.875*** (0.142)	1919.013 (11.380)	1917.108 (17.455)	1.905*** (0.106)
<i>Village-level variables</i>									
Latitude	-10.973 (6.955)	5.025 (7.643)	15.998*** (0.065)	-11.046 (6.805)	5.667 (8.235)	-16.714*** (0.075)	-11.007 (6.885)	5.288 (7.898)	-16.295*** (0.049)
Longitude	27.569 (11.371)	10.705 (19.016)	16.864*** (0.156)	27.709 (11.196)	5.996 (17.501)	21.713*** (0.067)	27.634 (11.290)	8.775 (18.555)	18.860*** (0.112)
Altitude	856.539 (389.289)	656.037 (535.854)	200.502*** (4.461)	879.954 (386.619)	534.783 (516.611)	345.171*** (4.679)	867.421 (388.220)	606.332 (531.406)	261.089*** (3.252)
Nightlight composite	2.563 (6.897)	1.601 (4.658)	0.962*** (0.044)	2.448 (6.808)	1.720 (4.857)	0.728*** (0.050)	2.510 (6.856)	1.650 (4.741)	0.860*** (0.033)
Pop. density (2010)	1297.833 (5170.735)	1056.708 (3149.892)	241.125*** (30.431)	1050.502 (4312.437)	958.689 (3103.504)	91.813*** (31.634)	1182.883 (4792.504)	1016.528 (3131.321)	166.355*** (21.984)
Distance (geodesic) lake/coastline	103.597 (113.291)	117.080 (111.648)	-13.483*** (9.66)	105.490 (109.640)	131.855 (112.878)	-26.365*** (1.060)	104.477 (111.611)	123.137 (112.388)	-18.660*** (7.15)
Distance (geodesic) international border	46.026 (41.586)	81.903 (74.721)	-35.877*** (6.11)	46.250 (40.845)	81.936 (79.587)	-35.687*** (7.00)	46.130 (41.243)	81.817 (76.752)	-35.787*** (4.60)
Average time urban center (2015)	77.479 (76.205)	78.707 (96.880)	-1.228*** (0.812)	78.590 (76.041)	86.703 (108.687)	-8.113*** (0.978)	77.995 (76.130)	81.985 (101.961)	-3.989*** (0.625)
Malaria incidence (2010)	0.379 (0.171)	0.368 (0.213)	0.011*** (0.002)	0.380 (0.174)	0.349 (0.215)	0.031*** (0.002)	0.379 (0.172)	0.360 (0.214)	0.019*** (0.001)
Vegetation index (2010)	3066.037 (582.642)	3233.712 (975.880)	-167.676*** (8.012)	3077.646 (590.357)	3143.876 (987.358)	-66.230*** (8.764)	3071.432 (586.259)	3196.886 (981.593)	-125.454*** (5.914)
Length growing season	8.008 (1.511)	9.248 (2.733)	-1.240*** (0.022)	7.967 (1.481)	8.769 (2.601)	-0.802*** (0.023)	7.989 (1.497)	9.502 (2.690)	-1.062*** (0.016)
Distance mine	3480.909 (1018.664)	2399.391 (1094.061)	1081.518*** (9.352)	3494.092 (993.290)	2107.588 (1027.745)	1386.504*** (9.641)	3487.036 (5.869)	2279.775 (2.755)	1207.261 (6.785)
Mine within 1000 km	0.034 (0.182)	0.041 (0.199)	-0.007*** (0.002)	0.035 (0.185)	0.054 (0.227)	-0.019*** (0.002)	0.035 (0.183)	0.035 (0.211)	-0.012*** (0.001)
Ethnic fractionalization	0.370 (0.280)	0.270 (0.265)	-0.100*** (0.002)	0.363 (0.281)	0.270 (0.265)	0.093*** (0.003)	0.367 (0.280)	0.270 (0.265)	0.097*** (0.002)
Ethnic polarization	0.622 (0.370)	0.486 (0.393)	0.136*** (0.003)	0.611 (0.372)	0.487 (0.396)	0.125*** (0.004)	0.617 (0.371)	0.486 (0.394)	0.131*** (0.002)
Observations	15,758	90,206		13,683	62,665		29,441	152,871	

Notes: This table reports summary statistics (means; and standard deviations in parenthesis) computed on my main regression sample (column 5 in Table 1). Standard errors of t-test of equality of means are reported in parenthesis in columns reporting mean differences. A description of variables is provided in subsection A.2 in appendix

Table 11: Ancestral Matrilineality, Genetic Diversity and Father Diversification (OLS)

	Country-Level		Individual-Level	
	Mobility index-predicted genetic diversity	Mobility index-predicted genetic diversity (ancestry andjusted)	Women’s children’s number of different fathers	Women’s children’s father diversification
	(1)	(2)	(3)	(4)
Matrilineality	0.005** (0.002)	0.007* (0.004)	0.007* (0.004)	0.014** (0.006)
Controls	Yes	Yes	Yes	Yes
Continent FE	Yes	Yes	No	No
Region-survey FE	No	No	Yes	Yes
Observations	124	124	66,895	66,895
R-squared	0.994	0.897	0.023	0.333
Clusters	N.A.	N.A.	1,407	1,407
Mean Dep. Var.	0.711	0.722	1.019	0.537

*Notes: In columns 1 and 2, OLS estimates are reported with robust standard errors in brackets. The unit of observation is a country. “Matrilineality” is the estimated proportion of countries’ citizens with ancestors that had matrilineal inheritance rule (source: Giuliano and Nunn (2018)). This variable ranges from 0 to 1. Outcomes are from Ashraf and Galor (2013a) and are computed at the country level. The “Mobility index-predicted genetic diversity” is the expected heterozygosity (genetic diversity) of a given country as predicted by migratory distance (human mobility index), calculated for the journey from Addis Ababa (Ethiopia) to the country’s modern capital city. “Mobility index-predicted genetic diversity (ancestry andjusted)” is the expected heterozygosity (genetic diversity) of a given country as predicted by migratory distance (human mobility index), calculated for the journey from Addis Ababa (Ethiopia) to each of the year 1500 CE locations of the ancestral populations of the country’s component ethnic groups in 2000 CE, as well as for the journey between each pair of these ancestral populations. The Controls are from Ashraf and Galor (2013a) and Ashraf and Galor (2013b), computed at the country level and include: the “aerial” great circle distance “as the crow flies” from Addis Ababa (Ethiopia) to the country’s modern capital city; the square of this previous measure of distance; the “migratory” great circle distance from Addis Ababa (Ethiopia) to the country’s modern capital city along a land-restricted path forced through one or more of five intercontinental waypoints; the square of this previous measure of distance; the absolute value of the latitude of a country’s geodesic centroid; a geospatial index of the suitability of land for agriculture (based on ecological indicators of climate suitability for cultivation); the percentage of a country’s arable land; the country’s mean distance to nearest waterway; the country’s total land area; the average monthly temperature of a country; the average monthly precipitation of a country; the fraction of a country’s land area that is located in tropical and subtropical climate zones; the total number of different types of infectious diseases in a country; a dummy indicating whether a nation is an island; and a dummy indicating whether a nation is landlocked. In column 3 and 4, OLS estimates are reported with standard errors clustered at the ethnicity-DHS (within-country) region-survey (year) level in brackets. The unit of observation is a woman surveyed in DHS and being part of my final sample, originating from either a traditionally matrilineal or a traditionally patrilineal ethnic group. “Matrilineality” indicates (dummy) whether a woman belongs to a traditionally matrilineal ethnic group. “Women’s children’s number of different fathers” is the number of different males with whom a female had her biological children (aged less than 17 and member of the household at the time of the survey). “Women’s children’s father diversification” is the number of different males with whom a female had her biological children divided by the number of a female’s biological children (aged less than 17 and member of the household at the time of the survey). The controls consist of the Individual controls, the (Ancestral) Ethnic Group Controls as well as the Village-Geographic Controls defined in subsection 3.3 and Table 1. The number of female’s biological children is not controlled for anymore in column 4. Region-survey is a subnational region defined in DHS, interacted with its survey-year. In columns 3 and 4, “R-squared” reported is adjusted; and “Mean Dep. Var.” reported is the mean of patrilineal women in the regression samples. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$*

Table 12: Selection And Falsification Tests

	Selection		Falsification Tests		
	Consent HIV Test (1)	Take HIV Test (2)	Anemia (3)	BMI (4)	Rohrer Index (5)
Matrilineality	-0.028*** (0.007)	-0.030*** (0.008)	-0.008 (0.018)	-0.105 (0.107)	-0.042 (0.071)
Female	0.018*** (0.002)	0.020*** (0.002)			
Female × Matrilineality	-0.000 (0.004)	0.001 (0.005)			
Ind. Controls	Yes	Yes	Yes	Yes	Yes
Ethnic Group Controls	Yes	Yes	Yes	Yes	Yes
Village-Geographic Controls	Yes	Yes	Yes	Yes	Yes
Region-survey FE	Yes	Yes	Yes	Yes	Yes
Gender	Both	Both	Female	Female	Female
Observations	145,619	145,877	69,999	83,908	83,908
Adj. R-squared	0.053	0.053	0.062	0.197	0.167
Clusters	1,204	1,204	1,027	1,324	1,324
Mean Dep. Var. (Patri. Males)	0.941	0.934	0.473	22.195	13.980

Notes: OLS estimates are reported with standard errors clustered at the ethnicity-DHS (within-country) region-survey (year) level in brackets. The unit of observation is an individual originating from either a traditionally matrilineal or a traditionally patrilineal ethnic group. “Matrilineality” indicates (dummy) whether an individual belongs to a traditionally matrilineal ethnic group. “Female” indicates (dummy) whether an individual is a female. “**Female × Matrilineality**” indicates (dummy) whether an individual is a female belonging to a traditionally matrilineal ethnic group. “Consent HIV Test” is a dummy indicating whether an individual consented to take the DHS HIV test. “Take HIV Test” is a dummy indicating whether an individual actually took the DHS HIV test. “Anemia” is a dummy indicating whether an individual has any level of anemia (available only for women). “BMI” is the Body Mass Index (available only for women). “Rohrer Index” is available only for women. Controls are defined in [Table 1](#). *Region-survey* is a subnational region defined in DHS, interacted with its survey-year. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 13: **Heterogeneous Effects by Subsamples: Common Law vs. Civil Law Countries / Polygynous vs. Non Polygynous Individuals**

Sample	HIV			
	Common Law Countries (1)	Civil Law Countries (2)	Ind. not in Polygynous Union (3)	Ind. in Polygynous Union (4)
Matrilineality	-0.001 (0.009)	-0.002 (0.007)	-0.000 (0.005)	0.002 (0.019)
Female	0.009*** (0.002)	0.008*** (0.001)	0.011*** (0.001)	-0.001 (0.003)
Female × Matrilineality	0.015*** (0.006)	0.002 (0.005)	0.014*** (0.004)	0.007 (0.013)
Ind. Controls	Yes	Yes	Yes	Yes
Ethnic Group Controls	Yes	Yes	Yes	Yes
Village-Geographic Controls	Yes	Yes	Yes	Yes
Region-survey FE	Yes	Yes	Yes	Yes
Observations	81,422	100,890	147,799	34,513
Adj. R-squared	0.089	0.027	0.083	0.060
Clusters	747	941	1,653	1,059
Mean Dep. Var. (Patri. Males)	0.043	0.014	0.023	0.024
Mean Female x Matri.	0.158	0.029	0.095	0.049
Std. Dev. Female x Matri.	0.365	0.167	0.294	0.215

Notes: OLS estimates are reported with standard errors clustered at the ethnicity-DHS (within-country) region-survey (year) level in brackets. The unit of observation is an individual originating from either a traditionally matrilineal or a traditionally patrilineal ethnic group. *Common Law Countries* and *Civil Law Countries* classification is from [La Porta et al. \(2008\)](#) dataset. “Matrilineality” indicates (dummy) whether an individual belongs to a traditionally matrilineal ethnic group. “Female” indicates (dummy) whether an individual is a female. “**Female × Matrilineality**” indicates (dummy) whether an individual is a female belonging to a traditionally matrilineal ethnic group. “**HIV**” is a dummy indicating whether an individual is HIV positive (from DHS HIV Tests). Controls are defined in [Table 1](#). Actual polygyny is not included in the controls in columns 3 and 4. *Region-survey* is a subnational region defined in DHS, interacted with its survey-year. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 14: **Considering Ancestral Matrilocality**

Sample	HIV		
	Full (Controlling for Matrilocality)	Not Matrilocal	Matrilocal
	(1)	(2)	(3)
Matrilineality	-0.003 (0.006)	0.001 (0.007)	
Female	0.008*** (0.001)	0.009*** (0.001)	-0.027** (0.011)
Female × Matrilineality	0.017*** (0.004)	0.014** (0.006)	0.048*** (0.012)
Ind. Controls	Yes	Yes	Yes
Ethnic Group Controls	Yes	Yes	Yes
Village-Geographic Controls	Yes	Yes	Yes
Region-survey FE	Yes	Yes	Yes
Ancestral Matrilocality	Yes		
Observations	182,312	160,079	22,233
Adj. R-squared	0.078	0.067	0.107
Clusters	1,688	1,494	194
Mean Dep. Var. (Patrilineal Males)	0.023	0.023	0.021
Mean Female x Matrilineality	0.086	0.031	0.488
Std. Dev. Female x Matrilineality	0.281	0.172	0.500

Notes: OLS estimates are reported with standard errors clustered at the ethnicity-DHS (within-country) region-survey (year) level in brackets. The unit of observation is an individual originating from either a traditionally matrilineal or a traditionally patrilineal ethnic group. The sample in column 2 is composed of individuals originating from an ethnic group which was not ancestrally matrilocal; while the sample in column 3 is composed of individuals originating from an ethnic group which was ancestrally matrilocal. “Matrilineality” indicates (dummy) whether an individual belongs to a traditionally matrilineal ethnic group (this is dropped in column 3 for collinearity reason). “Female” indicates (dummy) whether an individual is a female. “**Female × Matrilineality**” indicates (dummy) whether an individual is a female belonging to a traditionally matrilineal ethnic group. “**HIV**” is a dummy indicating whether an individual is HIV positive (from DHS HIV Tests). Controls are defined in Table 1. Ancestral matrilocality is additionally controlled for in column 1. *Region-survey* is a subnational region defined in DHS, interacted with its survey-year. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 15: Assessing the importance of bias from unobservables by controlling for observable characteristics

Robustness Tests:		(1)	(2)	(3)	(4)	(5)	(6)
		Coeff. Ratio Test (Altonji et al. (2005))	Minimum Coeff. Lower Bound (Oster (2017))			$\hat{\beta}_F - /+ 2.8 \text{ S.E.}$	
Controls in Restricted (R) set	Controls in Full (F) set		$R^2_{Max} =$ $\min(\mathbf{1.3}\hat{R}_F^2; 1)$	$R^2_{Max} =$ $\min(\mathbf{1.5}\hat{R}_F^2; 1)$	$R^2_{Max} =$ $\min(\mathbf{2}\hat{R}_F^2; 1)$	$R^2_{Max} = \mathbf{1}$	
Matrilineality; Female	Full set of controls from Equation 1	15.79	0.017	0.017	0.018	0.031	[0.004 ; 0.029]
Matrilineality; Female; Region-survey FE	Full set of controls from Equation 1	-4.89	0.014	0.012	0.008	-0.088	[0.004 ; 0.029]
Matrilineality; Female; Region-survey FE; Ind. Controls	Full set of controls from Equation 1	-509.27	0.016	0.016	0.015	-0.001	[0.004 ; 0.029]
Matrilineality; Female; Region-survey FE; Ind. Controls; Ethnic Group Controls	Full set of controls from Equation 1	167.42	0.019	0.020	0.024	0.099	[0.004 ; 0.029]

Notes: Each cell in column 1 report ratios based on the coefficient of *Matrilineality* \times *Female* in two regressions; in one regression a “restricted” set of controls is included and in the other, a “full” set of controls is included. In both regressions, the sample sizes are the same. The controls included in each set are listed on the left side of the table (see Table 1 for a full description of the full set of controls from Equation 1). If $\hat{\beta}_R$ is the coefficient in the restricted set and $\hat{\beta}_F$ is the coefficient in the full set, then the ratio is $\hat{\beta}_F / (\hat{\beta}_R - \hat{\beta}_F)$ (see Altonji et al. (2005)). Each cell in columns 2-5 report bias-adjusted coefficient lower bounds of *Matrilineality* \times *Female* based on Oster (2017). If \hat{R}_R^2 is the R^2 of the regression with the restricted set of controls, and \hat{R}_F^2 is the R^2 of the regression with the full set of controls, then the minimum coefficient lower bound is: $\hat{\beta}_F - (\hat{\beta}_R - \hat{\beta}_F) \times ((R^2_{Max} - \hat{R}_F^2) / (\hat{R}_F^2 - \hat{R}_R^2))$. Column 6 reports the bounds of the 99.5% confidence interval of the fully controlled estimate of *Matrilineality* \times *Female*.

Table 16: Ancestral Matrilineality and Fertility (OLS)

	Want another child (1)	Nb. children (2)
Matrilineality	-0.070*** (0.012)	-0.129*** (0.047)
Female	-0.247*** (0.006)	0.188*** (0.023)
Female × Matrilineality	0.127*** (0.014)	0.153*** (0.038)
Ind. Controls	Yes	Yes
Ethnic Group Controls	Yes	Yes
Village-Geographic Controls	Yes	Yes
Region-survey FE	Yes	Yes
Observations	120,501	182,312
Adj. R-squared	0.310	0.643
Clusters	1,497	1,688
Mean Dep. Var. (Patri. Males)	0.787	3.155

Notes: OLS estimates are reported with standard errors clustered at the ethnicity-DHS (within-country) region-survey (year) level in brackets. The unit of observation is an individual originating from either a traditionally matrilineal or a traditionally patrilineal ethnic group. “Matrilineality” indicates (dummy) whether an individual belongs to a traditionally matrilineal ethnic group. “Female” indicates (dummy) whether an individual is a female. “**Female × Matrilineality**” indicates (dummy) whether an individual is a female belonging to a traditionally matrilineal ethnic group. “Want another child” is a dummy indicating whether the individual wants another child. “Nb. children” is the number of biological children ever born. Controls are defined in [Table 1](#) (However, number of children is not controlled for anymore). *Region-survey* is a subnational region defined in DHS, interacted with its survey-year. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 17: **Heterogeneity in Condom Use by Perception of Risk of HIV Transmission (OLS)**

	Condom							
	Females				Males			
	Full sample		Ever tested	Never tested	Full sample		Ever tested	Never tested
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Matrilineality × Condom reduces HIV	-0.005				0.009			
	(0.006)				(0.011)			
Matrilineality × Condom not reduces HIV	-0.011*				-0.006			
	(0.006)				(0.011)			
Matrilineality × HIV		0.021**	0.034***	-0.003		0.048***	0.084***	-0.008
		(0.008)	(0.012)	(0.010)		(0.016)	(0.025)	(0.021)
Matrilineality × No HIV		-0.007	-0.003	-0.005		0.003	-0.008	0.006
		(0.005)	(0.008)	(0.006)		(0.010)	(0.018)	(0.011)
Ind. Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ethnic Group Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Village-Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region-survey FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	72,855	89,081	30,795	54,814	52,796	58,094	17,683	39,954
Adj. R-squared	0.073	0.073	0.079	0.067	0.237	0.239	0.240	0.231
Clusters	1,373	1,416	1,104	1,205	1,075	1,083	888	967
Mean Dep. Var. (Patri.)	0.038	0.031	0.055	0.021	0.151	0.138	0.236	0.106
F-Test Equality of coeff. (p-value)	0.117	0.000	0.000	0.356	0.009	0.001	0.000	0.472

Notes: OLS estimates are reported with standard errors clustered at the ethnicity-DHS (within-country) region-survey (year) level in brackets. The unit of observation is an individual originating from either a traditionally matrilineal or a traditionally patrilineal ethnic group. “Matrilineality” indicates (dummy) whether an individual belongs to a traditionally matrilineal ethnic group. “Condom (not) reduces HIV” is a dummy indicating whether an individual thinks that always using condoms during sex (does not) reduces chance of getting HIV. “(No) HIV” is a dummy indicating whether an individual is HIV (negative) positive according to DHS test. “Condom” is a dummy indicating whether an individual reports “condom (male)” as her current contraception method. Controls are defined in Table 1. *Region-survey* is a subnational region defined in DHS, interacted with its survey-year. Samples “Ever tested” in column 3 and 7 consist of individuals who have ever been tested for AIDS before DHS survey. Samples “Never tested” in column 4 and 8 consist of individuals who have never been tested for AIDS before DHS survey. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 18: Ancestral Matrilineality and Couples HIV Discordance (OLS)

	Dep. Var.: Couple's Serostatus							
	Wife - / Husband -		Wife + / Husband -		Wife - / Husband +		Wife + / Husband +	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Matrilineality	0.0103 (0.0127)	0.0107 (0.0136)	0.0133* (0.0076)	0.0131 ^a (0.0081)	-0.0018 (0.0073)	0.0025 (0.0072)	-0.0218*** (0.0076)	-0.0263*** (0.0081)
Ind. Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ethnic Group Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Village-Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region-survey FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Husband's Controls		Yes		Yes		Yes		Yes
Observations	25,272	23,917	25,272	23,917	25,272	23,917	25,272	23,917
Adj. R-squared	0.100	0.106	0.021	0.024	0.027	0.030	0.056	0.061
Clusters	1,100	1,043	1,100	1,043	1,100	1,043	1,100	1,043
Mean Dep. Var. (Patri.)	0.955	0.957	0.014	0.013	0.016	0.015	0.015	0.015
Prop. Mixed Ethnicity	0.193	0.156	0.193	0.156	0.193	0.156	0.193	0.156
Prop. Mixed Matrilineality	0.048	0.044	0.048	0.044	0.048	0.044	0.048	0.044

Notes: OLS estimates are reported with standard errors clustered at the ethnicity-DHS (within-country) region-survey (year) level in brackets. The unit of observation is a non-polygynous formally married couple with both wife and husband originating from either a traditionally matrilineal or a traditionally patrilineal ethnic group. "Matrilineality" is a dummy indicating whether the wife belongs to a traditionally matrilineal ethnic group. "Wife - / Husband -" is a dummy indicating whether both the wife and the husband are HIV negative. "Wife + / Husband -" is a dummy indicating whether the wife is HIV positive while the husband is HIV negative. "Wife - / Husband +" is a dummy indicating whether the wife is HIV negative while the husband is HIV positive. "Wife + / Husband +" is a dummy indicating whether both the wife and the husband are HIV positive. *Individual controls*, *(Ancestral) Ethnic Group Controls* and *Village-Geographic Controls* are computed for the wife and are defined in Table 1. *Husband's Controls* consist of the *Individual controls*, *(Ancestral) Ethnic Group Controls* and *Village-Geographic Controls* defined in Table 1 and computed for the husband, in addition of a dummy indicating whether the wife and the husband originate from a different Ethnographic Atlas ancestral ethnic group ("*Mixed ethnicity*"), as well as a dummy indicating whether the wife and the husband originate from ethnic groups with different ancestral kinship organizations (matrilineal vs. patrilineal) ("*Mixed matrilineality*"). *Region-survey* is a subnational region defined in DHS, interacted with its survey-year. Note: ^a $p=0.103$, * $p<0.10$, ** $p<0.05$, *** $p<0.01$

Table 19: **Ancestral Matrilineality and Sexual Debuts (OLS)**

	Age first sex (1)	Age first marriage (2)
Matrilineality	-0.783*** (0.103)	-0.739*** (0.125)
Female	-1.807*** (0.101)	-4.947*** (0.097)
Female × Matrilineality	1.257*** (0.143)	1.405*** (0.143)
Ind. Controls	Yes	Yes
Ethnic Group Controls	Yes	Yes
Village-Geographic Controls	Yes	Yes
Region-survey FE	Yes	Yes
Observations	150,197	128,860
Adj. R-squared	0.238	0.444
Clusters	1,659	1,615
Mean Dep. Var. (Patri. Males)	18.602	23.996

Notes: OLS estimates are reported with standard errors clustered at the ethnicity-DHS (within-country) region-survey (year) level in brackets. The unit of observation is an individual originating from either a traditionally matrilineal or a traditionally patrilineal ethnic group. “Matrilineality” indicates (dummy) whether an individual belongs to a traditionally matrilineal ethnic group. “Female” indicates (dummy) whether an individual is a female. “**Female × Matrilineality**” indicates (dummy) whether an individual is a female belonging to a traditionally matrilineal ethnic group. “Age first sex” is the age at first sexual intercourse. “Age first marriage” is the age at start of first marriage or union. Controls are defined in [Table 1](#). *Region-survey* is a subnational region defined in DHS, interacted with its survey-year. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 20: **Ancestral Matrilineality, Acknowledgment of HIV Risks and Access to Condom (OLS)**

	Acknowledgment of Risks				Access to condom	
	Heard of AIDS	Heard of STI	Think condom reduces HIV	Think having one partner reduces HIV	Know a source to get condom	Can get condom
	(1)	(2)	(3)	(4)	(5)	(6)
Matrilineality	-0.007*	-0.006	0.005	-0.002	-0.026**	0.007
	(0.004)	(0.004)	(0.011)	(0.008)	(0.011)	(0.012)
Female	-0.027***	-0.025***	-0.048***	-0.031***	-0.191***	-0.286***
	(0.004)	(0.003)	(0.008)	(0.005)	(0.009)	(0.012)
Female × Matrilineality	0.022***	0.021***	0.007	-0.007	0.092***	0.037**
	(0.005)	(0.004)	(0.016)	(0.012)	(0.015)	(0.018)
Ind. Controls	Yes	Yes	Yes	Yes	Yes	Yes
Ethnic Group Controls	Yes	Yes	Yes	Yes	Yes	Yes
Village-Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes
Region-survey FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	182,284	182,218	156,133	164,673	164,368	103,316
Adj. R-squared	0.121	0.104	0.048	0.051	0.337	0.161
Clusters	1,688	1,688	1,670	1,674	1,623	1,539
Mean Dep. Var. (Patri. males)	0.966	0.974	0.852	0.913	0.761	0.878

Notes: OLS estimates are reported with standard errors clustered at the ethnicity-DHS (within-country) region-survey (year) level in brackets. The unit of observation is a woman originating from either a traditionally matrilineal or a traditionally patrilineal ethnic group. “Matrilineality” indicates (dummy) whether an individual belongs to a traditionally matrilineal ethnic group. “Female” indicates (dummy) whether an individual is a female. “**Female × Matrilineality**” indicates (dummy) whether an individual is a female belonging to a traditionally matrilineal ethnic group. “Heard of AIDS” is a dummy indicating whether an individual has ever heard of AIDS. “Heard of STI” is a dummy indicating whether an individual has ever heard of any STI. “Think condom reduces HIV” is a dummy indicating whether an individual thinks that always using condoms during sex reduces chance of getting HIV. “Think having one partner reduces HIV” is a dummy indicating whether an individual thinks that having only one sexual partner reduces chance of getting HIV. “Know source to get condom” is a dummy indicating whether an individual knows a source to get male condoms. “Can get condom” is a dummy indicating whether an individual can get herself a male condom. Controls are defined in Table 1. *Region-survey* is a subnational region defined in DHS, interacted with its survey-year. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 21: Summary of Model's Parameters

Parameter	Value	Interpretation	Source
α_f^M	0.076	<i>Matrilineal</i> female's number of extramarital partners	The author (Table 7)
α_f^P	0.008	<i>Patrilineal</i> female's number of extramarital partners	The author (Table 7)
α_m	0.204	Male's number of extramarital partners	The author (Table 7)
γ_f^M	0.992	<i>Matrilineal</i> female's probability of not using condom	The author (Table 7)
γ_f^P	0.976	<i>Patrilineal</i> female's probability of not using condom	The author (Table 7)
γ_m	0.931	Male's probability of not using condom	The author (Table 7)
ρ_m	0.0045	Male's contamination risk for one-time unprotected sex	Greenwood et al. (2019)
ω_f	1.75	Gender difference in biological susceptibility to contract the virus	Greenwood et al. (2019)
δ_f^M	0.007812	<i>Matri.</i> female's proba. of infection per each contact with a HIV+ male ($= \gamma_f^M \times \rho_m \times \omega_f$)	The author (Calculation)
δ_f^P	0.007686	<i>Patri.</i> female's proba. of infection per each contact with a HIV+ male ($= \gamma_f^P \times \rho_m \times \omega_f$)	The author (Calculation)
δ_m	0.00419	Male's probability of infection per each contact with a HIV+ female ($= \gamma_m \times \rho_m$)	The author (Calculation)
μ_m	4	Male's nb. of sexual intercourses with LT committed partner	The author (Assumption)
μ_f	4	Female's nb. of sexual intercourses with LT committed partner	The author (Assumption)
τ_m	4	Male's nb. of sexual intercourses with each extramarital partner	The author (Assumption)
τ_f^M	4	<i>Matrilineal</i> female's nb. of sexual intercourses with each extramarital partner	The author (Assumption)
τ_f^P	3.836	<i>Patrilineal</i> female's nb. of sexual intercourses with each extramarital partner ($= 0.959 \times \tau_f^M$)	The author (Assumption)
$i_{f,0}$	0.01	Initial female's HIV rate	The author (Assumption)
$i_{m,0}$	0.01	Initial male's HIV rate	The author (Assumption)

Table 22: **Simulated HIV Rates (Alternative Scenarios)**

Scenario 2 ($\mu_m = 4$ and $\tau_m = 2$)							
Months	<i>Matrilineality</i>			<i>Patrilineality</i>			Raw
	Females	Males	Difference	Females	Males	Difference	Diff-in-Diff
1	0.010	0.010	0.000	0.010	0.010	0.000	0.000
25	0.020	0.017	0.003	0.020	0.016	0.004	-0.001
50	0.038	0.029	0.009	0.036	0.029	0.007	0.002
75	0.067	0.052	0.015	0.064	0.050	0.014	0.001
100	0.118	0.090	0.028	0.110	0.086	0.024	0.004
120	0.179	0.137	0.042	0.167	0.131	0.036	0.006
Scenario 3 ($\mu_m = \tau_m = 2$)							
Months	<i>Matrilineality</i>			<i>Patrilineality</i>			Raw
	Females	Males	Difference	Females	Males	Difference	Diff-in-Diff
1	0.010	0.010	0.000	0.010	0.010	0.000	0.000
25	0.015	0.013	0.002	0.014	0.013	0.001	0.001
50	0.021	0.017	0.004	0.020	0.017	0.003	0.001
75	0.029	0.023	0.006	0.028	0.023	0.004	0.002
100	0.040	0.032	0.008	0.037	0.031	0.006	0.002
120	0.052	0.041	0.011	0.048	0.039	0.008	0.003

Notes: This table reports the simulated HIV rates, based on the compartmental SI epidemic model detailed in [section 6](#), assuming that $\mu_m = 4$ and $\tau_m = 2$ in the upper part of the table; and that $\mu_m = \tau_m = 2$ in the lower part of the table.