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

Jordan Loper†

May 22, 2019

Most Recent Version Here

PRELIMINARY DRAFT, PLEASE DO NOT CIRCULATE

Abstract

While 80 percent of all HIV-positive women in the world live in sub-Saharan Africa, there is substantial variation in infection rates across the continent. This paper traces current variation in female HIV rates within Sub-Saharan Africa to women's position in pre-colonial ethnic group's kinship organizations. In matrilineal kinship systems, lineage and inheritance are traced through female members. The structure of matrilineal kinship systems implies that, relative to patrilineal kinship systems, women have greater support from their own kin groups, and husbands have less authority over their wives. Using within-country variation across about 280,000 individuals in their ethnic group's ancestral kinship organization, in 18 countries, I show that females originating from ancestrally matrilineal ethnic groups are today more likely to be infected by HIV. This result 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 further show that my result is not driven by differential selection into HIV testing, nor by differences in general health status, and that omitted variable bias is very unlikely. Computing distance from DHS villages to nearest ancestral matrilineal ethnic boundary, I use a number of alternative estimation strategies, including and instrumental variable (IV) strategy and a geographic regression discontinuity design (RDD) at the ancestral matrilineal ethnic boundary, and find robust estimates. I provide several mechanisms to this finding: first, benefiting from more sexual autonomy, matrilineal women adopt riskier sexual practices which are more conducive to HIV and suggest channels of transmission of the virus alternative to simple infection by the husband. Second, despite a better ability to impose safe sexual practices to their husbands, matrilineal women are less likely to use protective contraception methods such as male condom. I provide evidences that behavioural mechanisms are driving this counterintuitive result: I find that, relative to their patrilineal counterparts, matrilineal women are more likely to (1) perceive condom as an *ex-post protective* method against HIV *transmission* rather than an *ex-ante preventive* method against HIV *infection*; (2) underestimate their husband's propensity to be unfaithful. I conjecture that this latter result is an adverse effect of matrilineal women's overestimation of the efficiency of their intra-household threatening strategy against male infidelity.

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 Guirkinger, Joseph Henrich, Asim Khwaja, Jacob Moscona, Marc Sangnier, Heather Sarsons, Jonathan Schulz, Liyang Sun, Michèle Tertilt, Edoardo Teso, Romain Wacziarg and participants at 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 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 HIV[1], representing about 80% of the worldwide female HIV positive population, Sub-Saharan Africa is the most affected region in the world[2]. 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 population[3], and adult females represented about 57% of new adult infections in 2017[4]. This phenomenon has been referred to "Feminization of HIV" in Sub-Saharan Africa. Widely-aknowledged, this phenomenon has very recently raised researchers' interest in understanding its long-term origins and subsequent variation across the continent. However, evidences are still limited.

Institutions (UNAIDS in particular), 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 barganing power and ability to impose safe sexual pratice 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 (see section 2 for more details). Indeed, matrilineality, an ethnic group's inheritence rule according to which the tracing of kinship is made through the female line, is correlated with forms of social organizations where women benefit from higher social status, as well as more autononmy and barganing power, as compared to their patrilineal counterparts (more details are provided in section 2).

In this paper, exploiting variation in Sub-Saharan pre-industrial/pre-colonial ethnic group's inheritence norm and associated kinship organization, I look at ancestral matrilineality as a long-term determinant of female HIV infection. This paper argues that this counterintuitive pattern is most likely the result of relatively riskier female' sexual behaviour and relatively lower female's use of protective contraceptive methods due to fertility preferences and underestimation of male infidelity, in ancestrally matrilineal ethnic groups.

My key empirical strategy is to exploit within subnational regions ethnic groups variation and to link it with variation in ancestral inheritence norm and associated kinship organization (matrilineality vs. patrilineality). In particular, using 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

[1] 11.2 millions in "Eastern and Southern Africa" and 3.3 millions in "West and Central Africa"
[2] These numbers are from http://aidsinfo.unaids.org/
[3] 9.5 millions males lived with HIV in Sub-Saharan Africa
[4] Among the 1.02 millions newly infected adult individuals in Sub-Saharan Africa in 2017, 580,000 were females.

for some countries, amounting to 32 *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 industrilization 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 the main results, and has nothing to do with the channel of causality argued for in this paper. To address these concerns, my key identication strategy exploits the variation in ethnic groups' ancestral kinship organizations within subnational regions in Sub-Saharan countries, which enable me to include subnational region-survey fixed effects into my estimations. This notably allows me to control for a large set of potential national and subnation 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 potential differential socio-economic composition of matrilineal ethnic groups' individuals, in particular in education, marital behaviour and religion, which are plausible drivers of sexual and contraceptive behaviour (Case and Paxson (2013)). Additionaly, 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 which may be other long-term confounding factors affecting contemporaneous gender norms and subsequent female's sexual and contraception behaviour. Further, 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 trasmitted 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. In addition, 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 based on a large array of geographic observables. All these exercises provide very little support for the presence of an omitted variable bias in my OLS estimates.

Going beyond in correcting for potential omitted variable bias, I compute the minimum distance between the location of DHS villages and the nearest ancestral matrilineal ethnic boundary (based on digitized Murdock's map of ancestral ethnic groups in Africa, see Figure 2), and I implement an instrumental variable (IV) strategy, instrumenting for individual's ethnic group's ancestral matrilineality using such measure of geographical distance. This strategy provides additional support to the main findings found in my OLS estimations.

Finally, I conduct a geographic regression discontinuity (RD) identification strategy, exploiting the measure of distance that I computed and restricting my attention to individuals living in DHS villages which are located close to an ancestral matrilineal ethnic boundary, where 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 ancestral matrilineality as a driver of contemporaneous female HIV rates.

To explain such long-term effect of ancestral matrilineality on contemporaneous female HIV in Sub-Saharan Africa, 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 virus, *as compared to their patrilineal counterparts*. In line with this findings, I 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 which suggests alternative channels of transmission of the virus to the one highlighted by Anderson (2018) (see below), and therefore calls for complementary policies. Additionaly, I underline a negative effect of ancestral matrilineality on female's protective contraception use, despite matrilineal women's higher ability to impose safe sexual practice to their husbands. I show that this counterintuitive result can be explained by both matrilineal women being more likely to desire having more child, and matrilineal women's higher propensity of underestimating male's infidelity, *relative to patrilineal women*. I conjecture that this *relative* underestimation is an adverse effect of matrilineal women's better marriage outside options, which make them wrongly overestimate the efficiency of their intrahousehold threatening strategy against male infidelity.

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-Saharah 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 barganing 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. (ming) 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 *Garo* 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 transatalantic 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 papers also links more directly to the burgenoning 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. (2018) 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 and robustness checks are presented in section 4. I then explore the mechanisms in section 6, and I dig deeper into the discussion on condom use in section 7. Finally, I provide concluding thoughts in section 8.

2 Context and Conceptual Framework

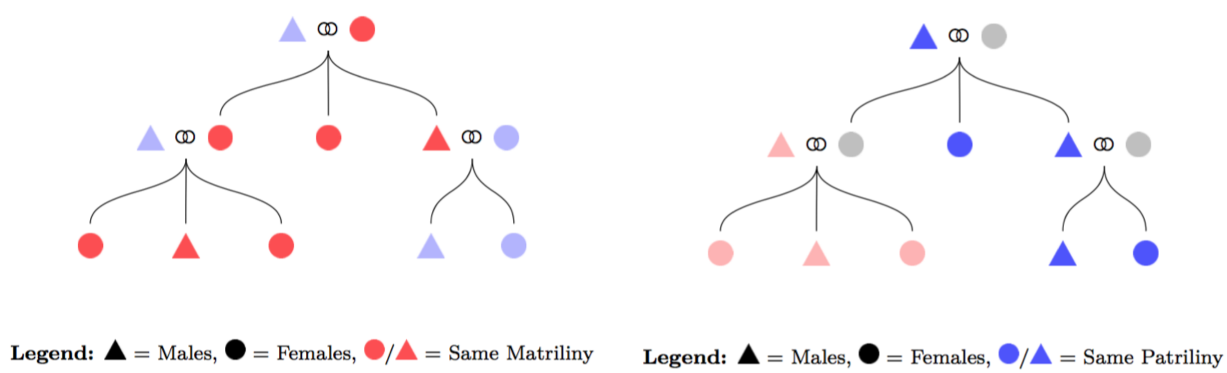
2.1 Ancestral Matrilineality in Sub-Saharan Africa

In matrilineal kinship systems, 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 - for all married couples one spouse is red and the other spouse is blue.

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))



(a) Matrilineal Kinship   (b) Patrilineal Kinship

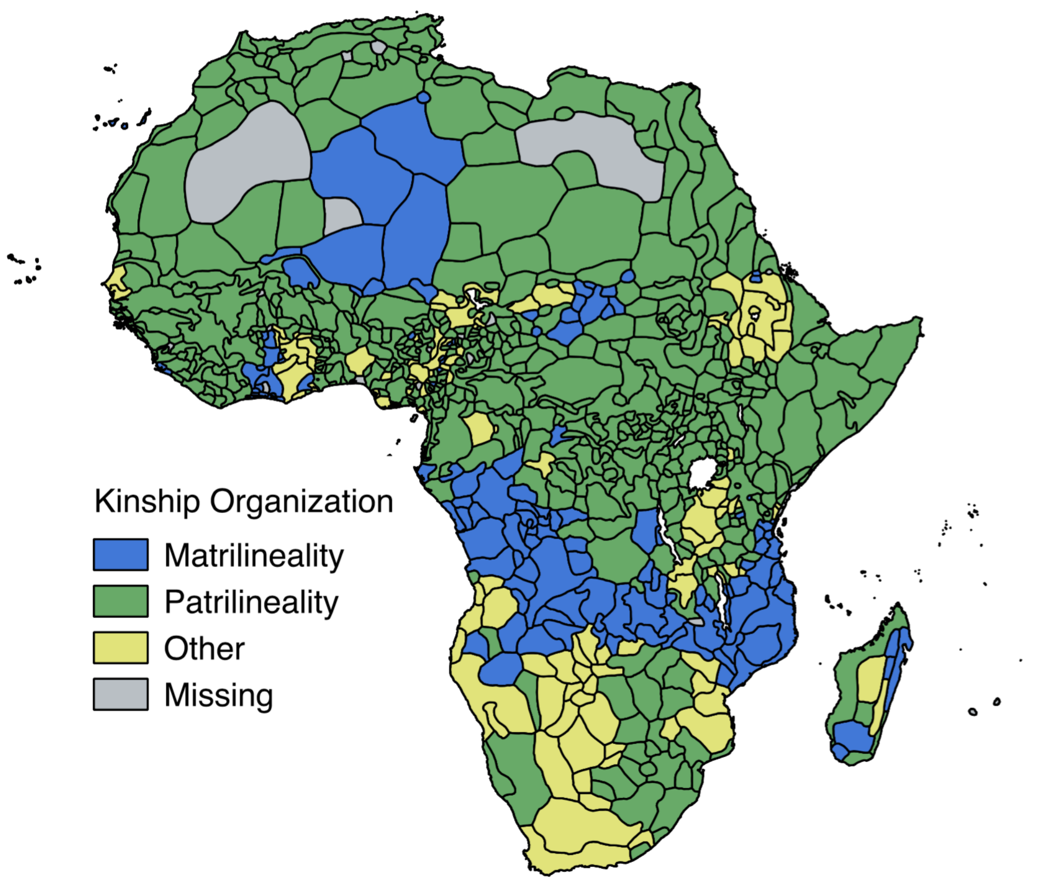
Historically, matrilineal kinship systems are correlated with other cultural traits, which have been shown to have long-term impact on gender roles and economic development. For example, Lowes (2018a) shows that, in Africa[5], matrilineality is negatelively correlated with the practice of bride price, the use of the plough as well as animal husbandry. However, she does not find any statistically significant correlation between ethnic group's historical matrilineality and historical jurisdictional hierarchy. Nevertheless, I will control for all these historical correlates, as described in subsection 3.3[6].

[5]Using data from the *Ethnographic Atlas*.

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

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 beige[7]. 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 these countries crossed by the matrilineal belt. In fact, ones talk about an *"AIDS Belt"*. I describe it in more detail in the next subsection.

Figure 2: Ancestral Ethnic Group Boundaries and Matrilineal Belt



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 worldwide female HIV positive population (UNAIDS (2018b)). Uniquely, it is the only place in the world

[7]This 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.

where more women than men live with HIV: adult female HIV positive population is about 1.5 larger than male HIV positive population[8], and adult females represented about 57% of new adult infections in 2017[9].This phenomenon has been referred to "Feminization of HIV" in Sub-Saharan Africa [10].

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 is represented in Figure 3[11], is computed from observations in my final sample, using DHS data and linking individuals to their ancestral ethnicity, and shows high prevalence rate in Easterna dn Southern Africa, the so-called *"AIDS Belt"*[12]. Incidentally, Zambia and Malawi, two the countries with highest female HIV rates are also two of the countries with the highest proportions of individuals originating from matrilineal ethnic groups[13].

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 barganing power and ability to impose safe sexual pratice 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.

2.3 Conceptual Framework

Structural elements of matrilineal societies may increase women's status, relative to patrilineal societies (Lowes (2018a)). Examples of structural elements of matrilineality include matrilocality (which is the practice

[8] 9.5 millions males live with HIV in Sub-Saharan Africa

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

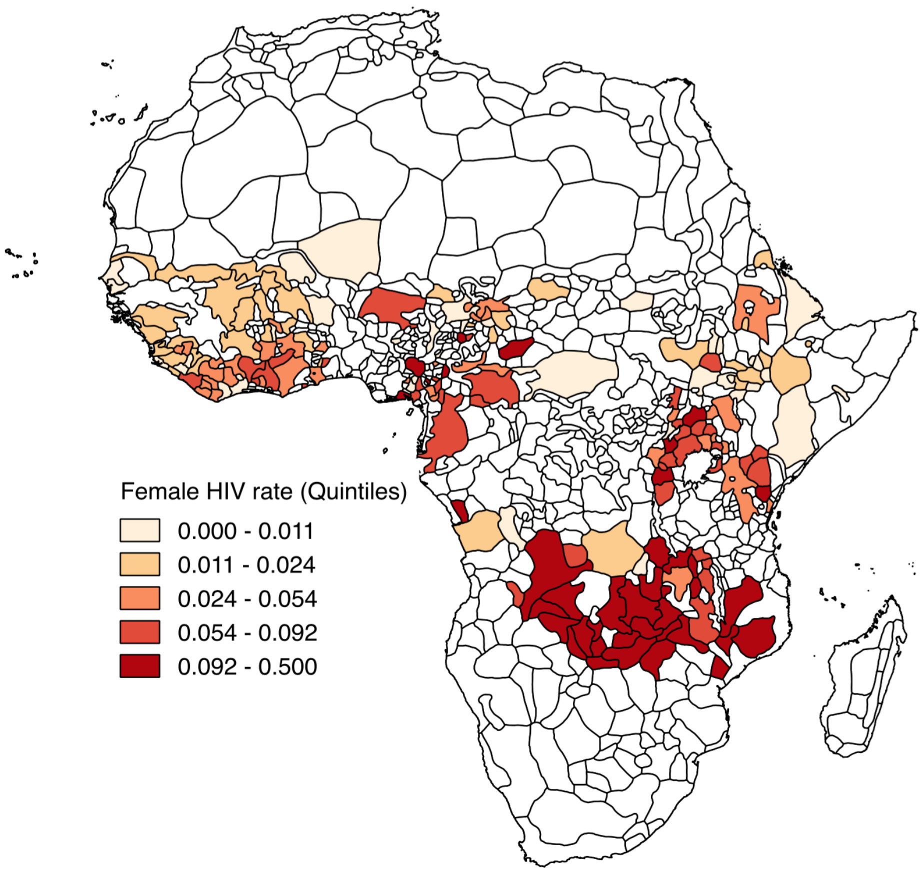
[10] These numbers are from http://aidsinfo.unaids.org/

[11] This 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.

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

[13] In 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.

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



of living close to the wife's relatives), and women's inheritance of land. In an intrahousehold bargaining framework, both these structural elements constitute better marriage outside options, and consequently enable women to benefit from higher bargaining power and ability to better implement their preferences. An alternative explanation of higher women's status is that matrilineal societies inherently value women more (Lowes (2018a)). This may be an internalized social norm that resulted from the structural factors discussed above, and thus persists even when these structural elements are no longer present.

This relative higher women's status within matrilineal societies, as compared to patrilineal societies, may foster female's sexual autonomy and potentially lead to riskier sexual behaviour, more conducive to HIV infection. This is the Emmanuel Todd's hypothesis (Todd (2017), p. 101)[14]:

"The spread of HIV virus infection provides a second measure, more indirect, of women's status. Women's sexual freedom is indeed negatively correlated with male's domination and the intensity of patrilineality. Where the level of control over women is weak, moral freedom has unfortunately permitted a vast diffusion of the virus. Where the level of control over women is strong, the spread of the disease has been better restrained."

An other potential driver of such female's sexual autonomy might be the fact that relative to patrilineal males, matrilineal males may care less about the certainty of their paternity. Indeed, matrilineality may be

[14] Source of the translation: the author

advantageous in environments with low paternal certainty. 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)).

Additionally, a recent litterature has highlighted the influence of matrilineality on gender related behaviour. In particular, Lowes (2018a) provides experimental evidences from DRC that, due to the social structures described above, men and women from matrilineal ethnic groups cooperate less with their spouses than their patrilineal counterparts. Further, Lowes (2018b) provides, in a similar context, experimental evidences that matrilineality may have gender-specific effect on social preferences and psychological traits. As such, while she finds that matrilineality does not significantly affect male individuals risk-preference, she finds that matrilineality closes the gender gap in risk-preference, with matrilineal women having a higher preference for risk than their patrilineal counterparts. These gender-specific differences in contemporaneous behaviours may also be expected to relate to sexual behaviour and within marriage contraception behaviour. Namely, it can be hypothesized that both preference for risk and lower spousal cooperation may increase, *relative to patrilineal females*, matrilineal female's propensity to behave in risky manner with respect to sexuality (e.g. more infidelity); and decrease matrilineal female's use of male condom within marriage.

Finally, in an intrahousehold bargaining framework, it can also be hypothesized that, benefiting from better marriage outside option, matrilineal females may overestimate the efficiency of their threatening strategy against male's infidelity, and consequently adversely underestimate their husband's propensity of behaving in an unfaithful manner. This would also lead in matrilineal females having a lower propensity to use condom within marriage, as compared to their patrilineal counterparts. This is an hypothesis I will explore in section 7.

All in all, in line with the geographical correlation highlighted above between the so-called *"Matrilineal Belt"* and the so-called *"AIDS Belt"* in Sub-Saharan Africa, all these channels point towards a higher prevalence of HIV among females originating from matrilineal ethnic groups, relative to women originating from patrilineal ethnic groups. This is exactly what I empirically test in this paper.

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-surveys[15]). 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 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 autononomy, actual sexual behaviour, and contraception use, which I investigate in section 6. Also, a set of questions pertain to individual acknowledgment about HIV risks and protective methods, as well as individual attitude towards life sex, which I investigate in section 6 and section 7.

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.

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 world[16], and that contains detailed ethnographic information on cultural aspects and ways of life of the portrayed ethnic groups, *prior to industrialization and colonial contact*[17], 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 datasets[18].

[15] 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)

[16] The majority of the ethnicities sampled are in Africa.

[17] "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))

[18] For 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

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) matching[19] to match some remaining unmatched individuals. Finally, I build on online sources[20] 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:

$$\begin{aligned} y_{ievrt} = \alpha + \beta_1 Matrilineality_e + \beta_2 Female_i + \beta_3 Female_i \times Matrilineality_e + \\ \mathbf{X'_{ievrt}}\Delta + \mathbf{X'_{ert}}\Omega + \mathbf{X'_{vrt}}\Pi + \lambda_{rt} + \varepsilon_{ievrt} \end{aligned} \tag{1}$$

with y_{ievrt} denoting an individual-level outcome (for example an 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. $Matrilineality_e$ 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 HIV. 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.

$\mathbf{X'_{ievrt}}$ represents a set of individual-level covariates which includes indicators of marital status; a dummy

because an alternative group's name is used.

[19] Teso (2019) matching procedure largely builds on Michalopoulos et al. (2019) matching procedure.

[20] One 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.

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 presence of clans; indicators of ancestral settlement patterns; indicators of ancestral juridictional 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; index of ethnic fractionalization; and index of ethnic polarization.

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 and geographic factors that could potentially affect contemporaneous HIV prevalence and also be correlated with the geographical distribution of ancestral matrilineality.

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_{ievrt}), throughout the analysis standard errors are clustered at the ethnic group $\times$ (within-country) DHS region $\times$ survey (year) level.

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 institutions [21], as well as modes of production more conducive of equal gender norm[22] 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

[21]e.g. Ashraf et al. (2019) on the role of bride price
[22]e.g. Alesina et al. (2013) on the role of plough use

De Walque (2012), mine workers' international migration is an other 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. 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, correlated with both ancestral matrilineality and contemporaneous HIV prevalance. To account for this, I control for village's population density (2010) and nightlight composite, two proxies of economic prosperity. Note also that 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 this.

Finally, institutional factors may also drive female HIV. As an illustration, Anderson (2018) claims that legal origins of Sub-Saharan African countries significantly determine current-day female HIV rates: she namely finds that female HIV rates are signicantly 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 1 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 $Female \times Matrilineality$ 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 possibles 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 $Female \times Matrilineality$ 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. This effect corresponds to 39% to 54% of the average HIV prevalence (4.1%) in my full regression sample.

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 "Female HIV in Sub-Saharan Africa".

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 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 signicantly 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 ($Female \times Matrilineality$) is in line with my conceptual framework according to which women's position in ancestral marilineal societies has a long-lasting effect specific on female's contemporaneous sexual behaviour. This result is confirmed by the analysis made by gender subsamples, and presented in the columns 6 and 7 of Table 1: I find that ancestral matrilineality significantly increases females' likelihood of being HIV positive (column 6) by 1.2 percentage points (representing 25% of the average HIV prevalence of females in my regression sample); while I do not find any significant effect for their male counterparts (column 7).

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

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

	HIV						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Matrilineality	0.058***	0.000	0.001	0.000	-0.001	0.012*	-0.004
	(0.006)	(0.005)	(0.004)	(0.005)	(0.006)	(0.007)	(0.006)
Female	0.016***	0.012***	0.009***	0.009***	0.008***		
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)		
Female × Matrilineality	0.020***	0.022***	0.019***	0.016***	0.017***		
	(0.004)	(0.004)	(0.004)	(0.004)	(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.	0.041	0.041	0.041	0.041	0.041	0.048	0.031

Notes: OLS estimates are reported with standard errors clustered at the ethnicity × (within-country) DHS 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 presence of clans; indicators of ancestral settlement patterns; indicators of ancestral juridictional 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); 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. * $p<0.10$, ** $p<0.05$, *** $p<0.01$

comparison between the characteristics of those who agreed to be tested and those who refused testing shows minimal bias[23]. Moreover, it is reasonable to expect that any selection will cause a downward bias since infected individuals should be less keen to be tested[24]. 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 of females on the probability of underreporting HIV infection, which in practice amounts to a refusal to consent

[23]See https://dhsprogram.com/topics/HIV-Corner/hiv-prev/index.cfm

[24]This is confirmed by Mishra et al. (2006), who find that the rate of HIV infection among individuals not tested for HIV is systematically larger than the rate among those not tested.

Table 2: **Selection And Falsification Tests**

	Selection		Falsification Tests		
	Consent HIV Test	Take HIV Test	Anemia	BMI	Rohrer Index
	(1)	(2)	(3)	(4)	(5)
Matrilineality	-0.029***	-0.031***	-0.008	-0.184*	-0.094
	(0.007)	(0.008)	(0.018)	(0.108)	(0.071)
Female	0.018***	0.020***			
	(0.002)	(0.002)			
Female × Matrilineality	-0.000	0.001			
	(0.004)	(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	0.196	0.166
Clusters	1,204	1,204	1,027	1,324	1,324
Mean Dep. Var.	0.947	0.941	0.460	22.231	14.041

Notes: OLS estimates are reported with standard errors clustered at the ethnic group × (within-country) DHS 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 actualy 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. * $p<0.10$, ** $p<0.05$, *** $p<0.01$

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 2 report estimates on the propability that DHS respondant consent to take HIV test (dummy outcome), and on the probability that she actualy takes it (dummy outcome). 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 the same 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 disease[25], as well as BMI and Rohrer index as other objective measures of health status[26], I perform falsification tests by estimating the long-term effect of ancestral matrilineality on such health outcomes[27]. 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 2, and I find that, unlike for HIV, ancestral matrilineality is not associated with an increase in the prevalence of anemia, neither with Rohrer index[28] of female individuals, and only very weakly associated[29] with their Body Mass Index (BMI)[30], corroborating the hypothesis that sexual behaviour is the actual driver of my results on HIV.

4.2.3 Robustness to Alternative Channels

The previous results demonstrate that ancestral matrineality is a long-term determinant of contemporaneous prevalence of female HIV in Sub-Saharan Africa. However, other long-term determinants of female HIV in Sub-Saharan Africa have also been recently highlighted in the literature. The aim of the heterogenous analysis performed in this subsection is to rule out the possibility that such alternative channels are in fact driving the long-term effect of ancestral matrilineality.

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 pratices 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 ethnical effect. In some sense, my identification strategy in symmetric since I exploit ethnic norm (and related kinship organization) variation *within region (within-country) - survey (year)*, therefore ruling 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 countries[31]. The two first columns of Table 9 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

[25] Bertocchi 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

[26] Only available for female respondants

[27] I 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.

[28] $mass/height^3$

[29] p-value=0.088; magnitude = 0.83% of the sample mean

[30] $mass/height^2$

[31] Following Anderson (2018), I use the dataset in La Porta et al. (2008) in order to identifiy common law and civil 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 in common law countries[32].

Polygyny. Bertocchi and Dimico (2018) have recently highlighted contemporaneous polygyny as an other driver fo 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 are meant to capture this alternative transmission channel.

Again, I also peform an heterogeneous analysis as an additional robustness check, by estimating Equation 5 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 9 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-poygynous individuals only. Similarly to the previous heterogenous 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-polygynyous females[33].

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 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), malaria incidence (2010), vegetation index, and indicators for the length of the growing season (see subsection 3.3). My main results being robust and only slightly changed 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.

[32]In 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.

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

Table 3: **Heterogeneous Effects by Subsamples: Common Law vs. Civil Law Countries / Polygynous vs. Non Polygynous Individuals**

	HIV			
Sample	Common Law Countries	Civil Law Countries	Ind. not in Polygynous Union	Ind. in Polygynous Union
	(1)	(2)	(3)	(4)
Matrilineality	-0.003	-0.002	-0.001	0.002
	(0.009)	(0.007)	(0.005)	(0.019)
Female	0.009***	0.008***	0.011***	-0.001
	(0.002)	(0.001)	(0.001)	(0.003)
Female × Matrilineality	0.015***	0.002	0.014***	0.007
	(0.006)	(0.005)	(0.004)	(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.	0.069	0.018	0.042	0.035
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 ethnic group × (within-country) DHS 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 Countrie* 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*

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 writes, 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 to social sanction individuals should be detected as having risky sexual behaviuor, 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, provides 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 are computed at the village level, following Montalvo and Reynal-Querol (2005) formulas:

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

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 Herfindhahl 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.

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

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, ... 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. I also control for such potential differences at the village-level by including, as proxies of village-level economic development, nighlight composite as well as population density. I finally control for such diferences at the ethnic group level by controlling for ancestral settlement patterns; ancestral juridictional 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 only slightly changed estimates fo the long-term effect of ancestral matrilineality on female HIV once these controls are included (see Table 1, column 5), 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 controling 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 Foreig Bible Society (BFBS) missions per square kilometer of an ethnic group's land during the colonial period. Additionaly, 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 theland 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 respondants in DHS[34], 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 prositution[35], which have both been emphasized in the literature as an impediment and a driver of HIV, respectively.

Table 4 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 1 percentage point more likely to suffer from HIV than their male counterpart, an effect representing about 22% of the mean HIV rate in my regression sample. Further, I still find that female originating from an ancesrally matrilineal ethnic group are 1.6 percentage point more likely to sugger from HIV than their patrilineal counterparts, an effect representing about 35% of the mean HIV rate in my regression sample.

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

[35]Unfortunately for this study, in DHS female respondants have not been interrogated about prostitution.

Table 4: **Robustness of OLS Estimates to Additional Controls**

	HIV								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Matrilineality	0.001	0.000	0.001	0.001	0.001	0.001	-0.100*	-0.010*	-0.012
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.08)
Female	0.008***	0.008***	0.008***	0.009***	0.009***	0.009***	0.009***	0.008***	0.010***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)
Female × Matrilineality	0.016***	0.016***	0.016***	0.016***	0.016***	0.016***	0.017***	0.017***	0.016***
	(0.04)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.004)	(0.005)
Ind. Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ethnic Group Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Village-Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region-survey FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Murdock Ethnic Groups Historical Controls:									
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
DHS Villages Contemporaneous Controls:									
Village's prop. of males circumcised							Yes		Yes
Village's prop. of males who paid for sex								Yes	Yes
Observations	166,419	166,419	166,419	166,983	166,983	166,983	115,620	136,586	106,362
Adj. R-squared	0.080	0.080	0.080	0.080	0.080	0.080	0.090	0.086	0.091
Clusters	1,489	1,489	1,489	1,499	1,499	1,499	1,088	1,190	964
Mean Dep. Var.	0.043	0.043	0.043	0.043	0.043	0.043	0.043	0.038	0.046

Notes: OLS estimates are reported with standard errors clustered at the ethnicity × (within-country) DHS 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/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 respondants circumcised. "Village's prop. of males who paid for sex" is the within DHS village's proportion of male respondants who report having paid for sex in the last 12 months. * $p<0.10$, ** $p<0.05$, *** $p<0.01$

5 Identifying Causal Relationships

The positive association between ancestral matrilineality and female HIV that is documented in the previous section is consistent with the hypothesis that ancestral matrilineality engendered a higher female sexual freedom and, adversely, riskier sexual behaviours more conducive to HIV. However, despite the great care provided in the last section on controlling for numerous alternative channels, remaining omitted variables that are correlated with both ancestral matrilineality and contemporaneous HIV as well as sexual and contraception behaviours might still explain this correlation. For example, if ethnic groups where females were inherently more likely or more free to adopt riskier sexual behaviours were also more likely to adopt matrilineal kinship organizations, and if females in these groups continue to adopt riskier sexual behaviours today, then this could generate a positive relationship between ancestral matrilineality and contemporaneous female HIV.

In this section, I pursue three strategies to assess whether the correlations documented to this point are causal. First, I use selection on observable variables to assess the likelihood that my estimates are being driven by unobserved heterogeneity across ethnic groups. Second, I use the distance from the individual's DHS village to the closest ancestral matrilineal ethnic boundary as an instrument for individual's ethnic group's ancestral matrilineality. Finally, I adopt a geographic regression discontinuity (RD) strategy, comparing locations that are sufficiently close and located at each side of an ancestral matrilineal ethnic boundary.

5.1 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 1 may still be biased by unobservable factors correlated with both ancestral matrilineality and females contemporaneous sexual and contraception behaviours, and subsequently HIV. In this subsection, I build on Moscona et al. (2018) to assess the likelihood that the OLS estimates are 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 5, where each row reports result for different set of restricted covariates. This yields four ratios that range from -1801.22 to 212.68. 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 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: $\beta^* = \hat{\beta_F} - (\hat{\beta_R} - \hat{\beta_F}) \times ((R^2_{Max} - \hat{R^2_F})/(\hat{R^2_F} - \hat{R^2_R}))$, where $\hat{\beta_R}$ and $\hat{\beta_F}$ are defined as above; $\hat{R^2_F}$ is the R-squared from the fully-controlled regression; $\hat{R^2_R}$ is the R-squared from the restricted regression; and R^2_{Max} is the R-squared from a regression that includes all observable and unobservable controls. Although in theory the maximum possible value of R^2_{Max} is one, as underlined by González and Miguel (2015), in the real world, where there is significant measurement error, the value of R^2_{Max} should be much lower than one. In fact, by definition, $R^2_{Max} \in [\hat{R^2_F}; 1]$. Oster (2017) provides some insights on how R^2_{Max} should be chosen, showing that $R^2_{Max} = 1$ may lead to over-adjustment in many cases. I follow her procedure and present bias-adjusted lower bound coefficients for $R^2_{Max} = min(\mathbf{1.3}\hat{R^2_F}; 1)$[36], $R^2_{Max} = min(\mathbf{1.5}\hat{R^2_F}; 1)$, and $R^2_{Max} = min(\mathbf{2}\hat{R^2_F}; 1)$ in column 2, 3 and 4 respectively (I also report bias-adjusted lower bound coefficients for $R^2_{Max} = 1$ in column 5 for informational purpose).

All lower biased-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 suggest 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.

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

Table 5: **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 S.E.**
Controls in Restricted (R) set	Controls in Full (F) set		$R^2_{Max} = min(\mathbf{1.3}\hat{R^2_F}; 1)$	$R^2_{Max} = min(\mathbf{1.5}\hat{R^2_F}; 1)$	$R^2_{Max} = min(\mathbf{2}\hat{R^2_F}; 1)$	$R^2_{Max} = \mathbf{1}$	
Matrilineality; Female	Full set of controls from Equation 1	15.47	0.017	0.017	0.018	0.031	[0.004 ; 0.029]
Matrilineality; Female; Region-survey FE	Full set of controls from Equation 1	-4.93	0.014	0.012	0.008	-0.087	[0.004 ; 0.029]
Matrilineality; Female; Region-survey FE; Ind. Controls	Full set of controls from Equation 1	-1801.22	0.017	0.016	0.016	0.012	[0.004 ; 0.029]
Matrilineality; Female; Region-survey FE; Ind. Controls; Ethnic Group Controls	Full set of controls from Equation 1	212.68	0.018	0.019	0.022	0.081	[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 $\hat{R^2_R}$ is the R^2 of the regression with the restricted set of controls, and $\hat{R^2_F}$ is the R^2 of the regression with the full set of controls, then the minimum coefficient lower bound is: $\hat{\beta_F} - (\hat{\beta_R} - \hat{\beta_F}) \times ((R^2_{Max} - \hat{R^2_F})/(\hat{R^2_F} - \hat{R^2_R}))$. Column 6 reports the bounds of the 99.5% confidence interval of the fully controlled estimate of $Matrilineality \times Female$.

Nearest Neighbor Matching. As an alternative 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 area (i.e. patrilineal or other)[37] that is the most similar in terms of geographic characteristics. The matching average treatment over treated effects are reported in Table 1, with village pairs being matched[38] 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 2.7 percentage points to their nearest neighbor village located in ancestrally non-matrilineal area, an effect significant at 5% and representing about half of the average village's female HIV rate in my regression sample of 13,176 DHS villages. Interestingly, no such effect is found on village's proportion of HIV positives males[39], consistent with my findings so far that the effect of ancestral matrilineality on contemporaneous HIV rates is specific to female populations.

Table 6: **Nearest Neighbor Matching (ATT)**

	(1) Village's proportion of HIV positive females	(2) Village's proportion of HIV positive males
Ancestrally matrilineal area	0.027** (0.011)	0.010 (0.011)
Observations	13,176	12,710
Mean Dep. Var.	0.055	0.037

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*

5.2 Instrumenting for Ancestral Matrilineality

While the results presented so far are robust to controlling for a wide array of observable individual, ethnic group's historical and village-level geographic factors, and unlikely driven by unobservables, there could still be unobserved omitted variables that are correlated with both an ethnic group's ancestral matrilineal kinship organization and current female 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 matrilinelity 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 environment[40]. In particular, matrilineality may be advantageous in environments with low paternal certainty since while it is difficult to confirm paternity,

[37] Based on Murdock's map of ancestral ethnic groups (see Figure 2).

[38] I 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).

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

[40] Lowes (2018a) appendix provides an extensive overview of the hypothesized origins of matrilineal kinship systems.

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 behaviours[41], and therefore drive the OLS estimates towards zero.

To address these concerns, I rely on an instrumental variable strategy which I present in this subsection. More specifically, I instrument 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 distance. Note that I 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 boundary.

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 contraception 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 subsection 4.2.

Table 7 presents estimates of my instrumental variable strategy. The OLS estimates of regressing my

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

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 very 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.

5.3 Accounting for Unobservables: Geographic RD Estimates

To further address identification concerns discussed in the previous subsections, and omitted variable bias concern in particular, I undertake a geographic regression discontinuity analysis as an additional alternative strategy. 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 practice matrilineal descent alongside groups that practice 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 similar (note that the ancestral matrilineal ethnic boundary does not coincide with any actual border).

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 matrilinral ethnic boundary. The benefit of this strategy is that it accounts for 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 orde to get the more conservative estimates, I directly control for the large array of individual,

[42] I 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 7: **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 × Matri.**	
Matrilineality	-0.010	-0.010				
	(0.038)	(0.246)				
Female	0.006***	0.008***	Yes	Yes	Yes	Yes
	(0.001)	(0.002)				
Female × Matrilineality	0.029***	0.020***				
	(0.007)	(0.007)				
	[0.018;0.04]	[0.011;0.032]				
Distance to matri. boundary			0.0003***	0.0002***		
			(0.000)	(0.000)		
Female × Distance to matri. boundary					0.0008***	0.0011***
					(0.000)	(0.000)
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	163,452	179,312	163,452	179,312	163,452
Adj. R-squared	0.032	0.034	0.835	0.883	0.617	0.673
Clusters	1,671	1,472	1,671	1,472	1,671	1,472
Mean Dep. Var.	0.040	0.042	0.158	0.171	0.085	0.092
F-Stat (test coeff. = 0)			18.86	15.85	63.54	79.81

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 × (within-country) DHS 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 × 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 4. ** $p<0.10$, ** $p<0.05$, *** $p<0.01$*

ethnic group's historical and village-level geographic control, as well as (within-country) region × survey (year) fixed effects described in the previous sections. More specifically, my RD specification takes the following form:

$$y_{ievrt} = \alpha + \beta_1 Matrilineality_v + \beta_2 Female_i + \beta_3 Female_i \times Matrilineality_v + \\ \mathbf{X'_{ievrt}}\Delta + \mathbf{X'_{ert}}\Omega + \mathbf{X'_{vrt}}\Pi + f(location_v) + \lambda_{rt} + \varepsilon_{ievrt} \quad (4)$$

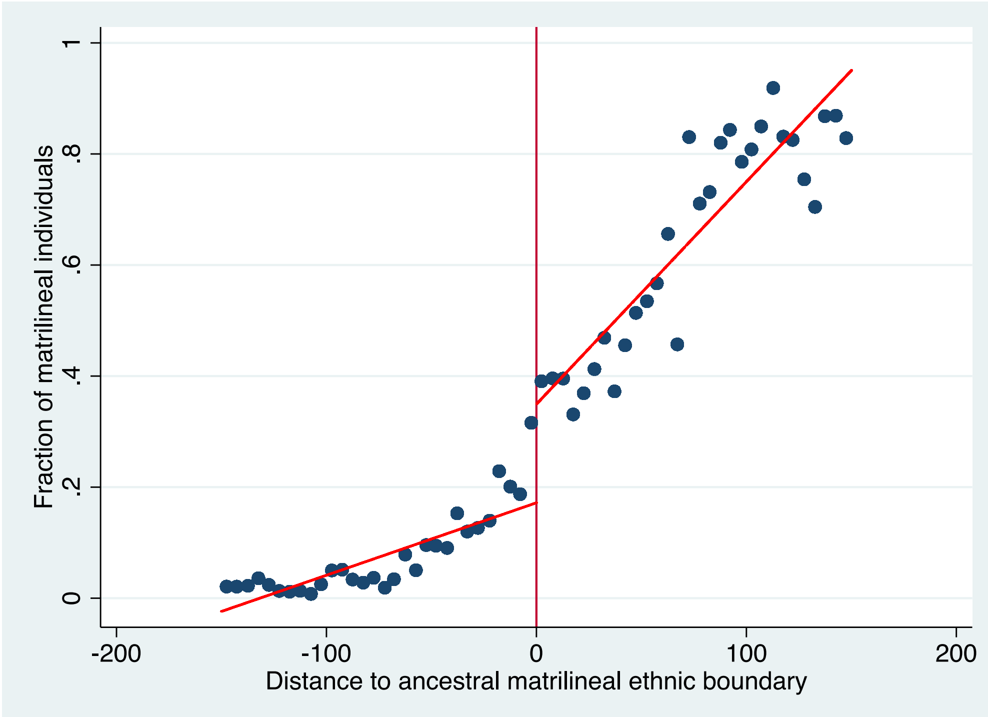
with y_{ievrt} denoting an individual-level outcome (indicator for whether the individual is HIV positive) for individual i from ethnic group e, living in village v in within-country DHS region r, and surveyed at year t. $Matrilineality_v$ is an indicator for whether an individual lives in a village 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. $f(location_v)$ denotes a 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"). $\mathbf{X'_{ievrt}}$; $\mathbf{X'_{ert}}$, 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_{ievrt} are clustered at the ethnic group $\times$ (within-country) DHS region $\times$ 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. (2018) 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 accurate[43]. 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 4, 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 borders[44].

[43] This 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.

[44] It 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.

Figure 4: 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.

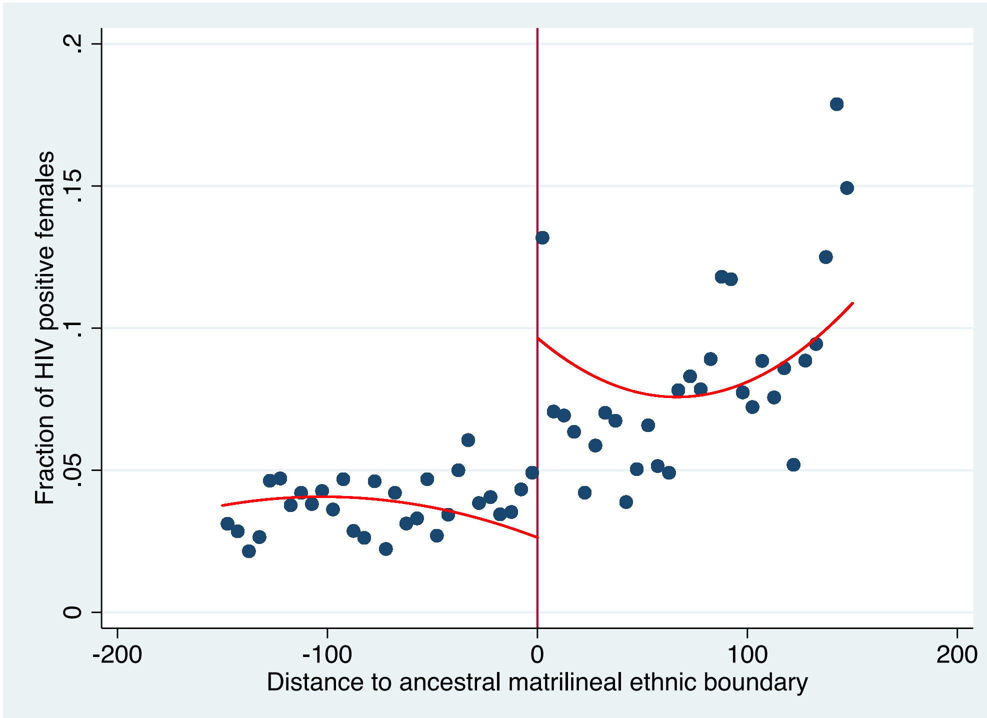
Geographic RD Estimates. Before turning to my estimates I first examine the raw data for the RD sample. Figure 5 shows a bin scatterplots of the predicted HV rate for females living in villages located within 150 km of the ancestral matrilineal ethnic border, using a flexible second-order RD polynomial conditionned 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 area 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 8 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 village located within the ancestral boundaries of a 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 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 matrilinal ethnic groups. As shown in Figure 4, 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, has it as been underlined by Tequame (2012), riskier sexual behaviours are more easy to conceal in fragmented societies. Alternatively, migration at the boundary could further 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 muted[45].

[45]The estimated effect of $Female \times Matrilineality$ 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.

Figure 5: 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) × survey (time) FE, and for which a quadratic 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.

6 Mechanisms

The previous section have uncovered a robust relationship between ancestral matrilineality and the contemporaneous spread of the HIV epidemic among female individuals in Sub-Saharan Africa. These findings suggest that the legacy of ancestral matrilineality may be conducive to social, sexual and contraception behaviours that favor a higher rate of transmission of HIV among women. More specifically, I conjecture that, benefiting from higher sexual autonomy, women originating from ancestrally matrilineal ethnic groups adopt riskier sexual and contraception behaviours. In this section, I turn to a more direct investigation of these channels of transmission, by exploring the empirical relationship between ancestral matrilineality and female's sexual autonomy, sexual behaviour and contraception use.

6.1 Female Sexual Autonomy

Preliminary, I explore whether women originating from ancestrally matrilineal ethnic groups benefit from a higher social status, and consequent higher bargaining power and sexual autonomy. First, building on the

Table 8: **Geographic RD Estimates**

	(1)	(2)	(3)	(4)	(5)	(6)
	Dep. Var.: HIV					
Distance from matrilineal boundary:	100km		150km		200km	
Running variable:	*Distance*	*Lat./Long.*	*Distance*	*Lat./Long.*	*Distance*	*Lat./Long.*
	Panel A: Linear Polynomial					
Matrilineality	-0.000	0.000	-0.001	-0.001	-0.002	-0.001
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Female	0.011***	0.011***	0.010***	0.010***	0.011***	0.011***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Female × Matrilineality	0.008*	0.008*	0.009**	0.009**	0.009**	0.009**
	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)
	Panel B: Flexible Linear Polynomial					
Matrilineality	0.005	0.002	0.000	0.001	-0.001	0.000
	(0.005)	(0.006)	(0.004)	(0.006)	(0.004)	(0.006)
Female	0.011***	0.011***	0.010***	0.010***	0.011***	0.011***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Female × Matrilineality	0.008*	0.008*	0.009**	0.009**	0.009**	0.009**
	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)
	Panel C: Quadratic Polynomial					
Matrilineality	-0.000	0.000	-0.001	-0.000	-0.001	-0.000
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Female	0.011***	0.011***	0.010***	0.010***	0.011***	0.011***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Female × Matrilineality	0.008*	0.008*	0.009**	0.009**	0.009**	0.009**
	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.004)
	Panel D: Flexible Quadratic Polynomial					
Matrilineality	0.005	0.000	0.005	0.001	0.001	0.001
	(0.008)	(0.008)	(0.007)	(0.008)	(0.006)	(0.008)
Female	0.011***	0.011***	0.010***	0.010***	0.011***	0.011***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Female × Matrilineality	0.008*	0.008*	0.009**	0.009**	0.009**	0.009**
	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.004)
	Panel E: Cubic Polynomial					
Matrilineality	-0.000	0.000	-0.001	-0.000	-0.001	-0.001
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Female	0.011***	0.011***	0.010***	0.010***	0.011***	0.011***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Female × Matrilineality	0.008*	0.008*	0.009**	0.009**	0.009**	0.009**
	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)
	Panel F: Flexible Cubic Polynomial					
Matrilineality	0.007	-0.007	0.000	-0.009	0.005	-0.009
	(0.010)	(0.010)	(0.008)	(0.009)	(0.008)	(0.009)
Female	0.011***	0.011***	0.010***	0.010***	0.011***	0.011***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Female × Matrilineality	0.008*	0.008*	0.009**	0.009**	0.009**	0.009**
	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(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.	0.044	0.044	0.044	0.044	0.045	0.045

Notes: Geographic RD estimates based on Equation 4 are reported with standard errors clustered at the ethnicity × (within-country) DHS 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. * $p<0.10$, ** $p<0.05$, *** $p<0.01$

household economics literature highighting the role of marriage outside option on intrahousehold bargaining power (Baland and Ziparo (2017); Bargain et al. (2018)), 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. 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.

A second set of outcomes pertains 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 womand could get a male condom herself, whether a wife is justified to ask husband to use condom if he has a STI (Sexually Trasmitted 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:

$$y_{ievrt} = \alpha + \beta_1 Matrilineality_e + \mathbf{X'_{ievrt}}\Delta + \mathbf{X'_{ert}}\Omega + \mathbf{X'_{vrt}}\Pi + \lambda_{rt} + \varepsilon_{ievrt} \tag{5}$$

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 $\times$ (within-country) DHS region $\times$ survey (year) level. My coefficient of interest is now β_1 and captures the long-term effect of ancestral matrilineality[46].

In Table 9, I report results from estimating Equation 5 on these outcomes. I find that ancestral matrilineality is indeed significantly associated with all these dimensions of female's social status, bargaining power and sexual autonomy. More specificaly, according to the estimates reported in column 1, matrilineal women are 1.5 percentage points significantly more likely to be divorced, a large effect as compared to the mean proportion of divorced female individuals in my regression sample. Also, according to the estimate reported in column 2, I find that divorced women originating from ancestrally matrilineal ethnic group are 25.1 percentage points (representing about 52% of the mean probability of the sample) more likely to own a house and/or a land, as compared to their divorced patrilineal counterparts. Further, estimates reported in column 3 to 6 indicates, in the same order, that matrilineal women are 4.2 percentage points more likely to be decision-maker regarding the use of contraception; 2.4 percentage points more likely to answer that they could get (male) condom themselves; 2.9 percentage points more likely to find justified for a woman to ask

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

her husband to use a condom if he has a STI; and 2.6 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 order to explain the highest HIV prevalence rate for females originating from ancestrally matrilineal ethnic groups, my results rule out intrahousehold bargaining mechanism highlighted by Anderson (2018), 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 next two subsections. More precisely, I explore whether the higher social status and sexual autonomy of women originating from ancestrally matrilineal ethnic groups translate into riskier sexual and contraception use behaviours.

Table 9: **Ancestral Matrilineality and Female Sexual Autonomy (OLS)**

	Divorced	Own house and/or land (divorced)	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)
Matrilineality	0.015***	0.251**	0.042**	0.024*	0.029***	0.026**
	(0.003)	(0.122)	(0.018)	(0.013)	(0.011)	(0.013)
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	105,964	1,191	13,709	56,270	91,420	95,604
Adj. R-squared	0.041	0.316	0.054	0.122	0.157	0.139
Clusters	1,550	238	981	1,392	1,392	1,392
Mean Dep. Var.	0.021	0.479	0.877	0.634	0.780	0.626

Notes: OLS estimates are reported with standard errors clustered at the ethnic group × (within-country) DHS 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 respondant 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. * $p<0.10$, ** $p<0.05$, *** $p<0.01$

6.2 Sexual Behaviour

In order to capture sexual behaviour that is a prominent risk factor for 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 10, which presents results from estimating Equation 1 on these sexual behaviour outcomes. To begin with, according to 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. However, 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 women are 3.2 percentage points more likely to have had any sexual activity in the last 4 weeks (column 1); 1.9 percentage points less likely to have use a condom during their last sexual intercourse (column 2), an effect representing about 15% of the mean probability of the full sample of males and females; and 2.6 percentage points (representing about 41% of the mean probability) more likely to report any extramarital affairs in the last 12 months (column 3). Estimate from column 4 also indicates that originating from an ancestrally matrilineal ethnic group increases the number of female's extramarital affairs in the last 12 months by 0.068, representing about 72% of the sample mean. Finally, and interestingly, it is worth noting that according to estimates reported in the first row, ancestral matrilineality effect is specific to females, since matrilineal males do not have a signficantly different propensity to adopt risky sexual behaviour than their patrilineal counterparts. If any, matrilineal males have a marginally significant lower number of extramarital affairs. Importantly, this latter result provides an explanation to the fact that the (ancestral) matrilineality long term effect on contemporaneous HIV found in this paper is specific to females.

The analysis on couples' HIV discordance I perform in Table 11 provides support to these results. As a matter of fact, 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 1.3 percentage points more likely to be HIV positive while having a HIV negative husband, relative to wives originating from ancestrally patrilineal ethnic groups, a large effect representing about 74% to 82% of my regressions samples' proportions of couples with a HIV positive wife and a HIV negative husband (see columns 1 and 2). Importantly, this result is robust 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"*, see column 2). 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 other channels than their husband. This is an important result, suggesting alternative channels of transmission of the virus to the one highlighted by Anderson (2018) (briefly that less empowered wives suffer more from HIV transmitted by their husbands, see section 1), that are also presumably at play when explaining the high rates of female HIV in Sub-Saharan Africa, and therefore call for complementary policies. Finally, I also find again that the effect of ancestral matrilineality on contemporaneous HIV is specific to female populations, since I do not find any significant effect of ancestral matrilineality on the probability that the husband is HIV positive while his wife is HIV negative (see columns 3 and 4).

While DHS data does not allow me to directly empirically test for the effect of matrilineality on gender-related social behaviour, it is worth noting that a recent litterature has highlighted the influence of matrilineality on such gender related behaviour. In particular, Lowes (2018a) provides experimental evidences from DRC that men and women 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 risky sexual behaviour found here, and within marriage contraception behaviour explored in the next subsection.

Table 10: **Ancestral Matrilineality and Sexual Behaviour (OLS)**

	Sexual activity	Condom during last sex	Infidelity (formally married)	Nb. of extramarital affairs (form. mar.)
	(1)	(2)	(3)	(4)
Matrilineality	-0.005	-0.008	-0.006	-0.037*
	(0.009)	(0.008)	(0.008)	(0.021)
Female	-0.079***	-0.064***	-0.127***	-0.196***
	(0.005)	(0.005)	(0.007)	(0.014)
Female × Matrilineality	0.032***	-0.019**	0.026***	0.068***
	(0.010)	(0.008)	(0.010)	(0.016)
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	181,108	134,888	104,489	104,489
Adj. R-squared	0.310	0.250	0.109	0.022
Clusters	1,688	1,630	1,558	1,558
Mean Dep. Var.	0.528	0.126	0.064	0.094

Notes: OLS estimates are reported with standard errors clustered at the ethnic group × (within-country) DHS region × survey (year) level in brackets. The unit of observation is an (formally married in columns 3 and 4) 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. “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 affair in the last 12 months. “Nb. of extramarital affairs” is the number of extramarital affairs in the last 12 months reported by an individual. “Any STI last 12 months” is a dummy indicating whether a respondant reports having had any STD (Sexually Transmitted Disease) in the last 12 months. 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 11: **Ancestral Matrilineality and Couples HIV Discordance (OLS)**

	Dep. Var.: HIV Discordant Couples			
	Wife + / Husband -	Wife + / Husband -	Wife - / Husband +	Wife - / Husband +
	(1)	(2)	(3)	(4)
Matrilineality	0.0133*	0.0131*	-0.0017	0.0024
	(0.0075)	(0.0079)	(0.0073)	(0.0072)
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
Husband's Controls		Yes		Yes
Observations	25,272	23,917	25,272	23,917
Adj. R-squared	0.021	0.024	0.027	0.030
Clusters	1,100	1,043	1,100	1,043
Mean Dep. Var.	0.018	0.016	0.021	0.020
Prop. Mixed Ethnicity	0.193	0.156	0.193	0.156
Prop. Mixed Matrilineality	0.048	0.044	0.048	0.044

Notes: OLS estimates are reported with standard errors clustered at the ethnic group × (within-country) DHS 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 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. *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. * $p<0.10$, ** $p<0.05$, *** $p<0.01$

6.3 Contraception Use

Infidelity may not be the only factor driving contemporaneous prevalence of HIV. According to the well-documented medical and economic literature, contraception methods constitute other factors influencing HIV contagion, and the result highlighted in the last subsection on matrilineal female's use of condom during their last sexual intercourse suggest that matrilineal females may also differ in their contraception behaviour. This is what I investigate in this subsection.

To do so, I exploit information in DHS on current contraception method used by respondant to create (1) a dummy for whether the respondant report using condom as her current contraceptive method, (2) a dummy for whether the respondant report using one of the methods classified by Anderson (2018) as male consent contraceptive methods (namely "male condom"; "periodic abstinence"; "withdrawal" ; or "prolonged abstinence"); (3) a dummy for whether the respondant report using a female consent contraceptive method (namely "pill", "IUD", "injection", "female sterilization", "implants/norplant", or "lactational amenorrhea (LAM)"); (4) and a dummy for whether the respondant use *any* contraceptive method. Importantly, according to Anderson (2018), male consent contraceptive methods are also *protective* methods against HIV, as compared to female consent contraceptive methods.

Table 12 presents results from estimating Equation 1 on these contraception use outcomes and reveal that, while not having a significantly different contraception behaviour at the extensive margin, the estimate of $Female \times Matrilineality$ being non significant on the probability of using any contraception method (column 4), matrilineal women significantly differ from their patrilineal counterparts in the contraceptive methods they use. More specifically, I find that matrilineal women are 6.2 percentage points less likely to report using condom as a current contraceptive method, as compared to patrilineal woman, an effect of large magnitude representing about 100% of the mean probability of the full sample of males and females (column 1). Note that this mean probability is even lower for female subsample, with only 2.6% of females in my final sample reporting condom as a current contraceptive method (not reported). This very low level of condom use within marriage in Sub-Saharan Africa is documented in the literature[47]. More generally, I also find that matrilineal women are 6.6 percentage points less likely to report using any male consent protective contraception methods, representing about 79% of the mean probability of the full sample (column 2). These results are in line with the effect of ancestral matrilineality on condom use in last sexual intercourse outcome, found in the previous subsection. Nevertheless, I also find that matrilineal women are 5.2 percentage points more likely than their patrilineal counterparts to report using a female consent contraceptive method, representing about 53% of the full sample mean.

All in all, it seems that, having a higher status and being more likely to be decision-maker regarding contraception, matrilineal women are more likely to bear the responsability of contraception, relying less on their husband, and therefore substitute male consent (which are incidentally more protective) contraception methods by female responsability contraception methods. This is fully in line with Islam et al. (2009), who find similar patterns in the context of Bangladesh, underlying that matrilineal *Garo* 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.

However, results presented in subsection 6.1 have revealed that matrilineal women benefit from a higher social status, sexual autonomy and ability to impose safe sexual practices to their husbands. Therefore, one may still reasonably wonder why matrilineal women adopt riskier sexual and contraception behaviours than their patrilineal counterparts. This is the object of the discussion of the next section.

7 Discussions on Condom Use

Condom use within marriage may be in conflict with the most salient social norms that regulate marital relations; reproduction and fidelity being the fundamental elements that guide spouses' behaviour (Cordero-Coma and Breen (2012)). In this section I explore how matrilineal women differ from their patrilineal

[47] Chimbiri (2007) calls condom an "intruder in marriage".

Table 12: **Ancestral Matrilineality and Contraception Use (OLS)**

	Condom	Male consent **protective** contraception method	Female consent contraception method	Any contraception method
	(1)	(2)	(3)	(4)
Matrilineality	0.029***	0.030***	-0.018**	0.005
	(0.010)	(0.011)	(0.008)	(0.011)
Female	-0.054***	-0.040***	0.051***	-0.058***
	(0.005)	(0.006)	(0.004)	(0.009)
Female × Matrilineality	-0.062***	-0.066***	0.052***	0.017
	(0.015)	(0.019)	(0.010)	(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	182,311	181,321	181,321	147,175
Adj. R-squared	0.120	0.116	0.164	0.156
Clusters	1,688	1,685	1,685	1,507
Mean Dep. Var.	0.062	0.084	0.099	0.233

Notes: OLS estimates are reported with standard errors clustered at the ethnic group × (within-country) DHS 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" is a dummy indicating whether an individual reports "condom (male)" as her current contraception method. "Male consent protective contraception method" is a dummy indicating whether an individual reports currently using a contraception method that requires male's consent (namely "male condom"; "periodic abstinence"; "withdrawal" ; or "prolonged abstinence")), and which is incidentally a protective method against HIV. "Any contraception 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*

counterparts along such dimensions, and how this affects their condom use behaviour.

7.1 Desired Fertility

First, condom serves a contraceptive function. Nevetheless, the use of contraceptives may conflict with the social norm related to the production of children within marriage (Chimbiri (2007); Cordero-Coma and Breen (2012)). But this particular conflict should be less acute the more children a couple already has. It can therefore been hypothesized that condom use would therefore be more common when both husband and wife agreed on stopping or spacing births.

To test this hypothesis, I exploit information in DHS on fertility preference. More specifically, I create a dummy indicating whether an individual reports not wanting child anymore (*"No more child"*) and I regress Equation 5 on this outcome[48]. Result is reported in column 1 of Table 13 and reveals that matrilineal women are 2.6 percentage points less likely to have reached their desired level of fertility. Further, the heterogeneous analysis performed in column 3 highlights that most of the negative effect of ancestral matrilineality on condom use found previously (and reported again in column 2) is in fact driven by women who do not have

[48] Importantly, I do not control for the number of children in this regression in column 1, neither in regression in column 3

reached their desired level of fertility yet, consistent with my hypothesis [49]. All in all, these results suggests that, despite having a higher sexual autonomy and ability to impose condom, matrilineal women are less likely to use this latter (as compared to patrilineal women) notably because they are less likely to have reached their desired level of fertility.

Table 13: **Ancestral Matrilineality, Desired Fertility and Condom Use (OLS)**

	No more child	Condom	Condom
	(1)	(2)	(3)
Matrilineality	-0.026***	-0.009**	
	(0.008)	(0.005)	
Matrilineality × No more child			-0.003
			(0.006)
Matrilineality × More child			-0.008
			(0.005)
Ind. Controls	Yes	Yes	Yes
Ethnic Group Controls	Yes	Yes	Yes
Village-Geographic Controls	Yes	Yes	Yes
Region-survey FE	Yes	Yes	Yes
Observations	79,845	105,964	79,845
Adj. R-squared	0.376	0.074	0.077
Clusters	1,401	1,550	1,401
Mean Dep. Var.	0.255	0.026	0.34
F-Test Equality of coeff. (p-value)			0.074

Notes: OLS estimates are reported with standard errors clustered at the ethnic group × (within-country) DHS 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. "No more child" is a dummy indicating whether an individual reports not wanting another child. "Condom" is a dummy indicating whether an individual reports "condom (male)" as her current contraception method. Controls are defined in Table 1. Number of children is not included in the controls in columns 1 and 3. *Region-survey* is a subnational region defined in DHS, interacted with its survey-year. * $p<0.10$, ** $p<0.05$, *** $p<0.01$

7.2 Acknowledgment of HIV Risks and Access to Condom

Condom also serves a protective function, in particular against HIV risks. However, my results indicate that despite being more able to decide on contraception, matrilineal women are less likely to use it. Could a lack of ackowledgment of HIV risks and/or access to condom explain such a behavioural difference in condom use ? This is what I explore in this subsection.

To do so, I create several outcomes from information in DHS on ackowledgment of HIV risks and role of condom in reducing these risks. I create a dummy indicating whether an individual has ever heard about

[49] Though *Matrilineal* × *More child* is not significant a conventional level (p-value=0.129), the estimates of *Matrilineal* × *No more child* and *Matrilineal* × *More child* are significantly different (p-value=0.074).

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 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 14. 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 ackowledgments of HIV/STI risks (in fact, the vast majority of both matrilineal and patrilineal females ackowledges such risks). Consequently, I can rule out differential ackowledgment of HIV risks and role of condom against HIV as a factor explaining differences in matrilineal's females condom use. Second, matrilineal females' lower use of condom cannot be explained by a lack of access to codom. If any, I conversely find that matrilineal women have a significantly easier access to condom, in line with their higher sexual autonomy.

Therefore, since, in addition of having a higher sexual autonomy and ability to impose safe sexual practices to their husbands, matrilineal women also highly acknowledge[50] HIV and condom as a protective method against HIV risks, why does ancestral matrilineality negatively affect female's current condom use ? I propose several behavioural explanations in the next subsection.

7.3 Misperception of HIV Risks

It can be expected that an important factor determining whether a couple uses condoms in their marital sexual relations is whether the husband or wife believes that his or her spouse has been unfaithful and that there is thus a risk of contracting HIV. Despite the prevalence of extramarital sexual activity, marriage in Sub-Saharan Africa is supposed to be based on trust, faithfulness, and legitimate sex (and reproduction as previously seen), among other social norms, and this militates against condom use within marriage as a preventive strategy (Chimbiri (2007); Cordero-Coma and Breen (2012); Smith and Watkins (2005); Tavory and Swidler (2009); Watkins (2004)). I explore such social norms channels in this subsection.

On the perception of condom: *Ex-Ante Preventive* Method against the Risk of *Contracting* HIV Virus vs. *Ex-Post Protective* Method against the Risk of *Transmitting* HIV virus. I explore first whether condom is mainly perceived as a preventive protective method, or alternatively a protective method adopted following the infection by HIV virus. To do so, I conduct an heterogeneous analysis by whether a female thinks that always using condoms during sex reduces chance of getting HIV or whether she does not think so. Doing so, I intend to explore whether differences in perception of condom as a preventive methods reducing the risk of *contracting* the HIV virus leads to differences in condom use behaviour. Results are reported in column 2 of Table 15 and does not provide any evidence of such a differential behaviour: de-

[50] In absolute level, see sample mean.

Table 14: **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	0.000	0.004	-0.004	-0.004	0.021**	0.024*
	(0.005)	(0.004)	(0.012)	(0.008)	(0.011)	(0.013)
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	105,941	105,930	87,473	94,038	99,412	56,270
Adj. R-squared	0.152	0.135	0.071	0.072	0.341	0.122
Clusters	1,550	1,550	1,504	1,522	1,503	1,392
Mean Dep. Var.	0.950	0.958	0.812	0.888	0.584	0.634

Notes: OLS estimates are reported with standard errors clustered at the ethnic group × (within-country) DHS 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. "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 wource 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. ** $p<0.10$, ** $p<0.05$, *** $p<0.01$*

spite a small difference in magnitude going in the expected direction (a lower negative *matrilineal effect* for the subsample of females acknowledging condom as a protective method), both matrilineal females acknowledging the preventive function of condom and matrilineal females not acknowledging it are significantly less likely to report using condom, relative to their patrilineal counterparts. Then, I perform an heterogeneous analysis by whether a female is HIV positive or whether she is HIV negative. Such analysis is intended to assess whether women having contracted the HIV virus now use condom as an *"ex-post protective"* method, aimed at reducing the risk of *transmitting* the virus (results are reported in column 3 of Table 15). Interestingly, I find that, contrary to HIV negative matrilineal females, who are less likely to use condom than patrilineal females, HIV seropositive matrilineal females adopt a very different behaviour as they are now 1.7 percentage points significantly more likely than patrilineal females to report using condom, an effect of large magnitude given the low absolute level of use of condom (2.6%) in my final female sample[51]. This important result suggests that, when there is less room for underestimation of HIV risks, matrilineal women exert their higher ability to impose safe sexual pratices by increasing their use of condom. Further, this result provides an explanation to the absence of effect of ancestral matrilineality on contemporaneous *male* HIV, since it suggests that,

[51] This result should not be interpreted causally since reverse causality may apply here.

when they are infected, matrilineal females are more likely to use condom and therefore protect their partner.

Table 15: **Heterogeneity in Condom Use by Perception of Condom: Ex-Ante Preventive vs. Ex-Post Protective Method (OLS)**

	Condom		
	(1)	(2)	(3)
Matrilineality	-0.009**		
	(0.005)		
Matrilineality × Condom reduces HIV		-0.010*	
		(0.005)	
Matrilineality × Condom not reduces HIV		-0.016***	
		(0.006)	
Matrilineality × HIV			0.017**
			(0.008)
Matrilineality × No HIV			-0.012***
			(0.005)
Ind. Controls	Yes	Yes	Yes
Ethnic Group Controls	Yes	Yes	Yes
Village-Geographic Controls	Yes	Yes	Yes
Region-survey FE	Yes	Yes	Yes
Observations	105,964	87,473	105,964
Adj. R-squared	0.074	0.076	0.074
Clusters	1,550	1,504	1,550
Mean Dep. Var.	0.026	0.031	0.026
F-Test Equality of coeff. (p-value)		0.086	0.000

Notes: OLS estimates are reported with standard errors clustered at the ethnic group × (within-country) DHS 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. "Condom reduces HIV" is a dummy indicating whether an individual thinks that always using condoms during sex reduces chance of getting HIV. "HIV" is a dummy indicating whether an individual is HIV positive according to DHS test. "No more child" is a dummy indicating whether an individual reports not wanting another child. "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. * $p<0.10$, ** $p<0.05$, *** $p<0.01$

On the underestimation of male infidelity. An other important factor expected to determine whether a couple uses condoms in their marital sexual relations is whether the husband or wife believes that his or her spouse has been unfaithful, and that there is consequently a risk of contracting the HIV virus (Cordero-Coma and Breen (2012)). I formally explore here if matrilineal women have a higher propensity, relative to patrilineal women, of underestimating male's infidelity, and whether this affects their condom use behaviour.

In a first time, in Table 16, I regress Equation 5 on different mesures of male's and female's infidelity. Working with outcomes on female's *beliefs* about male (and female) infidelity, as well as with *actual* measures of male (and female) infidelity allow me to assess the underestimation of matrilineal females in their male counterparts' infidelity. Dummy outcomes on female's *beliefs* about infidelity are computed from answers in DHS on their attitudes towards life sex: whether they think that most married men/women have sex with their wife/husband only (columns 1 and 4)[52]. To emphasize *actual* infidelity, I create a dummy for self-reported infidelity (equals to one if an individual reports any extramarital affair in the last 12 months, columns 3 and 6); and I also compute measures of village average actual male (female) infidelity, which are simply proportions of males (females) reporting any extramarital affair in the last 12 months, averaged at the village level (columns 2 and 5).

Several lessons can be drawn from results presented in Table 16. First, despite a higher *actual* infidelity of matrilineal males (relative to patrilineal males), matrilineal females are less likely to *believe* that males are unfaithful. Indeed, according to estimate reported in column 3, matrilineal males are 2.2 percentage points more likely to report infidelity, and the average proportion of unfaithful males in matrilineal female's village is significantly higher by 2 percentage points (column 2), as compared to patrilineal female's village. Nevertheless, matrilineal females are 12.8 percentage points less likely to believe in male infidelity according to estimate reported in column 1, a large effect representing about 43% of the proportion of females believing in male fidelity.

Second, results also show that such a discrepancy does not exist regarding female infidelity. In fact, it seems that matrilineal women are more able to align their belief about female infidelity, with actual infidelity of females. Indeed, I find that matrilineal females are both significantly more likely to report unfaithful behaviour (columns 5 and 6), and, though non significant, less likely to believe in female fidelity (column 4).

One potential explanation for the tendancy of matrilineal women to underestimate male infidelity (as compared to their patrilineal counterparts), could be an adverse effect of the fact they benefit from better marriage outside options. Indeed, results presented in subsection 6.1 have shown that matrilineal women are more likely to be divorced, and are more likely to benefit from property rights in case of divorce. Smith and Watkins (2005) claim that females may use divorce as a threatening strategy for dealing with spouse that may be unfaithful and put them at HIV risk. I conjecture that matrilineal females may wrongly overestimate the efficiency of such a threatening strategy, and consequently adversely underestimate their husband's fidelity.

[52]Unfortunately, these questions have been withdrawn in DHS6 and DHS7 surveys. This explains the limited sample size when I regress on these outcomes.

Table 16: **Ancestral Matrilineality and Perception of Infidelity (OLS)**

	Male Infidelity			**Female Infidelity**		
	Most married men have sex with their wife only	Village average actual male infidelity	Self-reported Infidelity	Most married women have sex with their husband only	Village average actual female infidelity	Self-reported Infidelity
	(1)	(2)	(3)	(4)	(5)	(6)
Matrilineality	0.128**	0.020**	0.022**	-0.046	0.018***	0.027***
	(0.061)	(0.009)	(0.010)	(0.063)	(0.006)	(0.008)
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
Sample	Female	Female	Male	Female	Female	Female
Observations	4,979	88,492	75,827	4,906	105,964	105,275
Adj. R-squared	0.211	0.409	0.241	0.103	0.496	0.246
Clusters	145	1,400	1,402	141	1,550	1,550
Mean Dep. Var.	0.296	0.262	0.274	0.408	0.128	0.128

Notes: OLS estimates are reported with standard errors clustered at the ethnic group × (within-country) DHS region × survey (year) level in brackets. The unit of observation is an individual (male in column 4, female in all other columns) 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. "Most married men have sex with their wife only" is a dummy indicating whether an individual thinks that most married men only have sex with their wives. "Village average actual male infidelity" is the within village of residence's average proportion of husbands reporting any extramarital affir in the last 12 months. "Self-reported Infidelity" is a dummy indicating whether an individual reports having had any extramarital affair in the last 12 months. "Most married women have sex with their husband only" is a dummy indicating whether an individual thinks that most married women only have sex with their husband. "Village average actual female infidelity" is the within village of residence's average proportion of wife reporting any extramarital affair in the last 12 months. 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 a second time, I test the hypothesis that matrilineal women's condom use behaviour should be reversed (i.e. they should be more likely to use condom, relative to patrilineal women) when they have correct beliefs about male infidelity. To do so, I perform an heterogeneous analysis to explore how beliefs about infidelity and actual infidelity shape matrilineal women's condom use. Results presented in Table 17 confirm this hypothesis, as they underline first that matrilineal women not believing in male fidelity are, relative to their patrilineal counterparts, 4.8 percentage points more likely to use condom, a very large effect as compared to the mean proportion of 3.2% of females using condom in my regression sample (column 1). Interestingly, I also find an effect of similar magnitude for females not believing in female fidelity (column 2). Finally, consistent with my previous findings that matrilineal females have false beliefs about males' infidelity but more correct beliefs about females' one, I find that, while matrilineal females living in village with high average male infidelity are not significantly more likely to use condom (column 3), the opposite is true in villages with high actual female infidelity (column 4): I find that an increase of 1 percentage point in village average proportion of unfaithful females increases matrilineal female's use of condom by 11.3 percentage points, a very large effect as compared to the mean proportion of females using condom in my regression sample (2.6%).

8 Conclusions

This paper shows that matrilineal kinship organizations adopted by ethnic groups in pre-industrial/pre-colonial period, and related women's position within such organizations, have long lasting effect on contemporaneous female HIV in Sub-Saharan Africa. 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. 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, as well as nearest neighbor matching treament over treated effect, and find very little support for the presence of an omitted variable bias in my OLS estimates. Additionaly, 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.

Going beyond in identifying causal relationships and correcting for potential omitted variable bias, 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 robust across all these specifications.

Consistent with their higher social status in matrilineal kinship organizations, I find that matrilineal

Table 17: **Heterogeneity in Condom Use by Perception of Infidelity (OLS)**

	Condom			
	(1)	(2)	(3)	(4)
Matrilineality × Most men sex with wife only	0.012			
	(0.030)			
Matrilineality × Not most men sex with wife only	0.048**			
	(0.022)			
Matrilineality × Most women sex with husband only		0.032		
		(0.025)		
Matrilineality × Not most women sex with husband only		0.046**		
		(0.022)		
Matrilineality			-0.004	-0.028***
			(0.005)	(0.005)
Matrilineality × Village average actual male infidelity			-0.008	
			(0.009)	
Matrilineality × Village average actual female infidelity				0.113***
				(0.018)
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	4,979	4,906	88,492	105,964
Adj. R-squared	0.037	0.036	0.072	0.075
Clusters	145	141	1,400	1,550
Mean Dep. Var.	0.032	0.031	0.031	0.026
F-test Equality of coeff. (p-value)	0.139	0.333		

Notes: OLS estimates are reported with standard errors clustered at the ethnic group × (within-country) DHS 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. "Most men sex with wife only" is a dummy indicating whether an individual thinks that most married men only have sex with their wives. "Most women sex with husband only" is a dummy indicating whether an individual thinks that most married women only have sex with their husbands. "Village average actual male infidelity" is the within village of residence's average proportion of husbands reporting any extramarital affir in the last 12 months. "Village average actual female infidelity" is the within village of residence's average proportion of wives reporting any extramarital affir in the last 12 months. "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. * $p<0.10$, ** $p<0.05$, *** $p<0.01$

women benefit from a higher sexual autonomy and ability to impose safe sexual practices to their husband. In line with their higher sexual autonomy, and consistent with a recent literature showing that matrilineal women have a higher preference for risk than their patrilineal counterparts, my first mechanism highlights that, relative to patrilineal women, matrilineal women adopt riskier sexual practices which are more conducive to HIV. In line with this, I also find that, relative to patrilineal women, matrilineal women are signficantly more likely to be HIV positive while having a HIV negative husband (while the opposite does not hold true for HIV positive husbands), an important finding that suggests channels of transmission of the virus alternative to wives being simply infected by their husbands.

My second mechanism highlights a significant negative effect of originating from an ancestrally matrilineal ethnic group on female's use of male consent protective contraception methods (substituted by less protective contraception methods relying on female consent), and male condom in particular. While I show that matrilineal women benefit from a higher ability to impose safe sexual practices to their husbands, and are not significanty less likely to ackowledge HIV risk, nor less likelly to have access to male condom (in fact, the opposite is true, if any), I provide evidences that this counterintuitive results has several behavioural explanations. First, I find that, relative to patrilineal women, matrilineal women are more likely to desire more children, which goes against the use of contraception methods. Second, I provide evidences that, relative to patrilineal women, matrilineal women are more likely to have misperception about HIV infection risks: (1) I show that they are more likely to actually use condom as an *ex-post* (i.e. once contaminated) protective method against the risk of *transmitting* the HIV virus rather than an *ex-ante* preventive method against the risk of *contracting* the virus; (2) I finally show that relative to their patrilineal counterparts, matrilineal women are also more likely to understimate actual male's infidelity. I conjecture that this underestimation is an adverse effect of the fact they benefit from better marriage outside options and consequently overestimate the effectiveness of their threatening strategy against male infidelity, relative to their patrilineal counterparts.

While recent research literature and policy recomendations 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 ability to impose safe sexual practices to their husbands, are relatively more likely to adopt sexual and contraception behaviours which put them at high risk of HIV. Further, I have highlighted that, once beliefs about actual risks of HIV and actual males' sexual behaviour are correctly internalized, matrilineal women actually make use of their higher decision-making to increase their protective contraception use, as compared to their patrilineal counterparts. This promising result calls for complementary policies targeting empowered women, and aimed at raising female's awareness about the actual riskiness of their own sexual behaviour, as well as on the actual riskiness of the sexual behaviour of their male counterparts. It is hoped that such policies will induce behavioural changes restraining the spread of HIV in Sub-Saharan Africa.

References

Abadie, A. and G. W. Imbens (2006). Large sample properties of matching estimators for average treatment effects. *Econometrica 74*(1), 235–267.

Abadie, A. and G. W. Imbens (2011). Bias-corrected matching estimators for average treatment effects. *Journal of Business & Economic Statistics 29*(1), 1–11.

Alesina, A., B. Brioschi, and E. L. Ferrara (2016). Violence against women: a cross-cultural analysis for africa. *NBER Working Paper*.

Alesina, A., P. Giuliano, and N. Nunn (2013). On the origins of gender roles: Women and the plough. *The Quarterly Journal of Economics 128*(2), 469–530.

Altonji, J. G., T. E. Elder, and C. R. Taber (2005). Selection on observed and unobserved variables: Assessing the effectiveness of catholic schools. *Journal of Political Economy 113*(1), 151–184.

Anderson, S. (2018). Legal origins and female hiv. *The American Economic Review 108*(6), 1407–39.

Andrews, I. (2018). Valid two-step identification-robust confidence sets for gmm. *Review of Economics and Statistics 100*(2), 337–348.

Ashraf, N., N. Bau, N. Nunn, and A. Voena (2019). Bride price and female education. *Journal of Political Economy*.

Baland, J.-M. and R. Ziparo (2017). Intra-household bargaining in poor countries. *WIDER Working Paper*.

Bargain, O., J. Loper, and R. Ziparo (2018). Traditional norm, access to divorce and women's empowerment: Evidence from indonesia. *Working Paper*.

Becker, A. (2018). On the economic origins of constraints on women's sexuality. *Working Paper*.

Bertocchi, G. and A. Dimico (2018). The long-term determinants of female hiv infection in africa: The slave trade, polygyny, and sexual behavior. *Working Paper*.

Besley, T. and M. Reynal-Querol (2014). The legacy of historical conflict: Evidence from africa. *American Political Science Review 108*(2), 319–336.

Bjorkman Nyqvist, M., L. Corno, D. De Walque, and J. Svensson (Forthcoming). Incentivizing safer sexual behavior: Evidence from a lottery experiment on hiv prevention. *American Economic Journal: Applied*.

Cagé, J. and V. Rueda (2018). Sex and the mission: The conflicting effects of early christian investments on the hiv epidemic in sub-saharan africa. *Working Paper*.

Case, A. and C. Paxson (2013). Hiv risk and adolescent behaviors in africa. *American Economic Review: Papers & Proceedings 103*(3), 433–38.

Chimbiri, A. M. (2007). The condom is an "intruder" in marriage: evidence from rural malawi. *Social Science & Medicine 64*(5), 1102–1115.

Cordero-Coma, J. and R. Breen (2012). Hiv prevention and social desirability: husband–wife discrepancies in reports of condom use. *Journal of Marriage and Family 74*(3), 601–613.

Corno, L. and D. De Walque (2012). Mines, migration and hiv/aids in southern africa. *Journal of African Economies 21*(3), 465–498.

Delavande, A. and H.-P. Kohler (2015). Hiv/aids-related expectations and risky sexual behaviour in malawi. *The Review of Economic Studies 83*(1), 118–164.

Fortunato, L. (2012). The evolution of matrilineal kinship organization. *Proceedings of the Royal Society London*.

Fox, R. (1934). *Kinship and Marriage: An Anthropological Perspective*. Cambridge: Cambridge University Press.

Giuliano, P. (2017). Gender: An historical perspective. *The Oxford Handbook of Women and the Economy*.

González, F. and E. Miguel (2015). War and local collective action in sierra leone: A comment on the use of coefficient stability approaches. *Journal of Public Economics 128*, 30–33.

Islam, M. A., M. R. Islam, and B. Banowary (2009). 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 14*(4), 301–306.

La Porta, R., F. Lopez-de Silanes, and A. Shleifer (2008). The economic consequences of legal origins. *Journal of Economic Literature 46*(2).

Lowes, S. (2018a). Kinship systems, gender norms, and household bargaining: Evidence from the matrilineal belt. *Working Paper*.

Lowes, S. (2018b). Matrilineal kinship and gender differences in competition. *Working Paper*.

Lowes, S. R. and E. Montero (2018). The legacy of colonial medicine in central africa. *Working Paper*.

Michalopoulos, S., L. Putterman, and D. N. Weil (2019). The influence of ancestral lifeways on individual economic outcomes in sub-saharan africa. *Journal of the European Economic Association*.

Mishra, V., M. Vaessen, J. Boerma, F. Arnold, A. Way, B. Barrere, A. Cross, R. Hong, and J. Sangha (2006). Hiv testing in national population-based surveys: experience from the demographic and health surveys. *Bulletin of the World Health Organization 84*, 537–545.

Montalvo, J. G. and M. Reynal-Querol (2005). Ethnic polarization, potential conflict, and civil wars. *The American Economic Review 95*(3), 796–816.

Moscona, J., N. Nunn, and J. A. Robinson (2018). Social structure and conflict: Evidence from sub-saharan africa. *Working Paper*.

Nunn, N. (2010). Religious conversion in colonial africa. *The American Economic Review: Papers & Proceedings 100*(2), 147–52.

Nunn, N. and L. Wantchekon (2011). The slave trade and the origins of mistrust in africa. *The American Economic Review 101*(7), 3221–52.

Oster, E. (2005). Sexually transmitted infections, sexual behavior, and the hiv/aids epidemic. *The Quarterly Journal of Economics 120*(2), 467–515.

Oster, E. (2012a). Hiv and sexual behavior change: Why not africa? *Journal of Health Economics 31*(1), 35–49.

Oster, E. (2012b). Routes of infection: Exports and hiv incidence in sub-saharan africa. *Journal of the European Economic Association 10*(5), 1025–1058.

Oster, E. (2017). Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics*, 1–18.

Paula, Á. D., G. Shapira, and P. E. Todd (2014). How beliefs about hiv status affect risky behaviors: Evidence from malawi. *Journal of Applied Econometrics 29*(6), 944–964.

Smith, K. P. and S. C. Watkins (2005). Perceptions of risk and strategies for prevention: responses to hiv/aids in rural malawi. *Social Science & Medicine 60*(3), 649–660.

Sun, L. (2018). Implementing valid two-step identification-robust confidence sets for linear instrumental-variables models. *The Stata Journal 18*(4), 803–825.

Tavory, I. and A. Swidler (2009). Condom semiotics: meaning and condom use in rural malawi. *American Sociological Review 74*(2), 171–189.

Tequame, M. (2012). Hiv, risky behavior and ethno-linguistic heterogeneity. *Journal of Economics and Statistics 232*(6), 606–632.

Teso, E. (2019). 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*.

Thornton, R. L. (2008). The demand for, and impact of, learning hiv status. *The American Economic Review 98*(5), 1829–63.

Todd, E. (2017). *Où en sommes-nous ? Une esquisse de l'histoire humaine*. Éditions du Seuil.

UNAIDS (2018a). Miles to go - closing gaps. breaking barriers. righting inustices. *Joint United Nations Programme on HIV/AIDS*.

UNAIDS (2018b). Women and girls and hiv. *Joint United Nations Programme on HIV/AIDS*.

Watkins, S. C. (2004). Navigating the aids epidemic in rural malawi. *Population and Development Review 30*(4), 673–705.