Fertility and HIV Risk in Africa

Yao Yao

Washington University in St. Louis

Job Market Paper

October 2015

Abstract

This paper examines the role of social and cultural norms regarding fertility in women's HIV risk in Sub-Saharan Africa. Fertility, or the ability to bear children, is highly valued in most African societies, and premarital fertility is often encouraged in order to facilitate marriage. This, however, increases women's exposure to HIV risk by increasing unprotected premarital sexual activity. I construct a lifecycle model that relates a woman's decisions concerning sex, fertility and education to HIV risk. The model is calibrated to match Kenyan women's data on fertility, marriage and HIV prevalence. Quantitative results show that fertility motives play a substantial role in women's, especially young women's, HIV risk. If premarital births did not facilitate marriage, or women's value of children was close to the US level, the HIV prevalence rate of young women in Kenya would be one-third lower. Policies that subsidize income, education, and HIV treatment are evaluated.

1 Introduction

HIV is a major health risk facing young women in Sub-Saharan Africa (SSA). Premarital sex is arguably the leading cause of their infection. However, it is documented that behavior response to HIV risk is limited in Africa, which is especially true for young adults[1]. It is therefore of critical importance to understand what influences risky behavior and what is the obstacle to behavior change. In this paper, I explore the role of social and cultural norms regarding fertility played in African women's HIV risk. In particular, I examine how African women's motives to bear premarital children may affect their HIV infection.

Fertility, or the ability to bear children, is highly valued in most African societies. Motherhood is important for women's social status and is an essential component of married women's identity. In many cultures, childless women suffer discrimination, stigma and ostracism (Sembuya, 2010). For unmarried women, premarital fertility is often encouraged in order to facilitate marriage.

> *Since unmarried women who want to marry a man compete with his other girlfriends they may need to prove that they will be able to have children. In many African societies the birth of the first child is customarily considered an essential step in the development of a marriage ... Consequently, some women are expected to have children before marriage in order to prove their fertility to their future husband; others favour premarital pregnancy hoping that marriage will follow. (Meekers and Calvès, 1997)*

Strong incentives to bear children intensify women's exposure to HIV infection by increasing unprotected sexual activities. The risk of these activities increases if they are premarital, a time when men commonly have several partners and are less committed to the relationship. Indeed, overlapping sexual relations are considered to be a major cause of the severity of the HIV epidemic in Africa (Epstein and Morris, 2011). Unprotected sex also enhances the chance of other sexually transmitted

[1]Oster (2012) reviews the literature and finds that the limited behavior change can be partially explained by bias in OLS estimates and low non-HIV life expectancy in Africa, but these explanations apply to married people only.

infections (STIs) which are found to increase the HIV transmission rate dramatically (Oster, 2005).

To check the validity of this hypothesis, I apply Kenyan women's data from the Demographic and Health Surveys (DHS) and find that a woman who has a premarital birth is 3-6 percentage points more likely to be HIV infected, which is significant compared to Kenyan women's average HIV rate of about 8 percent. In addition, my results suggest that better-educated and wealthier women are less likely to have premarital births, suggesting a reduction of the infection risk for advantaged women through the fertility channel.

Taking the observations delineated above into account, I construct a lifecycle model that relates a woman's lifetime decisions concerning sex, fertility and education to HIV risk. The key ingredients of the model are as follows. A woman values consumption and children, and decides her sex type between committed and casual, the number of children to bear before and after marriage, and the fraction of time allocated to education versus work. Her income comes from labor income and transfers from her sexual partner. Marriage brings more income through transfers, and HIV infection lowers life expectancy and productivity. The premarital fertility decision is crucial in this model and has both positive and negative effects. On the one hand, having more children not only brings more happiness, but also increases the arrival rate of marriage. On the other hand, childrearing incurs time costs to both work and education, and the efforts to get pregnant amplify the HIV infection risk by increasing unprotected premarital sex. The sex-type choice between casual and committed sex also affects income, marriage and HIV risk. Being involved in casual sexual relationships delays marriage and increases HIV risk compared to committed sex. But it brings a fixed amount of transfers from the partner. In contrast, transfers from the partner in a committed relationship are proportional to the woman's own labor income[2]. Hence, casual sex may be more appealing than committed sex if the woman's own labor income is very low.

I allow women to be heterogeneous in their efficacy of human capital accumulation and their preference for children, and then the model has the following implications

[2]I assume that committed relationships are assortative matching.

for women's behavior. First, women who are more efficient in human capital accumulation tend to have fewer children before and after marriage, since children incur more opportunity costs. Thus they are married at a later age. These women are also more likely to be engaged in casual sex, as long as the costs in terms of HIV risk and marriage delay are not too high. This is because their labor income is very low when they are young, a time when they choose to spend more time at school. For them casual sex is a means of smoothing consumption over the lifecycle. Second, women who have higher preference for children tend to have more premarital children and hence are married earlier. They are also more likely to choose committed sex which complements premarital births in facilitating marriage.

The model is calibrated to match the data from Kenyan DHS. In particular, I categorize women by their education efficacy and preference for children, and solve their lifetime decisions numerically. The calibrated model matches data on births, marriage age and HIV prevalence for women of both high- and low-educated groups very well. Furthermore, I conduct counterfactual experiments regarding fertility motives and policy experiments of subsidizing income, education, and HIV treatment. For each type of subsidy, I also experiment on the sources of funding, say, the tax revenue from labor income tax or lump-sum tax (internal funding), and the international support (external funding). I conduct revenue-netural exercises by fixing the total amount of funding to be the same for all experiments.

Counterfactual experiments show a substantial role of fertility motives in HIV risk for women, especially *young* women. First, if premarital births did not facilitate marriage, the HIV prevalence rate of all women in Kenya (aged 15-49) would be reduced from 8.0 percent to 6.5 percent, and that of young women (aged 15-24) would be reduced from 6.6 percent to 4.4 percent. The result is similar if women's preference for children was set to be close to the US level. This is mainly due to a reduction of premarital births. Second, even though the HIV transmission rate was reduced by two-thirds, the role of fertility behavior in HIV prevalence is considerable. The number of premarital children would be increased by about 30 percent in response to the lowered HIV transmission rate, and young women's HIV prevalence rate would be reduced to 4.3 percent. However, it would be reduced to 2.2 percent if fertility

behavior was not allowed to change.

Through the policy experiments, I find that among the three types of subsidies, the education subsidy is the most effective in reducing premarital fertility and *young* women's HIV rate, since subsidizing education raises the opportunity costs of premarital childrearing. The HIV treatment subsidy is the most effective in reducing the average HIV rate of *all* women since the treatment is also an effective means of prevention. But it is not as effective in reducing the HIV rate of *young* women as the education subsidy, because it incentivizes them to have more premarital children which cause lower risk to their health now. The income subsidy has little effect on HIV prevalence even though it disincentivizes early marriage, since more income allows women to have more premarital children due to the income effect. External funds are more effective than internal funds in reducing the HIV rate, because tax lowers women's income, and incentivizes them either to bear more premarital children to facilitate marriage, or to stick to casual sex. Internal funds with labor income tax is the least effective since they lower the opportunity costs of premarital children due to the substitution effect.

Related literature

This paper is first related to the literature of the relationship between HIV and fertility. Research has looked at the effect of HIV on fertility, as the latter influences long-run economic growth by affecting population size. For example, Fortson (2009) constructs estimates of regional total fertility rate over time of SSA countries using data from the DHS, and he finds that fertility response to the HIV epidemic in SSA is insignificant. Juhn et al. (2013) find that although HIV infection significantly lowers an individual woman's fertility, local community HIV prevalence has little effect on non-infected women's fertility. On the contrary, Young (2005, 2007) finds large negative fertility responses to the HIV epidemic in Africa that contribute to welfare improvement in the long run. Kalemli-Ozcan (2008), however, finds a positive effect of HIV/AIDS on fertility in Africa in a cross-sectional, country-level framework. Sociological and medical literature shows that many HIV-infected men and women still

intend to have children, despite the health risk imposed on their partners, children and themselves (Cooper, et al., 2007; Paiva et al., 2003; Myer, et al, 2007; Nakayiwa et al., 2006; Nattabi, et al., 2009; Smith and Mbakwem, 2007; Peltzer et al., 2009). Although they do not examine the reverse effect, these studies, in general, indicate strong and pervasive fertility motives in Africa that affect behavior.

These studies, however, are subject to the reverse causality problem by ignoring the effects of people's fertility choice on their HIV risk. In contrast, I argue that not only does HIV affect fertility, but also HIV is an endogenous outcome of women's fertility choice. Furthermore, under my framework, fertility behavior affects HIV infection not only through its effect on sexual behavior, but also through its interactions with human capital accumulation and labor income status.

This paper is also related to the strand of literature that investigates socioeconomic factors that affect the HIV epidemic in Africa. For instance, economic activities such as exports (Oster, 2012) and migration (Corno and de Walque, 2012) are found to be positively associated with HIV prevalence due to the increasing concurrent sexual contacts by movements of people. On the high HIV infection rate among young African women, it is argued that transactional sex among young adults (Epstein, 2007) and premarital sexual relations (Meekers and Calvès, 1997) expose young women to high risk. Spousal search behavior, which leads to a frequent partner turnover before marriage, increases young women's HIV risk in South Africa (Magruder, 2011). Using a fidelity network model, Pongou and Serrano (2013) argue that the gender gap of HIV rates in Africa can be simply explained by the configuration of sexual networks. Nonetheless, Oster (2012) documents that despite the high HIV rates and corresponding high mortality risk associated with risky sexual behavior, behavior response has been limited. She proposes that this can be partially explained by bias in OLS estimates and low non-HIV life expectancy in Africa, but her explanation applies to married people only. The lack of behavior response among young women remains an open issue.

Much of the literature studying the relationship between HIV and education or wealth does so empirically. The results are mixed, though most support a positive relationship due to the positive association between education and premarital sex.

Using data of five SSA countries from the DHS, Fortson (2008) finds a positive relationship between education level and HIV rates in most sample countries. de Walque (2009), however, finds no correlation between the two when adding more controls using the same dataset. Alsan and Cutler (2013) attribute the rapid HIV decline in Uganda in the early 1990s to the rise in female secondary school enrollment that increased the likelihood of abstaining from sex. Duflo et al. (2015) evaluate HIV interventions in Kenyan schools and find that education subsidies, while reducing teenage pregnancies, do not reduce STIs since they do not affect casual sex. Thus they argue that distinguishing committed sex and casual sex is important for understanding policy effects on teen pregnancy and STIs. Case and Paxson (2013) argue that the effect of education on HIV infection depends on the availability of knowledge about HIV: when knowledge of HIV was unknown, regions with higher rates of female education had more prevalent nonmarital adolescent sexual activity and have higher HIV rates. Some other studies (e.g., de Walque, 2007; Lorio and Santaeulàlia-Llopis, 2011) incorporate a time-varing dimension to the analysis of the education gradient and find a dynamic relationship between HIV and education over time or over AIDS epidemic stages.

The lack of consensus on the relationship between HIV and education suggests that the channels through which education is associated with HIV is unclear. In this paper, I provide a framework in which education is associated with HIV through a fertility channel and a sex-type channel. Under the mechanism of this framework, better-educated women have a longer history of premarital sex and are more likely to choose casual sex, which increases their HIV exposure, but they tend to reduce early pregnancy, which reduces their HIV risk, all of which are consistent with the empirical findings in the literature. The overall effect of education on HIV remains a quantitative issue. Moreover, I find that even though better-educated women have fewer premarital births, their behavior is also strongly influenced by fertility motives.

In terms of methodology, Greenwood et al. (2013) offer the first quantitative general equilibrium model of disease transmission with rational decision making. In their paper, they model markets of different types of sexual activities and allow monetary transfers that clear the markets. Males and females make rational choices about

risky sexual behavior based on beliefs about its riskiness. Prior to their paper, most economic research on HIV adopts empirical methods, including field experiments (e.g., Duflo et al., 2015; Godlonton et al., 2014; Thornton, 2008). The few formal models of risky behavior and HIV/AIDS that exist are more mechanic (e.g., Kremer, 1996; Magruder, 2011). There is a large literature on HIV/AIDS transmission in epidemiology, but they do not model decision-making and take sexual behavior as exogenously given (see a review in Greenwood et al. 2013).

My paper adds to the relatively thin literature of modeling rational decision-making related to disease transmission. I offer a much richer framework than most other studies in the literature. Moreover, unlike Greenwood et al. (2013) who mainly model sexual behavior based on updated beliefs about its riskiness, I model more fundamental channels such as fertility and human capital accumulation that intensely interact with sex decisions. With this framework, I conduct policy experiments that are less stylized. In Greenwood et al. (2013), policy experiments are conducted by directly changing the value of certain parameters, for instance, the odds of infection. On the contrary, I start with the funding for policies and show how different use of the same amount of funding has different effects on behavior and HIV prevalence.

Finally, this paper is related to the broader literature of demographic transition and economic development. In particular, it is linked to the fertility literature of quality-quantity tradeoff and the scant literature of birth timing. Empirical studies find a clear negative relationship between income and fertility (e.g., Jones and Tertilt, 2008). A common interpretation is that the price of children is largely time, and hence children are more expensive for parents with higher income who have a higher demand for child quality, making quantity more costly (e.g., Becker, 1960; Easterlin, 1968; Barro and Becker, 1989; Becker, Murphy and Tamura, 1990; Wang, Yip and Scotese, 1994). The birth-timing literature mostly focuses on the relationship between fertility decisions of birth timing and income, employment and human capital accumulation. A general finding is that higher income or better education delays births (Happel, Hill and Low, 1984; Cigno and Ermisch, 1989; Conesa, 2002; Iyigun, 2000; Caucutt, Guner and Knowles, 2001; Mullin and Wang, 2002; Heckman and Walker, 1990a, b; Bloemen and Kalwij, 2001; Gutiérrez-Domènech, 2008; Tsay and Chu, 2005).

Another bulk of research focus on teenage births and the marital status of women at the time of birth (e.g., Hoynes, 1997). However, most existing models in the literature only allow binary or a small number of discrete decisions regarding fertility, and do not allow women to optimize in a dynamic setting with interactions of multiple decisions. In contrast, my model allows women to decide the number of children before and after marriage, which, combined with education and sex type choices, affects income, marriage, and health.

To summarize, this paper contributes to the literature by proposing a new angle of HIV risk components in Africa rooted in social and cultural norms regarding fertility. I offer a rich, dynamic framework of a lifetime setting in which women's sex, fertility, education and HIV infection are interrelated. This framework incorporates different channels of HIV infection, say, a fertility channel and a sex-type channel, which work differently for women with different education levels. This is the first paper that evaluates the effects of both channels on HIV risk under a theoretical framework. Furthermore, this paper has important policy implications. My results suggest that it is important for policy makers to take social factors, such as fertility motives, into account to make HIV intervention programs more effective. Moreover, I compare the effects of different policies by fixing the total amount of funding. This is highly relevant to reality since SSA countries are generally resource constrained and utilizing the funds is a critical issue.

The paper proceeds as follows. Section 2 introduces the background about HIV and fertility in Africa, and shows evidence of the relationship between the two in Kenya. Section 3 presents the model. Section 4 presents calibration, followed by quantitative analysis in Section 5, and Section 6 concludes.

2 Background and Evidence

2.1 HIV and fertility background in Africa

HIV/AIDS is a major cause of death, currently killing about 2 million people worldwide each year. The most affected continent is Africa. It is estimated that more

than 24 million people are living with HIV in Sub-Saharan Africa, and more than 1 million people die for AIDS-related cause each year, accounting for about 70 percent of all people living with HIV and of all AIDS deaths (UNAIDS, 2014). Unlike many western countries where HIV transmissions occur largely through male-to-male sexual contact and drug use, in Africa the principle mode of transmission is heterosexual contact, and most of the infected population is female.

Young women are particularly vulnerable to HIV infection compared to young men. HIV prevalence rate of young women (15-24 years old) in SSA is 3.2 percent, while that of young men is 1.5 percent (1990-2013 average. UNAIDS, 2014). This number has large variation across countries[3]. My study spotlights one part of Africa: Kenya, where HIV prevalence rate is about 7.4 percent (that of women is about 8.6 percent), higher than the average of SSA countries (about 5 percent), and the HIV rate of young women is about 4.3 percent, versus 1.5 percent of young men.

The HIV determinant this paper highlights is fertility. Fertility is highly valued in most African societies, and is a major means for women to achieve higher social status. Africa's fertility rate is much higher than many other parts of the world. In Sub-Saharan Africa, the fertility rate is 4.9 births per woman, far above the US level of 1.9 (World Bank).

Fertility and marriage are closely linked. A woman acquires an identity through marriage and, most importantly, when marriage is fertile. Inability to have children may lead to divorce or extra-marital relationships (polygamy). For unmarried women, premarital births are often encouraged to prove women's fertility and to facilitate marriage. However, premarital pregnancy is always risky, since not all births result in marriage and infidelity of men is even more common (Meekers and Calvès, 1997). In Kenya, nearly half of women who have children had their first pregnancy before the first marriage, and only about half of them had marriage follow within one year after pregnancy (the DHS for Kenya, or KDHS, 2008-2009)[4]. Better-educated

[3]For example, it is 4.1 versus 2.4 in Uganda, 7.8 versus 4.2 in Malawi, 11.5 versus 6.3 in Botswana, and 12.3 versus 6.8 in Swaziland.

[4]Premarital pregnancy is computed from the KDHS data using year and month of the first birth and age at the first union. Due to the way of computation, the number of premarital pregnancies is a lower bound estimation since miscarriages and abortions are not accounted.

women have lower premarital fertility rate. Women who have secondary school or higher education have an average of 0.74 pregnancies before marriage, while those with primary school or lower education have 1.19 on average (KDHS, 2008-2009).

Premarital sexual and fertility behavior also largely depends on ethnicity and religion. While some ethnic groups have premarital sexual relations freely permitted and subject to no sanctions, others insist on virginity and prohibit premarital sexual behavior (Murdock, 1967). Consequently, the prevalence of premarital fertility has large variation across ethnic groups in Sub-Saharan Africa (Garenne and Zwang, 2006)[5]. Kenya is a country where prevalence of premarital fertility is about 24 percent, about the median level of SSA countries. The prevalence of premarital fertility in Kenya varies substantially across ethnic groups, from 8.0 percent in Somali to 31.7 percent in Kamba (1989-2008, KDHS, computed by the author of the current paper).

Surprisingly, despite the high fertility rate in Africa, infertility is viewed as an important health issue in many African countries. In the so called "African infertility belt" that stretches across central Africa from Tanzania in the east to Gabon in the west, the infertility rate exceeds 30 percent in some countries (World Health Organization, 2003a)[6]. Major causes of infertility include sexually transmitted infections (STIs), pregnancy complications and unsafe abortion practices, which are aggravated by poverty and substandard medical care (Cates et al., 1985). Although men and women have the same rates of infertility biologically, the social burden falls disproportionately on women. A woman may be forced to divorce or enter polygamy if she

[5]Using the DHS data of 25 SSA countries, Garenne and Zwang (2006) study the relationship between premarital fertility and ethnicity in Sub-Saharan Africa. They find that the prevalence of premarital fertility, defined as the proportion of women who had premarital births, varies considerably among the 263 ethnic groups examined, from 0.1 percent (Kanem-Bornou in Chad) to 76.2 percent (Herero in Namibia).

[6]There are two types of infertility – primary infertility and secondary infertility. The former refers to couples who have not become pregnant after at least one year having sex without using birth control methods. The latter refers to couples who have been able to get pregnant at least once, but now are unable. Primary infertility rate is about 3 percent in Sub-Saharan Africa and do not vary much across countries. Most infertility in Africa is secondary infertility which varies substantially across countries, and it exceeds 30 percent in some countries. According to Larsen (2000) who adopts a stricter demographic definition of infertility, Kenya's infertility rate of women aged 20-44 is 16 percent, slightly below the median of 28 sample countries.

failed to conceive with her husband. The high infertility rates and social burdens of infertility imposed on unconceived couples in Africa provide the rationale for social expectations of women to prove their fertility before marriage.

2.2 Evidence

In this section, I show evidence on the relationship between HIV infection, premarital fertility, education and wealth in Kenya. I use data from the DHS for Kenya (KDHS). The surveys provide household and individual level data, which includes women's sexual behavior, marriage, births history and other personal information[7]. Table 2.1 summarizes statistics of the key variables using KDHS (2008). Below I show the main evidence, while leaving details of the regressions and data to the Appendix.

First, I examine the relationship between HIV status and premarital births. Table 2.2 shows the results[8]. It can be seen that in most specifications the coefficients of premarital births are positive and significant (mostly at 1 percent significance level). The magnitude is also large: a woman who has a premarital birth is 3-6 percentage points more likely to be HIV infected, which is significant compared to Kenya's HIV rate of about 7.4 percent and women's HIV rate of 8 percent. This indicates a substantial impact of premarital fertility motives on women's HIV infection risk.

I also conduct regressions using two subsamples. The first subsample includes only women who had premarital sex, and the second excludes commercial sex[9]. The

[7]I choose data of Kenya for two main reasons. First, Kenya is representative in the context of this research. Kenya's average HIV rate during 1990-2013 is about 7.4 percent (USAIDS, 2014), above the average of SSA countries. The proportion of Kenyan women who have premarital births is about 24 percent, about the median of SSA countries. Second, the refusing rate of HIV test in Kenya is relatively low (13 percent for 2003 and 8.5 percent for 2008), which reduces the potential selection bias problem. I also run regressions using data of other countries such as Malawi, Zambia and Swaziland, and most results are qualitatively and quantitatively similar.

[8]The coefficients displayed are transformed to marginal effects computed at the sample mean, so as the results below. I also examine the relationship between HIV status and premarital pregnancy and the results are very similar.

[9]I use the first subsample because in some regions or ethnic groups, premarital sexual behavior is prohibitive, and women are married at an early age, sometimes even before 15. This makes premarital births almost impossible. I exclude this group of women since the fertility motives focused in this paper are less relevant to them. The reason for using the second subsample is that female commercial sex workers may have a high chance to have premarital births due to high

results are shown in Table 2.3 and 2.4. It can be seen that the coefficients of premarital births are again significant and positive in all specifications, and the marginal effect is about 3-7 percentage points.

In addition, I examine the relationship between education, wealth, and premarital births, and Table 2.5 shows the results. As can be seen, more education is associated with lower premarital fertility, which is consistent for both education measures. Wealthier women are also less likely to have premarital births in all specifications. While not displayed in this table, women whose husbands have more wives and whose rank among the wives is lower are more likely to have premarital births.

In summary, casual empirical results shows that women who have premarital births are 3-6 percentage points more likely to be HIV infected, and women who are better-educated or wealthier are less likely to have premarital births.

3 A Lifecycle Model

In this section, I construct a continuous-time lifecycle model that relates a woman's decisions concerning sexual behavior, fertility decision and education investment to HIV risk. A woman values consumption and children, and chooses sex type, the number of pre- and post-marital children, and education time to optimize her lifetime utility. Sexual relationships bring her transfers from the partner, and children incur time costs to work and education. During her lifetime, she may get married, which brings her more transfers from the partner. She may also be HIV infected, which lowers her life expectancy and productivity. Premarital births facilitate marriage, but increase HIV risk. Casual sex delays marriage and increases HIV risk compared to committed sex, but may bring more transfers from the partner if the woman's own labor income is very low.

frequency of sexual activities and possibly higher pay if the sex had no protection. Their premarital births are very likely unwilling outcomes and are not relevant to the fertility motives underscored in this paper.

3.1 Model setup

Time is continuous. It starts from a woman's beginning of fertile age (F), and ends at the death date (D). A woman values her consumption and children. At the beginning of her fertile life, she decides the total number of children to give birth to during her life (n), the number of children before marriage (n_1), and sex type (s, committed or casual). At each point of time she chooses the fraction of time allocated to education (q)[10]. Her lifetime utility function is

$$V = \int_F^D [ln(c) + \theta ln(1 + n_1 + n_2 \cdot \mathbf{1}(M)) - \delta ln(1 + n_1) \cdot (1 - \mathbf{1}(M))]e^{-\rho t}dt \quad (1)$$

where n_2 is the number of children after marriage (i.e., $n_2 = n - n_1$), $\mathbf{1}(M)$ is the indicator of being married, and ρ is the discount factor. θ measures her preference for children relative to her preference for consumption, and δ measures the disutility cost of premarital children due to the negative social aspect of premarital fertility[11]. I assume that once one is married, she will not divorce or separate, that is, marriage lasts until her death.

A woman maximizes her utility subject to the following conditions.

The budget constraint:

$$c + \phi y(n_1 + n_2 \cdot \mathbf{1}(M)) = y + x \quad (2)$$

Women are hand-to-mouth[12]. Their living costs come from two sources: consumption and childrearing. The childrearing cost is modeled as a fraction ϕ of the labor income

[10] I assume that women start sexual activities from the beginning of their fertile age. So there is no choice between abstinence or not.

[11] The disutility cost of premarital children is a reduced form of another aspect of the social norm of premarital fertility. In particular, a woman faces risk when bearing premarital children, say, there is a possibility that she would be abandoned by her partner after her pregnancy or child birth. In this case, her reputation may be hurt as premarital fertility can be a signal of flighty sexual behavior, which would also hurt her future marital prospect.

[12] I abstract from asset accumulation decision in this setting for simplicity which is also relevant to African women many of whom live around the poverty line and do not have much savings.

per child, as literature shows that childrearing costs are largely time costs. y is one's labor income at time t if she had no children and x is the transfer she gets from her sexual partner. y and x are determined as follows.

$$y = wh(1-q)(1-\eta \cdot \mathbf{1}(I)) \tag{3}$$

where w is the hourly wage per efficient unit of labor, h is one's human capital, q is the fraction of time allocated to education and thus $1-q$ is the fraction of time to work, η is the percentage of productivity drop due to the HIV infection, and $\mathbf{1}(I)$ is the indicator of the HIV infection.

$$x = \begin{cases} x_0 \cdot (1-\mathbf{1}(M)) + wh(1-q)(\varphi_0+\varphi_M)\cdot \mathbf{1}(M), & if\ s=0\ (casual\ sex) \\ wh(1-q)(\varphi_0+\varphi_M \cdot \mathbf{1}(M)), & if\ s=1\ (committed\ sex) \end{cases} \tag{4}$$

that is, there are two types of transfers from one's sexual partner depending on the sex type she chose. If she chose casual sex, she gets a constant transfer x_0 before marriage regardless of her own income status, and a proportion $(\varphi_0+\varphi_M)$ of her own labor income after marriage (assuming marriage is an assortative matching). If she chose committed sex, she is matched to a partner assortatively and receive a proportion φ_0 of her own labor income before marriage, which increases by a proportion φ_M after marriage. I assume that the sex type decision is only relevant before marriage – once married, the woman always has committed sex.

Finally, human capital accumulation are determined by

$$\dot{h} = \Phi q h[1-\gamma(n_1+n_2\cdot \mathbf{1}(M))](1-\eta\cdot\mathbf{1}(I)) \tag{5}$$

which depends on Φ, efficacy of human capital accumulation, q, the fraction of time allocated to education, the number of children (n_1 before marriage and $n=n_1+n_2$ after marriage), and her HIV status. Childrearing costs to education also appears as time costs (a fraction γ per child). In particular, it is the custom in Africa that one would be expelled from school upon pregnancy or the child birth, hence γ is expected

to be large. Moreover, HIV infection reduces one's productivity of education by a percentage η.

There are three important events in one's life that arrive at Poisson arrival rates, and there is no particular order for them.

One is marriage (M), the arrival rate of which is determined by

$$\lambda_M = \overline{\lambda}_M(1+\alpha n_1)(1+\zeta s)\cdot(1-\mathbf{1}(M)) \tag{6}$$

that is, premarital children increase the arrival rate of marriage, and casual sex reduces the chance of marriage due to flawed reputation.

The second one is HIV infection which is determined by

$$\lambda_I = \overline{\lambda}_I[1+n_1^{\beta}\cdot(1-\mathbf{1}(M))][1+\kappa(1-s)\cdot(1-\mathbf{1}(M))] \tag{7}$$

that is, premarital fertility and casual sex increase the probability of HIV infection. Even if one has no premarital births and chooses committed sex, she is still subject to the risk (with the arrival rate $\overline{\lambda}_I$) since her partner may be infected. I abstract from more sources of HIV risk during postmarital periods since this paper focuses on HIV of young women and premarital behavior.

The last one is the death arrival at the rate $\overline{\lambda}_D$ if one is not HIV infected and $\overline{\lambda}_A$ if one is infected ($\overline{\lambda}_A > \overline{\lambda}_D$).

3.2 Model characterization

In the model economy, a woman's problem is to choose her fertility n and n_1, sex type s and fraction of time for education q to maximize her lifetime utility given by equation (1) subject to the budget constraint and human capital accumulation given by equations (2) - (5), taking marriage and HIV risk into account.

There are two binary state variables (m, i) corresponding to marriage and HIV status (there is another state variable – sex type s when premarital), and three continuous state variables (h, n, n_1). Therefore, there are four possible (binary) states in a woman's life regarding marriage and HIV status: (M, I), $(M, -I)$, $(-M, I)$,

$(-M, -I)$. The HJB equations of the four states are as follows[13].

$$\begin{aligned}\rho V(h|M,I;n,n_1) &= max_q\{ln(c) + \theta ln(1+n) \\ &+ \overline{\lambda}_A[-V(h|M,I;n,n_1)] + \frac{\partial V(h|M,I;n,n_1)}{\partial h} \cdot \dot{h}\}\end{aligned} \tag{8}$$

$$\begin{aligned}\rho V(h|M,-I;n,n_1) &= max_q\{ln(c) + \theta ln(1+n) \\ &+ \overline{\lambda}_I[V(h|M,I;n,n_1) - V(h|M,-I;n,n_1)] \\ &+ \overline{\lambda}_D[-V(h|M,-I;n,n_1)] + \frac{\partial V(h|M,-I;n,n_1)}{\partial h} \cdot \dot{h}\}\end{aligned} \tag{9}$$

$$\begin{aligned}\rho V(h|-M,I;n,n_1,s) &= max_q\{ln(c) + (\theta-\delta) ln(1+n_1) \\ &+ \overline{\lambda}_M(1+\alpha n_1)(1+\zeta s)[V(h|M,I;n,n_1,s) - V(h|-M,I;n,n_1,s)] \\ &+ \overline{\lambda}_A[-V(h|-M,I;n,n_1,s)] + \frac{\partial V(h|-M,I;n,n_1,s)}{\partial h} \cdot \dot{h}\}\end{aligned} \tag{10}$$

$$\begin{aligned}\rho V(h|-M,-I;n,n_1,s) &= max_q\{ln(c) + (\theta-\delta) ln(1+n_1) \\ &+ \overline{\lambda}_M(1+\alpha n_1)(1+\zeta s)[V(h|M,-I;n,n_1,s) - V(h|-M,-I;n,n_1,s)] \\ &+ \overline{\lambda}_I(1+n_1^{\beta})(1+\kappa(1-s))[V(h|-M,I;n,n_1,s) - V(h|-M,-I;n,n_1,s)] \\ &+ \overline{\lambda}_D[-V(h|-M,-I;n,n_1,s)] + \frac{\partial V(h|-M,-I;n,n_1,s)}{\partial h} \cdot \dot{h}\}\end{aligned} \tag{11}$$

Equation (8) – (11) say that when one is in state (M, I) she is only possible to

[13]Note that in this model, death arrives at a constant arrival rate $\overline{\lambda}_D$ which is independent of women's age, so there is no term $\frac{\partial V}{\partial t}$ in the HJB equations.

On the notation of binary states, M (I) represents the state of being married (HIV-infected), and $-M$ $(-I)$ represents the state of being unmarried (non-infected)

change into the state of death (equation(8)); when she is in state $(M, -I)$, she may change into HIV infection state with an arrival rate $\overline{\lambda}_I$ or death (equation (9)); when she is in state $(-M, I)$, she may get married with an arrival rate $\overline{\lambda}_M(1+\alpha n_1)(1+\zeta s)$ or die (equation (10)); and when she is in state $(-M, -I)$, she may enter marriage, be HIV infected or die (equation (11)).

The Bellman equations (8)-(11) can be used to characterize the choice of education time q. For example, by taking first order condition of equation (8) one gets:

$$\frac{1}{h(1-q)} = \frac{\partial V(h|M, I; n, n_1)}{\partial h} \cdot \Phi(1-\eta)(1-\gamma n) \tag{12}$$

which says that suppose $\frac{\partial V(h|M,I;n,n_1)}{\partial h}$ is fixed, then q increases in education efficacy Φ, and decreases in productivity drop due to HIV η and the number of children n. This is intuitive since the higher the education efficacy, the more beneficial is the time spent on education; the more productivity drop due to HIV, the less does one benefit from education; the more children one has, the less efficient are both work and education. The characterizations of q of other states are similar.

The fertility decision is more complicated since it needs to be solved back to the beginning of one's life and there is no analytical solution to the value functions. But qualitatively it is affected by the following factors. First, the benefit and cost of children are happiness they bring about versus childrearing costs. For the former, higher income increases fertility since children can be viewed as normal goods (i.e., the income effect). But for the latter, since childrearing costs are time costs, women with higher education efficacy (Φ) (hence higher potential labor income) would have fewer children due to higher opportunity costs of childrearing (i.e., the substitution effect). Second, the choice of premarital children has some additional effects. The additional benefit is that premarital births facilitate the arrival of marriage, and the additional costs are that they increase the chance of HIV infection, and incur a disutility cost. Hence, women with higher education efficacy may also have fewer premarital children due to higher opportunity costs of childrearing, as long as the costs of marriage delay is not too high. Finally, the higher chance do premarital births bring about marriage (i.e., higher α), the more premarital children would

women have. The more likely do they cause HIV infection (i.e., higher β), the more costly is HIV (i.e., higher η or $\overline{\lambda}_A$), or the more disutility does premarital fertility incur (i.e., higher δ), the fewer premarital children would be delivered by women.

The sex type decision is also made at the beginning of life and is affected by the following. The benefit of casual sex is that it brings more current transfers from the partner when the woman's labor income is so low that the proportional transfers from her partner in a committed relationship are lower than what she makes in a casual relationship. The cost is that it delays marriage due to flawed reputation and increases HIV risk. Therefore, a woman with higher education efficacy (i.e., higher Φ) may choose casual sex since she receives less from a committed relationship when she spends much time in school and has low labor income (hence low transfers proportional to her own income). But she may also choose committed sex if casual sex delayed marriage by so much that she would lost more (proportional) transfers from marriage when she has accumulated enough human capital and spends more time on working (thus has higher labor income). Moreover, premarital children and committed sex are complements in facilitating marriage (see equation (6)), hence women who have more premarital children are more likely to choose committed sex, *ceteris paribus*.

4 Calibration

I calibrate the model based on Kenyan women's data from the DHS (2008). I group all women aged 15-49 into four categories depending on their education efficacy (Φ_H and Φ_L) and preference for children (θ_H and θ_L), hence there are two-by-two categories in total. The former is determined by education level which is secondary school or higher (Φ_H) versus primary school or lower (Φ_L), with a population ratio of $1/2$ based on the KDHS (2008). The latter is determined by the number of children (i.e., a larger number of children reveals higher preference for children conditional on the same education level), and the population ratio of θ_H over θ_L is $2/1$, which is taken based on the population ratio of different sex types that will be discussed later.

There are 21 parameters to be determined: θ_H, θ_L, ρ, δ, ϕ_H, ϕ_L η, φ_0, φ_M, x_0, Φ_H, Φ_L, γ, $\overline{\lambda}_M$, α, ζ, $\overline{\lambda}_I$, β, κ, $\overline{\lambda}_A$, $\overline{\lambda}_D$. The main strategy is that I first compute most of the parameters from literature or data, and then solve the rest ones by quantitatively solving the model using value function iterations to match data moments. I transform the continuous-time model to a discrete-time one to solve the value function and use one year as the base period. The details are as follows.

Calibration of ρ, η, $\overline{\lambda}_D$, $\overline{\lambda}_A$, $\overline{\lambda}_M$. These five parameters are taken or computed directly from literature or data. ρ, the discount factor is set to be 0.05, as is typical in the literature. η, the percentage of productivity drop due to HIV, is set to be 0.05 as Manuelli (2015) does. For the parameters of the three arrival rates, $\overline{\lambda}_D$ is computed from female life expectancy at birth in Kenya, which is 58.9 (World Bank, 2003-2013 average). Since the fertile age starts from menarche the median of which is 15 (Leenstra et al., 2005), $\overline{\lambda}_D = 1/(58.9 - 15) = 0.023$. $\overline{\lambda}_A$ is computed from the mean life expectancy from the time of being HIV infected. In Africa, the average spacing between HIV infection and AIDS is 9.4 years, and between AIDS to death is 9 months (Morgan et al., 2002), hence $\overline{\lambda}_A = 1/(9.4 + 0.75) = 0.099$. $\overline{\lambda}_M$ is the arrival rate of marriage if the woman had no premarital children and chose committed sex. Of the four categories of women the highest mean marriage age is 23.47 (group (θ_L, Φ_H)). In data, this group of women have a positive mean number of premarital children and are most likely to have casual sex. Thus I reasonably assume the mean marriage age of committed-sex women without premarital children to be 24, which generates $\overline{\lambda}_M$ to be $1/(24 - 15) = 0.111$.

Calibration of Φ_H, Φ_L, γ. These are parameters related to human capital accumulation and are computed using equation (5) $\frac{\dot{h}}{h} = \Phi q[1 - \gamma(n_1 + n_2 \cdot \mathbf{1}(M))](1 - \eta \cdot \mathbf{1}(I))$. I use this equation for two groups of women with high and low education efficacy (Φ_H and Φ_L), and employ the average of Kenyan women's education efficacy Φ_{ave} (0.08) computed from Schultz (2003, Table 5)[14], hence obtain three equations to

[14]The average education efficacy Φ_{ave} is computed using Table 5 of Schultz (2003) by assuming full-time schooling education (i.e., $q = 1$) and no birth or HIV infection, hence $\Phi_{ave} = \frac{\dot{h}}{h}$. Since in my model the source of wage differences is all human capital differences, I equal wage growth rate to human capital growth rate. I first compute implied private wage returns in percent per annum of primary to secondary education (8.9) and secondary to university education (16.9) using

compute these three parameters. I approximate the growth rate of human capital $\frac{\dot{h}}{h}$ of the two groups of women with the growth rate of their labor income $\frac{\dot{y}}{y}$ by assuming hypothetical balanced growth path, and proxy labor income y by wealth index from the KDHS (use data of 2003 and 2008 to compute the growth rate)[15]. I use years of schooling of the two groups to compute q, and the number of children from data for n_1 and n_2 of the two groups (KDHS, 2008)[16].

Calibration of φ_0, φ_M, x_0. These are parameters about transfers from the sexual partner. I first use the equation that computes the mean income transfers from nonmarital partners for all women $Pr(casual) \cdot x_0 + Pr(committed) \cdot wh\varphi_0 = 12.5\% \cdot w_M h_M$, where $Pr(casual)$ ($Pr(committed)$) is the proportion of non-married women who have casual (committed) sex, w_M (h_M) is the average efficiency wage (human capital) of men, and 12.5% is the average fraction of a man's income that is transferred to his nonmarital sexual partner computed from Luke (2005). I also use the equation $\varphi_0 + \varphi_M = \frac{1+lc_M/lc_F}{2} - 1$, where lc_M (lc_F) is the average labor income

his regression coefficients, and then compute human capital h of each education level. I normalize human capital of women with primary or lower education to be one, then human capital of secondary education equals one plus years of secondary schooling multiplied by the wage returns per annum from primary to secondary education computed above ($h_{sec} = 1.54$), and human capital of university education equals h_{sec} plus years of university schooling multiplied by the wage returns per annum from secondary to university education computed above ($h_{unv} = 2.21$); finally $\frac{\dot{h}}{h}$ is the average of the growth rate of human capital from primary to secondary education and from secondary to university education, computed using the computation results from above.

[15]I use wealth index (variable hv270) from the KDHS 2003 and 2008 to proxy women's labor income. For each year, average wealth index of more- and less-educated women (aged 15-19 for 2003, and 20-24 for 2008) are employed to proxy average labor income of these two education groups, and then growth rate of labor income is computed based on the increase of wealth index from 2003 to 2008 of each group. The results are that the growth rate of income of more-educated women is 3.07 percent, and of less-educated women is 0.50 percent.

[16]The fraction of one's time allocated to schooling q is computed using fraction of schooling years over a certain period for each education group. The years of schooling is about 12 years for the better-educated group and 6 years for the less-educated group. Assuming schooling starts at the age 7, then it ends at ages 19 and 13 respectively for the two groups. Hence q for group Φ_H and group Φ_L can be computed by: $q_H = (19-17)/(22-17) = 2/5$, and $q_L = (13-12)/(17-12) = 1/5$. That is, for the Φ_H group, for the sample used to compute income growth rate (aged 15-19 in 2003 and 20-24 in 2008) a typical woman ages 17 in 2003 and 22 in 2008, and during these five years she continued education until the age 19. For women in the Φ_L group, since they do not have education after age 15, I assume their income growth rate from age 12 to 17 is the same as that from age 17 to 22, and during the five years from age 12 to 17, they continued education until the age 13.

of men (women). This equation assumes that a woman equally shares the household income with her husband after marriage. Additionally I assume x_0 to be 0.6, which is 60 percent of labor income of women with primary education or lower per unit of time. This number is chosen to generate reasonable proportions of two sex types consistent with data. The data that I use for the above computation are the following. $Pr(casual)$ and $Pr(committed)$ are taken from the KDHS data. However, there is no clear data about a woman's sex type. Greenwood et al. (2013) define casual sex as having sex with nonmarital or non-cohabiting partner. Using their definition, I obtain the proportion of casual sex to be 67.6 percent. But it may overestimate the proportion of casual sex type since sex with non-cohabited boyfriend may also be committed. I therefore choose the population ratio of θ_H/θ_L to be $2/1$ and presume all groups except for (θ_L, Φ_L) group choose casual sex, which is relatively more consistent with the KDHS data for sex types of different groups. This produces the proportion of casual sex to be $5/9$ (or 55.6 percent, a smaller number than 67.6 percent). Women's efficiency wage w is normalized to be one, and men's efficiency wage w_M equals male-to-female hourly wage ratio in South Africa taken from Winter (1999). Average human capital of women and men (h and h_M) is computed using years of schooling from the KDHS and Φ_{ave} (see footnote 14), and labor incomes lc_M and lc_F are computed accordingly by multiplying the efficiency wage and human capital taking labor participation rates into account [17].

Calibration of ϕ_H, ϕ_L. These are parameters about childrearing costs to work. I allow heterogeneity of ϕ for the two different education groups Φ_H and Φ_L since literature shows that better-educated people have higher demand of quality of children, hence spend more time on childrearing for each child. I calibrate the two parameters by transforming equation (2) for postmarital stage to $\frac{c}{y} + \phi n = 1 + \frac{x}{y} = 1 + \varphi_0 + \varphi_M$, where $\frac{c}{y}$ is the consumption-output ratio taken from PennWorld Table (6.3), and the number of children n for the two groups are taken from the KDHS.

Calibration of α, ζ, β, κ. These are parameters of Poisson arrival rates of marriage

[17] The female/male hourly wage ratio is 0.85 (Winter, 1999, Table 11), and labor participation rates of female and male are 27.2 percent and 42.9 percent respectively (Winter, 1999, Table 3). Hence ratio of labor income ratio lc_M/lc_F equals the multiplication of male/female hourly wage ratio, human capital ratio and labor participation ratio.

and HIV infection. α and ζ are calibrated from equation (6) using the mean marriage age and the number of premarital children of the two education groups (Φ_H and Φ_L) and $\overline{\lambda}_M$. β and κ are calibrated from equation (7) using HIV prevalence rates and the number of premarital children of the two education groups and $\overline{\lambda}_I$. $\overline{\lambda}_I$ is calibrated from quantitatively solving the model using value function iterations.

Calibration of $\overline{\lambda}_I$, δ, θ_H, θ_L. I calibrate these last four parameters by solving the model quantitatively. I transform the continuous-time model to a discrete-time one, discretize the space of n, n_1 and h, and use value function iterations to solve policy functions of each group of women. In particular, I use the simplex-search method to find the parameter values to match the four targets: n and n_1 of Φ_H and Φ_L groups, by minimizing the distance of n and n_1 of the two education groups from policy functions and data.

Table 4.1 shows the calibrated parameter values, and Table 4.2 shows the model and targeted values of certain variables, which are matched very well. Furthermore, I simulate HIV prevalence rates using the calibrated parameters, and obtain the HIV prevalence rate of all women to be 8.01 percent, comparable to the data value of 9.45 percent, and the HIV rate of young women to be 6.56 percent, very close to the data 6.96 percent [18]. HIV prevalence rates by age of the model and the data (premarital sex only, five-year moving average) are shown in Figure 4.1. It can be seen that the HIV prevalence rate from the model is matched to the data particularly well for young women.

5 Quantitative Analysis

5.1 Counterfactual analysis

In this section, I conduct counterfactual analysis to examine the effect of fertility motives quantitatively.

The first set of experiments either turns off the channel that premarital births

[18] The HIV prevalence rates of data are for women with premarital sex only, who are more relevant to my model.

facilitate marriage or alters women's preference for children. The first experiment makes premarital births having no effect on marriage by setting α to be zero and $\overline{\lambda}_M$ equal λ_M of the benchmark model excluding the multiplier of sex type (i.e., $\overline{\lambda}_M = \overline{\lambda}_{M\_BM}(1 + \alpha_{BM} n_{1\_BM})$). The second one chooses θ that makes women's income elasticity of fertility close to that of the US[19]. The third experiment sets δ to be equal to θ_L so that for the group of women with low preference for children, the disutility cost premarital fertility totally offsets the happiness from premarital children.

Table 5.1 shows the results. It can be seen that removing the effect of premarital births on marriage arrival rate ($\alpha = 0$) or increasing disutility cost premarital births ($\delta = 0.87$) have little effects on the total number of children since they do not directly affect postmarital behavior, while reducing preference for children ($\theta = 0.55$) reduces the total number of children by about 40 percent. In all experiments, the number of premarital children (n_1) is changed significantly. When premarital children do not facilitate marriage, or when disutility cost of premarital fertility is larger, women choose to have only about 0.7 premarital children instead of 1.05. Premarital fertility is even lower when women have lower preference for children. Interestingly, better-educated women are more affected in these experiments as their number of premarital children drops more significantly than that of less-educated women, hence their HIV infection rate also drops significantly. This suggests that even though better-educated women have fewer premarital children, their behavior and health are also deeply influenced by fertility motives. In addition, sex type of some groups of women are changed in the second and the third experiments. In the second one, women of (θ_H, Φ_L) type choose casual sex instead of committed sex (hence all women choose casual sex) since they are now willing to have fewer premarital children (and fewer

[19]The US income elasticity of fertility is -0.38, taken from Jones et al. (2008). I use the simplex search method to find the θ that makes income elasticity of fertility closest to the US level, which produces the value of θ to be 0.55 and income elasticity of fertility to be -0.51. The reason that I cannot find a value of θ that makes income elasticity of fertility exactly equal the US level is related to the way I compute the elasticity: there are only two groups of women with different income (Φ_H and Φ_L) when setting θ identical for all women, and hence the income elasticity of fertility is computed only based on difference of fertility and income (or human capital) of the two groups, i.e., $\epsilon_y = \frac{n_H - n_L}{h_H - h_L} \cdot \frac{h_{ave}}{n_{ave}}$.

total children as well), and premarital children and committed sex are complements in facilitating marriage. In the third experiment, women of group (θ_L, Φ_L) choose committed sex instead of casual sex, since now premarital fertility is less favored, but choosing committed sex helps bringing about marriage after which they would enjoy more children. As a consequence, HIV prevalence rates decline in all experiments. HIV prevalence rates of all women drop to about 6.5 percent from 8.0 percent, and that of young women drop even more, to about 4.4 percent from 6.6 percent.

Result 1. *Had premarital births not facilitated marriage, or women's value of children been close to the US level, the HIV prevalence rate of young women in Kenya would have been one-third lower.*

In the second set of experiments, I look at how behavior matters for HIV prevalence. In particular, I change the HIV transmission rate by setting $\overline{\lambda}_I$ to be one-third of its original value, since it is argued that the large difference of HIV prevalence rate between Africa and the western world comes from differences in transmission rates, and that the transmission rate per partnership in Africa is about three times of that in the US (Oster, 2005). I examine how this change of HIV transmission rate affects HIV prevalence when behavior response is allowed or not. Table 5.2 shows the results. It can be seen that when HIV transmission rate is lower, premarital children increase by about 30 percent when behavior is allowed to change, and the HIV prevalence rate is reduced to about 4 percent for all women and young women. However, when behavior is not allowed to change, that is, when the number of (premarital) children and sex type are fixed to be the same as in the benchmark model, the HIV prevalence rate can be reduced to 2.7 percent for all women and 2.2 percent for young women. This suggests that even though the HIV transmission rate matters, fertility behavior plays a considerable role in HIV prevalence in Africa.

Result 2. *Had the HIV transmission rate been two-thirds lower, the HIV prevalence rate of young women in Kenya would have been one-third lower. But this rate would have been two-thirds lower had sexual and fertility behavior been fixed.*

Figure 5.1 and 5.2 show HIV prevalence rates by age of these experiments and the benchmark model.

5.2 Policy experiments

In this section, I conduct policy experiments of the income subsidy, the education subsidy and the HIV treatment subsidy to examine their effects on women's fertility choice and HIV risk. For each type of subsidy, I experiment on sources of funding, including the internal funding (domestic tax revenues) and the external funding (the international support). For the internal funding, I experiment with a labor income tax and a lump-sum tax. Compared to the lump-sum tax, the labor income tax adds a substitution effect that makes childrearing less costly. Income affects premarital children in two opposite directions. Lower income may reduce premarital children when children are viewed as normal goods. But it may also increase premarital children since they facilitate marriage, and marriage brings more income. For comparison purposes, all policy experiments are conducted on a revenue-neutral basis in the sense the total amount of subsidy is the same across all experiments.

5.2.1 Income subsidy

The effect of the income subsidy is examined since in my model women's fertility motives largely come from income incentives – premarital fertility brings about marriage which improves their income status. I conduct five income-subsidy policy experiments which differ in the source of funding and groups of women to subsidize. In particular, the funding source can be either the internal funding (tax revenues) or the external funding (the international support that generates a positive income effect), and women to subsidize can be either women with lower education efficacy (hence lower education and income) or all women. For experiments that subsidize only low-income women with tax revenues, I experiment with a labor income tax and a lump-sum tax. For the first experiment (labor income taxes, subsidizing low-income women), all labor income is taxed with a 5 percent rate[20]. The total amount of tax revenues in this experiment is used as a benchmark for all policy experiments. That is, for other experiments, I keep the total amount of funding the same as in the

[20]Kenya's labor income tax follows a progressive rule and the tax rates ranges from 10 to 30 percent (Kenya Revenue Authority, 2007). I take 20 percent, the average of the tax rates, and assume one-fourth of the tax revenue is used for income subsidy, that is, a tax rate of 5 percent.

first experiment and compute the subsidy rate (and the tax rate in internal-funding experiments) to make the total amount of subsidies (and tax revenues in internal-funding experiments) equal the total amount of funding[21]. Lump-sum subsidies are applied in the income subsidy experiments[22].

Table 5.3 shows results of fertility and HIV rates of income-subsidy experiments. It can be seen that the total number of children increases in nearly all experiments. In the external-funding experiments, women give births to more children due to the income effect, and in the internal-funding, labor income tax experiments even more children are produced because opportunity costs of childrearing become lower (i.e., the substitution effect dominates the income effect).

The number of premarital children also increases in nearly all experiments, though very slightly. For low-income women, the number of premarital children increases in the EL experiment since more of them choose committed sex, and premarital children and committed sex are complements in facilitating marriage. It also increases in the internal-funding, labor income tax experiments due to lowered childrearing costs to work. For high-income women, premarital fertility increases under the internal-funding, labor income tax policies due to the substitution effect of tax. It also increases under the EU policy since more income allows more high-income women to choose committed sex which, combined with premarital children, increases the chance of marriage.

As a consequence, HIV prevalence rates decline under the external-funding policies for all and for young women, but it increases under the internal-funding, labor income tax policies, though the magnitude of changes is moderate in all experiments. Figure 5.3 shows the HIV prevalence rates by age of the benchmark model and the five income subsidy experiments.

These results suggest that the income subsidy may not be an effective means to reduce HIV rates, because even though higher income disincentivizes early marriage and hence premarital births, it allows more premarital children due to the income

[21]For all policy experiments, the simplex search method is used to find the subsidy and/or tax rate.

[22]The internal-funding policy experiments assume that subsidies to women are all funded by tax revenues from women, i.e., there are no funds from men.

effect. Subsidizing with external funds is more effective than with internal funds, because taxing labor income reduces opportunity costs of childrearing that encourages women to bear more premarital children, and the external funding has the potential of switching women's sex type toward committed sex.

Result 3. *An income subsidy policy would have little effect on women's premarital fertility behavior and HIV prevalence in Kenya.*

5.2.2 Education subsidy

The effect of the education subsidy is examined since data shows that better-educated women have fewer premarital children, which lowers their HIV risk through the fertility channel. Three experiments are conducted. In the first two, the subsidies are funded with labor income taxes and lump-sum taxes respectively (i.e., the internal funding), and in the third one they are funded with the international support (i.e., the external funding). In all experiments, women who receive school education are subsidized proportionally to their education time (q). The tax rate and the subsidy rate are computed so that the total amount of subsidies (or the total amount of tax revenues) equals that in the income subsidy experiments.

Table 5.4 shows the results. It can be seen that the total number of children increases under the IU_{lc} policy, which is mainly due to an increase of fertility of less-educated women, whose opportunity costs of childrearing become even lower due to the tax on their labor income (the education time of women with lower education efficacy does not increase even though education is subsidized, hence the only beneficiaries of this policy are women with higher education efficacy). The total number of children decreases under the EU policy as fewer children are given birth to by more-educated women; it decreases more in the IU_{ls} policy since less-educated women also produce fewer children due to the income effect.

The number of premarital children declines significantly for women with higher education efficacy under all policies, since subsidizing education increases their opportunity costs of having premarital children. But premarital fertility behavior of those with lower education efficacy is inelastic to the policy (actually increases slightly

in the internal funding experiments, because lower income incentivizes premarital births to facilitate marriage). Women's sex type does not change in any of these experiments. As a consequence, HIV rates of all women decline from 8.0 percent to 7.3 percent (a 9-percent reduction), and HIV rates of young women decline from 6.6 percent to 5.6 percent (a 15-percent reduction). The decline almost all comes from fertility behavior changes of women with higher education efficacy. The HIV rate under the *EU* policy is slightly lower than under the *IU* policies. Figure 5.4 shows the HIV prevalence rates by age of the benchmark model and the three education subsidy experiments.

These results suggest that first, education-subsidy policies are more effective than income-subsidy policies in reducing premarital fertility and HIV rates, since they increase opportunity costs of childrearing to education. Second, with greater benefit, women with higher education efficacy respond more to the education-subsidy policy since this policy is more beneficial to them. Finally, the education subsidy funded by the international support is more effective than that funded by domestic tax revenues, as the latter reduces income of less-educated women and encourages them to bear more premarital children to facilitate marriage.

Result 4. *An education subsidy policy would reduce the HIV prevalence rate of young women in Kenya by 15 percent, primarily driven by a reduction of premarital children by women with higher education efficacy.*

5.2.3 HIV treatment subsidy

This set of experiments subsidizes HIV treatment (antiretroviral therapy, or, in short, "ART"). Again, three experiments are conducted depending on the source of funding, i.e., the internal funding – either labor income taxes or lump-sum taxes, and the external funding – the international support. The funding is given to all HIV-infected women for HIV treatment in the form of lump-sum subsidies. The tax rate and the subsidy rate are computed to make the total amount of subsidy (or tax revenues) equal to that in the income subsidy experiments.

In order to know the effectiveness of the subsidy on a patient's health, one needs

to know the price and the effectiveness of the therapy. In recent years the price of HIV drugs have dropped significantly and a generic copy of a triple-therapy antiretroviral costs about 300 US dollars per patient per year. This transforms to about 37.5 percent of Kenya's GDP per capita. Jordan et al. (2012) show that using three antiretrovirals reduced progression to AIDS or death by about 75 percent. Since WHO recommends that, in resource-limited settings, ART should be taken at the advanced stage of HIV (World Health Organization, 2003b), the effect of ART can be considered to lengthen an HIV patient's life from 10 to 13 years, an 30 percent increase. This transforms to a decrease of $\overline{\lambda}_A$, the death arrival rate of an HIV patient, by about 23 percent. I also assume that η, the productivity drop due to the HIV infection, decreases by the same percentage as the death rate drop, because of the improvement of life quality due to the treatment. In addition, ART is found to be an effective way to prevent HIV (i.e., treatment as prevention), as it reduces HIV transmission rate by about 70 percent (Baeten et al., 2012). This transforms to a reduction of $\overline{\lambda_I}$ by 70 percent[23]. Since the subsidy each patient receives may not be adequate to purchase the full therapy, I make the proportion of the effectiveness of the subsidy to the effectiveness of a full therapy equal to the proportion of subsidy received per patient to the full therapy price per patient per year.

Table 5.5 shows results of the ART subsidy. It can be seen that the total number of children increases under the IU_{lc} policy, but not under other policies, since the former lowers opportunity costs of childrearing. The number of premarital children increases under all policies for all groups of women, since with the treatment not only is the cost of HIV risk lower, but the risk itself is lower. Under the IU_{lc} policy women have slightly more premarital children than under other policies for the same reason as the difference in the total number of children. As a result, HIV prevalence rates of all women decline by nearly 20 percent due to the reduction of the HIV

[23]Medical literature shows that when a HIV-positive person receives ART, the likelihood of HIV transmission to his or her sexual partner can be greatly reduced. My paper primarily focus on women whose HIV risk comes from intercourses with their infected partners. In this sense, I assumes men are equally treated as women who receive ART subsidy, hence the transmission rate to women declines. The more complicated network effect of the treatment is not taken into account here.

transmission rate, but HIV rates of young women do not decline as much as under the education-subsidy policy, since women are incentivized to have more premarital children. Internal funds with labor income tax are less effective than other sources of funding due to the substitution effect. Figure 5.5 shows the HIV prevalence by age in these experiments.

These experiments suggest that, by means of prevention, HIV treatment policy may be effective in reducing HIV rates. But its effectiveness over young women may be dampened because it encourages them to bear more premarital children.

Result 5. *An HIV treatment policy would effectively reduce the HIV prevalence rate for Kenyan women. But its effectiveness over young women would be dampened because of the encouragement for premarital fertility behavior.*

6 Conclusion

In this paper, I have explored the role of social and cultural norms regarding fertility in women's HIV risk in Sub-Saharan Africa. Using Kenyan women's data, I have documented evidence linking HIV risk with premarital fertility. I have thus constructed a lifecycle model that relates a woman's lifetime decisions concerning sex, fertility and education to HIV risk. Premarital fertility decision is the key element of the model. Premarital births increase the chance of marriage, but enhance HIV risk. I have subsequently calibrated the model based on the DHS data for Kenya and conducted counterfactual analysis regarding fertility motives and policy experiments of subsidizing income, education and HIV treatment.

Counterfactual experiments show that fertility motives play a substantial role in HIV risk for women, especially young women. Had premarital births not facilitated marriage, or women's preference for children been close to the US level, the HIV rate of young women in Kenya would have been one-third lower. Policy experiments show that the education subsidy is the most effective in reducing premarital fertility and the HIV rate of *young* women, by raising the opportunity cost of premarital childrearing. The HIV treatment subsidy is the most effective in reducing the average

HIV rate of *all* women since it effectively prevents HIV transmission. But it is not as effective in reducing *young* women's HIV rate as the education subsidy since it causes women to bear more premarital children. The income subsidy has only moderate effects on the HIV rate, since more income allows women to have more premarital children. For most types of subsidies, the external funding is the most effective, and the internal funding with labor income taxes is the least effective in reducing HIV rates. Overall, my results suggest that strong fertility motives in Africa impose considerable risk on women's sexual health, and policy makers need to take this factor into account to make HIV intervention programs more effective.

Future research may build general equilibrium effects into the fertility-HIV model. One channel can be that fertility behavior affects female labor supply and hence women's wage in the equilibrium. This in turn affects their fertility choice and hence HIV risk. Social network effects may also be on the future agenda. The current paper primarily focuses on women, but both genders engage in sexual activities and thus connect with the sexual network of both partners. Policy effects may be amplified through the behavior change of multiple individuals in the sexual network. Furthermore, the risk of infertility is an interesting issue for research, since it may incentivize women to bear children earlier over the course of their lifecycle. Finally, this paper emphasizes the importance of fertility motives to women's health, but how to reduce fertility motives, especially premarital fertility motives, remains an open issue for further studies.

Tables and Figures

Table 2.1 Summary statistics

HIV prevalence rate	all women	7.98%
	young women (age 15-24)	4.55%
	edu: primary or lower	8.54%
	edu: secondary or higher	6.91%
% of women who have premarital births		24.60%
% of women who have premarital pregnancy		47.40%
average number of children	edu: primary or lower	5.38
	edu: secondary or higher	3.35
average number of premarital children	edu: primary or lower	1.19
	edu: secondary or higher	0.74
average first marriage age	edu: primary or lower	19.26
	edu: secondary or higher	22.17
% of women who have secondary/higher education		33.63%
average years of schooling		7.70

Notes: Data are from the DHS for Kenya (2008). This table deserves more explanations. First, HIV prevalence rates listed are for all women aged 15-49, but women who have premarital sex are more relevant to the context of this paper (the fraction of this group is about 87 percent of all women). Their HIV rates are higher than the average rates of all women. For these women, HIV rates are 9.45% for all of them, 6.96% for young women, 10.15% for the less-educated, and 8.17% for the better-educated. Second, the number of children and premarital children, and the age at the first marriage the average of women aged 35-40 who had premarital sex from the KDHS (2008). The reason for choosing this age cohort is that for younger women, they may continue to have more children and will be married in the future if not yet, so including them does not well capture women's life profile of marriage and fertility; for older women, their marriage and fertility may be very different from a typical woman on average due to cohort effects. Third, education level and years of schooling are for women aged 25 or older, since younger women may not have finished education yet. Finally, all statistics are sample-weighted using weights provided in the DHS. Most of the variable values are used as targets in the calibration section.

The number of premarital children is computed in the following way. The DHS contains information of women's age at the first union and at the first child birth, and the current age of all living children. Using this data I compute women's age at the birth of each child, and then the number of (living) children born before their first marriage. Note that here birth ages are transformed to pregnancy ages.

Table 2.2. Premarital births and HIV infection

VARIABLES	(1) hiv	(2) hiv	(3) hiv	(4) hiv	(5) hiv	(6) hiv
prembth	0.030***	0.029***	0.026*	0.031***	0.047***	0.057***
	(0.132)	(0.133)	(0.197)	(0.133)	(0.174)	(0.254)
edu: primary	-0.009	-0.008	-0.056**	-0.015	-0.025	-0.072**
	(0.235)	(0.237)	(0.337)	(0.239)	(0.344)	(0.450)
edu: secondary	-0.020	-0.022	-0.064**	-0.028	-0.046	-0.097**
	(0.263)	(0.267)	(0.383)	(0.269)	(0.386)	(0.522)
edu: higher	-0.046*	-0.040	-0.076**	-0.054**	-0.049	-0.051
	(0.340)	(0.354)	(0.484)	(0.348)	(0.462)	(0.622)
urban	-0.053***	-0.047***	-0.068***	-0.046***	-0.052***	-0.086***
	(0.144)	(0.151)	(0.217)	(0.184)	(0.244)	(0.336)
ln_NoPartners			0.061***			0.053***
			(0.147)			(0.206)
occupation	NO	YES	YES	NO	NO	YES
wealthID				YES	YES	YES
No_wife & wife_rank				NO	YES	YES
region, year, cohort, age, age^2, ethnicity, religion, age at 1st union, knowledge	YES	YES	YES	YES	YES	YES
Observations	3,696	3,696	2,110	3,696	2,783	1,664

Notes: Results are from the DHS for Kenyan women (2003, 2008). The logistic model is used for regressions. Coefficients displayed are transformed to marginal effects computed at the sample mean. The dependent variable is a woman's HIV test result (1 if positive, and 0 if negative). *Prembth* is computed using a woman's age at the first birth of her child and her age at the first union (1 if a woman's first child was born before her first marriage, and 0 otherwise). *Edu* is a dummy of a woman's education level. *urban* is a dummy of whether a woman lives in an urban area. $Ln\_NoPartners$ is a woman's number of lifetime partners in logs. *Occupation* is a dummy of standardized respondent's occupation groups. $WealthID$ is the wealth index of a woman's household ranging from 1 to 5 (poorest to richest). No_wife is the number of wives of a woman's husband, and $wife\_rank$ is the woman's rank among her husband's wives. *Region* is the region where a woman lives (dummy). $Year$ is the year of the survey (dummy). *Cohort* is a dummy of a woman's birth year. *Age* is a woman's age the time of survey. *Ethnicity* and *Religion* are dummies of a woman's ethnic and religious groups. *Age at 1st union* is a woman's age at her first union (marriage). $Knowledge$ is a woman's knowledge about contraception methods and HIV prevention methods (dummy).

***Significant at the 1 percent level. **Significant at the 5 percent level. *Significant at the 10 percent level.

Table 2.3. Subsample 1 – premarital sex

VARIABLES	(1) hiv	(2) hiv	(3) hiv	(4) hiv	(5) hiv	(6) hiv
prembth	0.032**	0.030**	0.032*	0.032**	0.052***	0.068***
	(0.137)	(0.138)	(0.200)	(0.138)	(0.180)	(0.259)
edu: primary	-0.027	-0.026	-0.081**	-0.032	-0.056	-0.109**
	(0.270)	(0.271)	(0.378)	(0.272)	(0.382)	(0.501)
edu: secondary	-0.043	-0.044	-0.094**	-0.049*	-0.088**	-0.144***
	(0.295)	(0.299)	(0.418)	(0.301)	(0.424)	(0.572)
edu: higher	-0.088**	-0.077**	-0.118**	-0.093***	-0.100**	-0.109*
	(0.376)	(0.390)	(0.522)	(0.385)	(0.503)	(0.677)
urban	-0.068***	-0.062***	-0.082***	-0.065***	-0.068***	-0.104***
	(0.154)	(0.161)	(0.227)	(0.198)	(0.259)	(0.353)
ln_NoPartners			0.069***			0.053***
			(0.155)			(0.217)
occupation	NO	YES	YES	NO	NO	YES
wealthID				YES	YES	YES
No_wife & wife_rank				NO	YES	YES
region, year, cohort, age, age^2, ethnicity, religion, age at 1st union, knowledge	YES	YES	YES	YES	YES	YES
Observations	3,696	3,696	2,110	3,696	2,783	1,664

Notes: Results are from the DHS for Kenyan women (2003, 2008). I restrict the sample for regressions to women who had premarital sex. The logistic model is used for regressions. Coefficients displayed are transformed to marginal effects computed at the sample mean. The dependent variable is a woman's HIV test result (1 if positive, and 0 if negative). All the independent variables are the same as in Table 2.2.

***Significant at the 1 percent level. **Significant at the 5 percent level. *Significant at the 10 percent level.

Table 2.4. Subsample 2 – non-commercial sex

VARIABLES	(1) hiv	(2) hiv	(3) hiv	(4) hiv	(5) hiv	(6) hiv
prembth	0.030***	0.029***	0.027*	0.031***	0.047***	0.057***
	(0.132)	(0.133)	(0.197)	(0.133)	(0.174)	(0.254)
edu: primary	-0.001	-0.001	-0.056**	-0.001	-0.002	-0.072**
	(0.236)	(0.237)	(0.338)	(0.239)	(0.344)	(0.450)
edu: secondary	-0.002	-0.002	-0.064**	-0.002	-0.004	-0.097**
	(0.264)	(0.267)	(0.384)	(0.270)	(0.386)	(0.522)
edu: higher	-0.046*	-0.003	-0.074*	-0.054**	-0.004	-0.004
	(0.340)	(0.355)	(0.485)	(0.348)	(0.462)	(0.622)
urban	-0.053***	-0.046***	-0.068***	-0.046***	-0.052***	-0.086***
	(0.145)	(0.152)	(0.218)	(0.185)	(0.244)	(0.336)
ln_NoPartners			0.060***			0.053***
			(0.150)			(0.206)
occupation	NO	YES	YES	NO	NO	YES
wealthID				YES	YES	YES
No_wife & wife_rank				NO	YES	YES
region, year, cohort, age, age^2, ethnicity, religion, age at 1st union, knowledge	YES	YES	YES	YES	YES	YES
Observations	3,693	3,693	2,108	3,693	2,783	1,664

Notes: Results are from the DHS for Kenyan women (2003, 2008). I restrict the sample for regressions to women who are not involved in commercial sex. The logistic model is used for regressions. Coefficients displayed are transformed to marginal effects computed at the sample mean. The dependent variable is a woman's HIV test result (1 if positive, and 0 if negative). All the independent variables are the same as in Table 2.2.

***Significant at the 1 percent level. **Significant at the 5 percent level. *Significant at the 10 percent level.

Table 2.5. Education, wealth, and premarital births

VARIABLES	(1) prembth	(2) prembth	(3) prembth	(4) prembth	(5) prembth	(6) prembth	(7) prembth	(8) prembth	(9) prembth	(10) prembth	(11) prembth	(12) prembth
edu: primary	0.015	0.015	0.017*	0.020**	0.025**	0.022						
	(0.127)	(0.127)	(0.136)	(0.128)	(0.160)	(0.171)						
edu: secondary	-0.037***	-0.036***	-0.015	-0.027**	-0.023	-0.008						
	(0.140)	(0.141)	(0.150)	(0.143)	(0.176)	(0.190)						
edu: higher	-0.153***	-0.150***	-0.118***	-0.138***	-0.125***	-0.098***						
	(0.186)	(0.194)	(0.200)	(0.189)	(0.224)	(0.250)						
ln_eduyear							-0.075***	-0.067***	-0.050***	-0.066***	-0.068***	-0.040***
							(0.0786)	(0.0813)	(0.0869)	(0.0814)	(0.0975)	(0.110)
urban	0.023***	0.021***	-0.153***	-0.002	0.000	0.005	0.032***	0.027***	0.035***	0.001	0.002	0.008
	(0.0811)	(0.0841)	(0.0866)	(0.102)	(0.121)	(0.130)	(0.0848)	(0.0884)	(0.0919)	(0.106)	(0.125)	(0.136)
2.wealthID				-0.014*	-0.006	-0.006				-0.008	0.000	0.002
				(0.103)	(0.118)	(0.125)				(0.112)	(0.127)	(0.137)
3.wealthID				-0.020**	-0.019**	-0.017*				-0.017*	-0.016	-0.013
				(0.104)	(0.119)	(0.127)				(0.114)	(0.130)	(0.138)
4.wealthID				-0.026***	-0.028***	-0.020*				-0.027***	-0.028***	-0.016
				(0.107)	(0.124)	(0.132)				(0.116)	(0.133)	(0.143)
5.wealthID				-0.057***	-0.064***	-0.054***				-0.064***	-0.073***	-0.055***
				(0.135)	(0.160)	(0.172)				(0.143)	(0.168)	(0.183)
occupation	NO	YES	NO	NO	NO	YES	NO	YES	NO	NO	NO	YES
sexage	NO	NO	YES	NO	NO	YES	NO	NO	YES	NO	NO	YES
No_wife & wife_rank	NO	NO	NO	NO	YES	YES	NO	NO	NO	NO	YES	YES
region, year, cohort, age, age^2, ethnicity, religion, age at 1st union, knowledge	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	8,611	8,611	8,563	8,611	6,656	6,613	7,132	7,132	7,086	7,132	5,570	5,527

Notes: Results are from the DHS for Kenyan women (1989, 1993, 1998, 2003, 2008). The logistic model is used for regressions. The dependent variable is a woman's premarital birth (1 if she has a premarital birth, and 0 if not). All the independent variables are the same as in Table 2.2, except for *ln_eduyear*, which is a woman's years of schooling in logs, and *sexage*, which is a dummy of a woman's age at her first sex.

***Significant at the 1 percent level. **Significant at the 5 percent level. *Significant at the 10 percent level.

Table 4.1. Parameter values, lifecycle model

Tastes	$\theta_H = 1.34$, $\theta_L = 0.87$, $\rho = 0.05$, $\delta = 0.25$
Childrearing cost	$\phi_H = 0.25$, $\phi_L = 0.16$
Human capital accumulation	$\Phi_H = 0.16$, $\Phi_L = 0.04$, $\gamma = 0.72$
Transfers from the partner	$\varphi_0 = 0.13$, $\varphi_M = 0.60$, $x_0 = 0.60$
HIV productivity drop	$\eta = 0.05$
Arrival rates	$\overline{\lambda}_M = 0.11$, $\alpha = 0.35$, $\zeta = 0.74$
	$\overline{\lambda}_I = 0.0072$, $\beta = 3.79$, $\kappa = 0.077$
	$\overline{\lambda}_D = 0.023$, $\overline{\lambda}_A = 0.0985$

Notes: This table reports the calibrated parameter values of the lifecycle model.

Table 4.2. Calibration results, model and targeted values

		Model	Target
Total number of children	n_H	3.20	3.35
	n_L	5.37	5.38
Number of premarital children	n_{1H}	0.79	0.74
	n_{1L}	1.18	1.19
Marriage age	$Mage_H$	22.24	22.17
	$Mage_L$	19.59	19.26
HIV prevalence rate	$HIVpr_H$	7.94%	8.04%
	$HIVpr_L$	8.04%	11.31%

Notes: This table reports the values of eight variables from the calibrated model and the target. The targeted values are based on the DHS (2008) for Kenyan women.

Table 5.1. Counterfactual analysis: fertility motives

	BM	$\alpha = 0$	$\theta = 0.55$	$\delta = 0.87$
n	4.64	4.62	2.73	4.62
n_1	1.05	0.73	0.60	0.68
n_{1H}	0.79	0.23	0.00	0.01
n_{1L}	1.18	0.98	0.90	1.02
$HIVpr$	8.01%	6.50%	6.65%	6.49%
$HIVpr_{young}$	6.56%	4.35%	4.47%	4.40%
$HIVpr_H$	7.94%	5.79%	5.79%	5.80%
$HIVpr_L$	8.04%	6.85%	7.10%	6.85%
s_{change}			$\Phi_L \to 0$	$\Phi_L \to 1$

Notes: This table reports the results of the counterfactual analysis regarding fertility motives. n is the total number of children per woman. n_1 is the number of premarital children per woman, n_{1H} (n_{1L}) is the number of premarital children per more-educated (less-educated) woman. $HIVpr$ is the HIV prevalence rate of all women (aged 15-49). $HIVpr_{young}$ is the HIV prevalence rate of young women (aged 15-24). $HIVpr_H$ ($HIVpr_L$) is the HIV prevalence rate of more-educated (less-educated) women. s_{change} is the change of sex type, where $\Phi_i \to s$ means more women of education group i choose sex type s. Column one shows the variable values from the benchmark model, and column two to four show the results of the four counterfactual experiments.

Table 5.2. Counterfactual analysis: behavior change

	BM	YES	NO
n	4.64	4.62	4.64
n_1	1.05	1.40	1.05
n_{1H}	0.79	1.03	0.79
n_{1L}	1.18	1.58	1.18
$HIVpr$	8.01%	4.03%	2.71%
$HIVpr_{young}$	6.56%	4.27%	2.23%
$HIVpr_H$	7.94%	3.85%	2.70%
$HIVpr_L$	8.04%	4.13%	2.72%

Notes: This table reports the results of counterfactual analysis regarding behavior change. $\overline{\lambda}_I$ is set to be one-third of its original value (i.e., $\overline{\lambda}_I = 0.0024$), and its effect on HIV rates is examined when behavior response is allowed or not. The variables reported in this table are the same as in Table 5.1. Column one (BM) shows variable values from the benchmark model. Column two shows the results when behavior change is allowed, and column three shows the results when behavior is fixed to be the same as in the benchmark model.

Table 5.3. Policy experiment, income subsidy

	BM	IL_{lc}	IL_{ls}	EL	IU_{lc}	EU
n	4.64	4.79	4.64	4.71	4.81	4.71
n_1	1.05	1.06	1.05	1.06	1.06	1.05
n_{1H}	0.79	0.81	0.79	0.79	0.81	0.80
n_{1L}	1.18	1.19	1.18	1.19	1.19	1.18
$HIVpr$	8.01%	8.11%	8.01%	7.75%	8.11%	7.68%
$HIVpr_{young}$	6.56%	6.72%	6.56%	6.28%	6.72%	6.18%
$HIVpr_H$	7.94%	8.07%	7.94%	7.94%	8.07%	6.97%
$HIVpr_L$	8.04%	8.13%	8.04%	7.66%	8.13%	8.04%
s_{change}				$\Phi_L \to 1$		$\Phi_H \to 1$

Notes: This table reports the results of the income subsidy policy. The variables reported in this table are the same as in Table 5.1. Column one (BM) shows the variable values from the benchmark model. Column two (IL_{lc}) shows the results of the income subsidy funded by the internal funding (labor income tax) and subsidizing only less-educated women. Column three (IL_{ls}) shows the results of the income subsidy funded by the internal funding (lump-sum tax) and subsidizing only less-educated women. Column four (EL) shows the results of the subsidy funded by the external funding and subsidizing less-educated women. Column five (IU_{lc}) shows the results of the income subsidy funded by the internal funding (labor income tax) and subsidizing all women uniformly. Column six (EU) shows the results of the subsidy funded by the external funding and subsidizing all women.

Table 5.4. Policy experiment, education subsidy

	BM	IU_{lc}	IU_{ls}	EU
n	4.64	4.72	4.58	4.62
n_1	1.05	0.92	0.92	0.91
n_{1H}	0.79	0.37	0.36	0.37
n_{1L}	1.18	1.19	1.19	1.18
$HIVpr$	8.01%	7.37%	7.37%	7.32%
$HIVpr_{young}$	6.56%	5.68%	5.67%	5.59%
$HIVpr_H$	7.94%	5.91%	5.89%	5.91%
$HIVpr_L$	8.04%	8.13%	8.13%	8.04%

Notes: This table reports the results of the education subsidy policy. The variables reported in this table are the same as in Table 5.1. Column one (BM) shows the variables values of the benchmark model. Column two (IU_{lc}) shows the results of the education subsidy funded by the internal funding (labor income tax) and subsidizing all women. Column three (IU_{ls}) shows the results of the education subsidy funded by the internal funding (lump-sum tax) and subsidizing all women. Column four (EU) shows the results of the education subsidy funded by external funding and subsidizing all women.

Table 5.5. Policy experiment, HIV treatment subsidy

	BM	IU_{lc}	IU_{ls}	EU
n	4.64	4.74	4.58	4.62
n_1	1.05	1.28	1.28	1.28
n_{1H}	0.79	0.96	0.94	0.95
n_{1L}	1.18	1.45	1.45	1.45
$HIVpr$	8.01%	6.60%	6.49%	6.49%
$HIVpr_{young}$	6.56%	5.98%	5.85%	5.85%
$HIVpr_H$	7.94%	6.42%	6.23%	6.23%
$HIVpr_L$	8.04%	6.69%	6.62%	6.62%

Notes: This table reports the results of the HIV treatment subsidy policy. The variables reported in this table are the same as in Table 5.1. Column one (BM) shows the variables values of the benchmark model. Column two (IU_{lc}) shows the results of the treatment subsidy funded by the internal funding (labor income tax) and subsidizing all HIV-infected women. Column three (IU_{ls}) shows the results of the treatment subsidy funded by the internal funding (lump-sum tax) and subsidizing all HIV-infected women.Column four (EU) shows the results of the treatment subsidy funded by external funding and subsidizing all HIV-infected women.

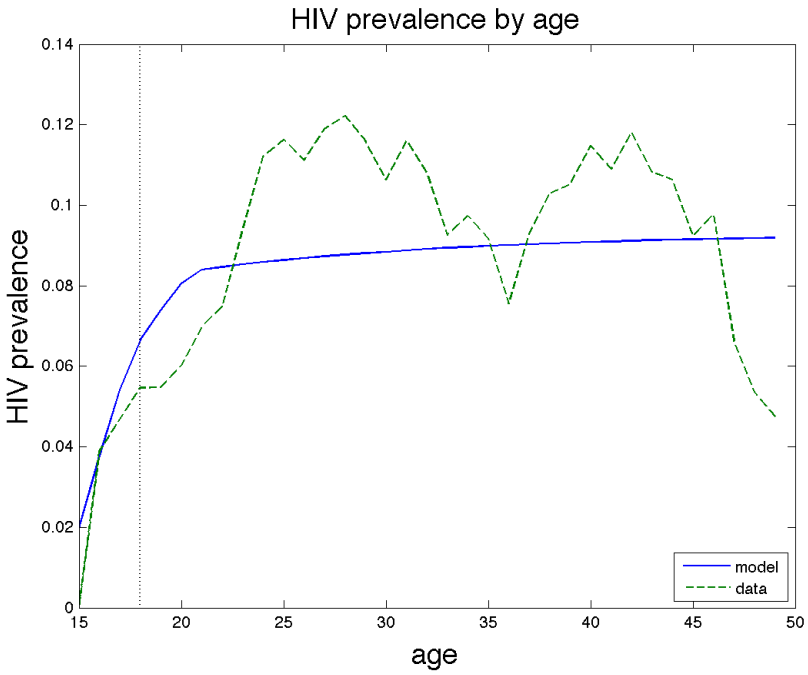

Figure 4.1. HIV prevalence by age: data and model

Notes: This figures shows the HIV prevalence rate of all Kenyan women by age (15-49) from the model and the data. The data are from the DHS (2008) for Kenyan women.

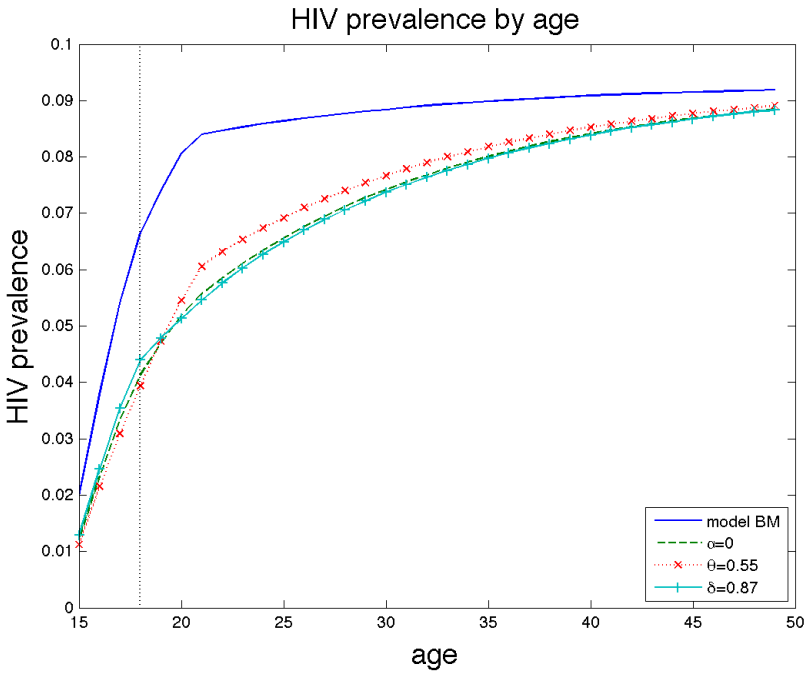

Figure 5.1. HIV prevalence, counterfactual experiments: fertility motives

Notes: This figure shows the HIV prevalence rate of all Kenyan women by age (15-49) from the benchmark model and the three counterfactual experiments regarding fertility motives.

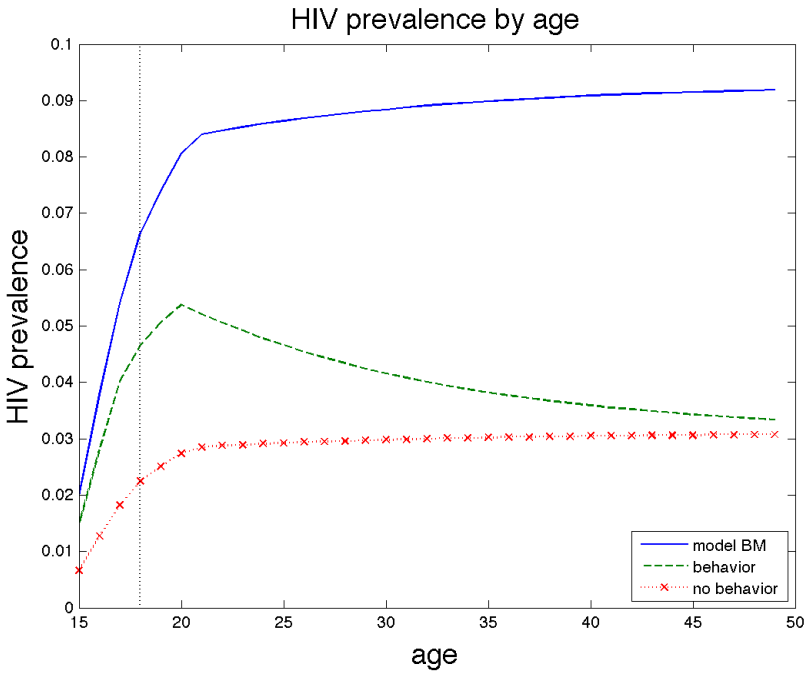

Figure 5.2. HIV prevalence, counterfactual experiments: behavior change

Notes: This figure shows the HIV prevalence rate of all Kenyan women by age (15-49) from the benchmark model and the counterfactual experiments regarding behavior change. $\overline{\lambda}_I$ is set to be one-third of its original value (i.e., $\overline{\lambda}_I = 0.0024$), and its effect on HIV rates is examined when behavior response is allowed or not. "Behavior" means behavior change is allowed, and "no behavior" means behavior is fixed to be the same as in the benchmark model.

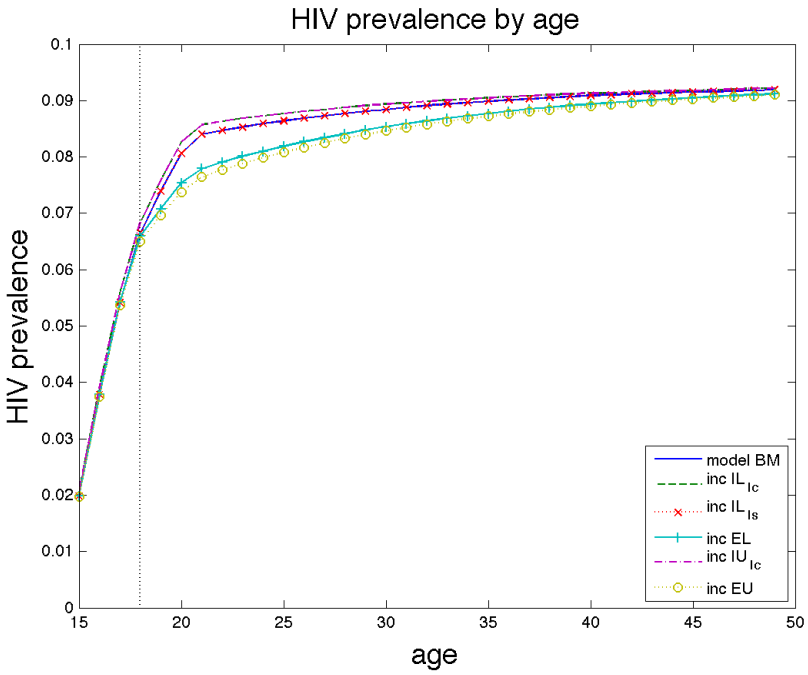

Figure 5.3. HIV prevalence, income subsidy

Notes: This figure shows the HIV prevalence rate of all women by age (15-49) from the benchmark model and four income subsidy policy experiments. "inc IL" is the result of the income subsidy funded by the internal funding and subsidized only to less-educated women. "inc IU" is the result of the income subsidy funded by the internal funding and subsidized uniformly to all women. "inc EL" is the result of the subsidy funded by the external funding and subsidizing less-educated women. "inc EU" is the result of the subsidy funded by the external funding and subsidizing all women. The subscript "lc" means internal funding is raised from labor income tax, and the subscript "ls" means internal funding is raised from lump-sum tax.

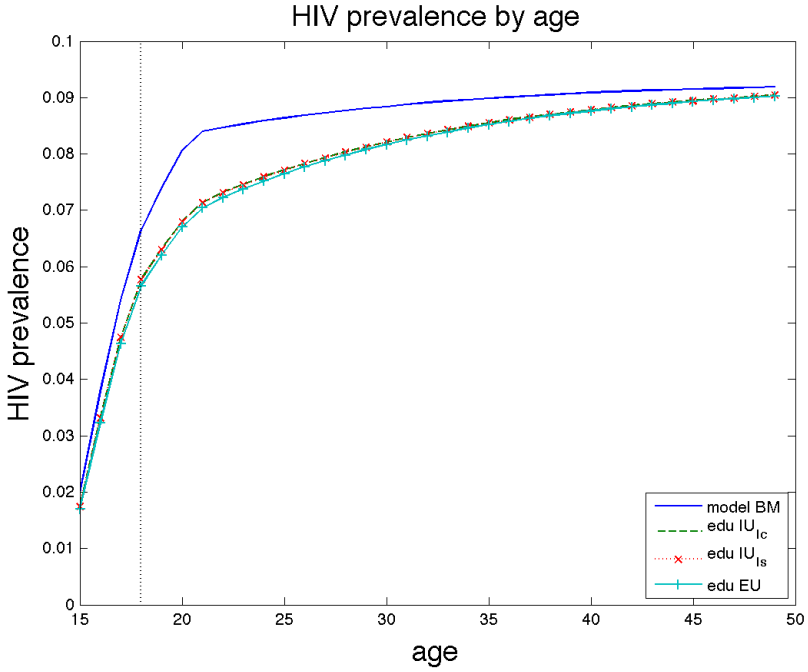

Figure 5.4. HIV prevalence, education subsidy

Notes: This figure shows the HIV prevalence of all Kenyan women by age (15-49) from the benchmark model and two education subsidy policy experiments. "edu IU" is the result of the education subsidy funded by the internal funding and subsidized to all women. "edu EU" is the result of the education subsidy funded by external funding and subsidized to all women. The subscript "lc" means internal funding is raised from labor income tax, and the subscript "ls" means internal funding is raised from lump-sum tax.

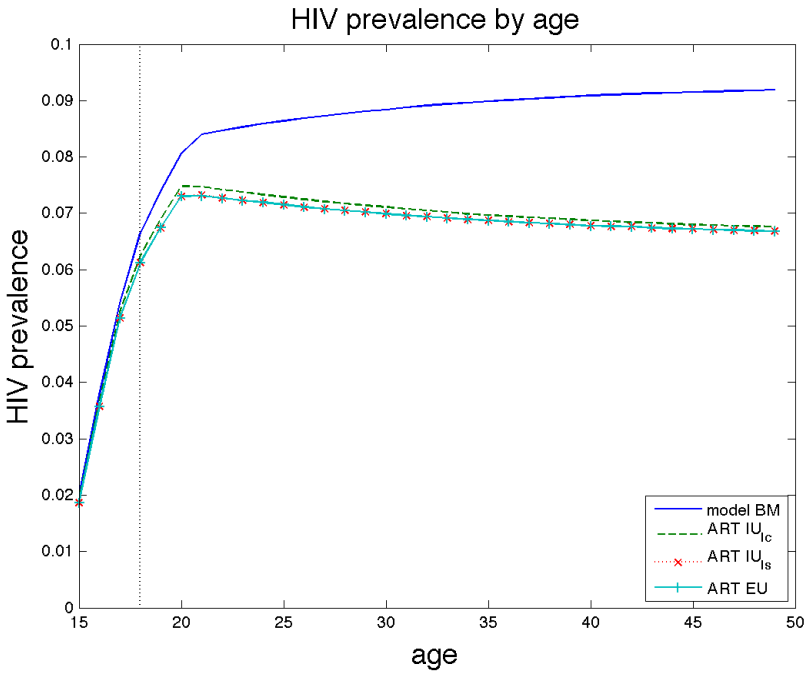

Figure 5.5. HIV prevalence, treatment subsidy

Notes: This figure shows the HIV prevalence of all Kenyan women by age (15-49) from the benchmark model and two HIV treatment subsidy policy experiments. "ART IU" is the result of the education subsidy funded by the internal funding and subsidized to all HIV-infected women. "ART EU" is the result of the education subsidy funded by external funding and subsidized to all HIV-infected women. The subscript "lc" means internal funding is raised from labor income tax, and the subscript "ls" means internal funding is raised from lump-sum tax.

References

[1] Alsan, Marcella M., and David M. Cutler. "Girls' education and HIV risk: Evidence from Uganda." Journal of health economics 32, no. 5 (2013): 863-872.

[2] Baeten, Jared M., Deborah Donnell, Patrick Ndase, Nelly R. Mugo, James D. Campbell, Jonathan Wangisi, Jordan W. Tappero et al. "Antiretroviral prophylaxis for HIV prevention in heterosexual men and women." New England Journal of Medicine 367, no. 5 (2012): 399-410.

[3] Barro, Robert J., and Gary S. Becker. "Fertility choice in a model of economic growth." Econometrica: journal of the Econometric Society (1989): 481-501.

[4] Becker, Gary S. "An economic analysis of fertility." In Demographic and economic change in developed countries, pp. 209-240. Columbia University Press, 1960.

[5] Becker, Gary, Kevin Murphy, and Robert Tamura. "Economic growth, human capital and population growth." Journal of Political Economy 98, no. 5 (1990): S12-S137.

[6] Bloemen, Hans, and Adriaan S. Kalwij. "Female labor market transitions and the timing of births: a simultaneous analysis of the effects of schooling." Labour Economics 8, no. 5 (2001): 593-620.

[7] Case, Anne, and Christina Paxson. "HIV Risk and Adolescent Behaviors in Africa." The American Economic Review 103, no. 3 (2013): 433-438.

[8] Cates, W., Timothy MM Farley, and Patrick J. Rowe. "Worldwide patterns of infertility: is Africa different?." The Lancet 326, no. 8455 (1985): 596-598.

[9] Caucutt, Elizabeth M., Nezih Guner, and John Knowles. "The timing of births: A marriage market analysis." (2001).

[10] Cigno, Alessandro, and John Ermisch. "A microeconomic analysis of the timing of births." European Economic Review 33, no. 4 (1989): 737-760.

[11] Conesa, Juan Carlos. "Educational attainment and timing of fertility decisions." Documents de treball (Facultat d'Economia i Empresa. Espai de Recerca en Economia), 2002, E02/78 (2002).

[12] Cooper, Diane, Jane Harries, Landon Myer, Phyllis Orner, and Hillary Bracken. ""Life is still going on": Reproductive intentions among HIV-positive women and men in South Africa." Social science & medicine 65, no. 2 (2007): 274-283.

[13] Corno, Lucia, and Damien De Walque. "Mines, migration and HIV/AIDS in Southern Africa." Journal of African economies (2012): ejs005.

[14] De Walque, Damien. "How does the impact of an HIV/AIDS information campaign vary with educational attainment? Evidence from rural Uganda." Journal of Development Economics 84, no. 2 (2007): 686-714.

[15] De Walque, Damien. "Does education affect HIV status? Evidence from five African countries." The World Bank Economic Review 23, no. 2 (2009): 209-233.

[16] Duflo, Esther, Pascaline Dupas, and Michael Kremer. "Education, HIV, and early fertility: Experimental evidence from Kenya." American Economic Review 105, no. 9 (2015): 2757-97.

[17] Easterlin, Richard A. "Population, labor force, and long swings in economic growth: The American experience." NBER Books (1968).

[18] Epstein, Helen, and Martina Morris. "Concurrent partnerships and HIV: an inconvenient truth." Journal of the International AIDS Society 14, no. 1 (2011): 13.

[19] Epstein, Helen. The invisible cure: why we are losing the fight against AIDS in Africa. Macmillan, 2008.

[20] Fortson, Jane G. "The gradient in sub-Saharan Africa: socioeconomic status and HIV/AIDS." Demography 45, no. 2 (2008): 303-322.

[21] Fortson, Jane G. "HIV/AIDS and fertility." American Economic Journal: Applied Economics 1, no. 3 (2009): 170-194.

[22] Garenne, Michel, and Julien Zwang. "Premarital fertility and ethnicity in Africa." (2006).

[23] Greenwood, Jeremy, Philipp Kircher, Cezar Santos, and Michèle Tertilt. "An Equilibrium Model of the African HIV/AIDS Epidemic." No. w18953. National Bureau of Economic Research, 2013.

[24] Godlonton, Susan, Alister Munthali, and Rebecca Thornton. "Responding to risk: Circumcision, information, and HIV prevention." Review of Economics and Statistics 0 (2014).

[25] Gutiérrez-Domènech, Maria. "The impact of the labour market on the timing of marriage and births in Spain." Journal of Population Economics 21, no. 1 (2008): 83-110.

[26] Happel, Stephen K., Jane K. Hill, and S. A. Low. "An economic analysis of the timing of childbirth." Population Studies 38, no. 2 (1984): 299-311.

[27] Heckman, James J., and James R. Walker. "The relationship between wages and income and the timing and spacing of births: evidence from Swedish longitudinal data." Econometrica: journal of the Econometric Society (1990a): 1411-1441.

[28] Heckman, James J., and James R. Walker. "The third birth in Sweden." Journal of Population Economics 3, no. 4 (1990b): 235-275.

[29] Hoyne, Hilary Williamson. "Does welfare play any role in female headship decisions?." Journal of Public Economics 65, no. 2 (1997): 89-117.

[30] Jones, Larry E., and Michèle Tertilt. "Chapter 5 An Economic History of Fertility in the United States: 1826–1960." In Frontiers of family economics, edited by Peter Rupert, pp. 165-230. Emerald Group Publishing Limited, 2008.

[31] Jordan, Rachel, Lisa Gold, Carole Cummins, and Chris Hyde. "Systematic review and meta-analysis of evidence for increasing numbers of drugs in antiretroviral combination therapy." Bmj 324, no. 7340 (2002): 757.

[32] Juhn, Chinhui, Sebnem Kalemli-Ozcan, and Belgi Turan. "HIV and fertility in Africa: first evidence from population-based surveys." Journal of Population Economics 26, no. 3 (2013): 835-853.

[33] Kalemli-Ozcan, Sebnem. "AIDS,"reversal" of the demographic transition and economic development: evidence from Africa." Journal of Population Economics 25, no. 3 (2012): 871-897.

[34] Kenya Revenue Authority, Domestic Taxes Department. Employer's guide to pay as you earn in Kenya. Revised edition, 2007.

[35] Kremer, Michael. "Integrating behavioral choice into epidemiological models of the AIDS epidemic." No. w5428. National Bureau of Economic Research, 1996.

[36] Iorio, Daniela, and Raul Santaeulalia-Llopis. "Education, HIV Status, and Risky Sexual Behavior: How Much Does the Stage of the HIV Epidemic Matter?." Unpublished Manuscript, Washington University (2011).

[37] Iyigun, Murat F. "Timing of childbearing and economic growth." Journal of Development Economics 61, no. 1 (2000): 255-269.

[38] Larsen, Ulla. "Primary and secondary infertility in sub-Saharan Africa." International Journal of Epidemiology 29, no. 2 (2000): 285-291.

[39] Leenstra, Tjalling, L. T. Petersen, S. K. Kariuki, A. J. Oloo, P. A. Kager, and F. O. Ter Kuile. "Prevalence and severity of malnutrition and age at menarche; cross-sectional studies in adolescent schoolgirls in western Kenya." European journal of clinical nutrition 59, no. 1 (2005): 41-48.

[40] Magruder, Jeremy R. "Marital shopping and epidemic AIDS." Demography 48, no. 4 (2011): 1401-1428.

[41] Manuelli, Rodolfo E. "AIDS, human capital and development." Working paper.

[42] Meekers, Dominique, and Anne-Emmanuèle Calvès. "'Main'girlfriends, girlfriends, marriage, and money: the social context of HIV risk behaviour in sub-Saharan Africa." Health Transition Review (1997): 361-375.

[43] Morgan, Dilys, Cedric Mahe, Billy Mayanja, J. Martin Okongo, Rosemary Lubega, and James AG Whitworth. "HIV-1 infection in rural Africa: is there a difference in median time to AIDS and survival compared with that in industrialized countries?." Aids 16, no. 4 (2002): 597-603.

[44] Mullin, Charles H., and Ping Wang. "The timing of childbearing among heterogeneous women in dynamic general equilibrium." No. w9231. National Bureau of Economic Research, 2002.

[45] Murdock, George Peter. "Ethnographic atlas: a summary." Ethnology (1967): 109-236.

[46] Myer, Landon, Chelsea Morroni, and Kevin Rebe. "Prevalence and determinants of fertility intentions of HIV-infected women and men receiving antiretroviral therapy in South Africa." AIDS patient care and STDs 21, no. 4 (2007): 278-285.

[47] Nakayiwa, Sylvia, Betty Abang, Laura Packel, Julie Lifshay, David W. Purcell, Rachel King, Enoch Ezati, Jonathan Mermin, Alex Coutinho, and Rebecca Bunnell. "Desire for children and pregnancy risk behavior among HIV-infected men and women in Uganda." AIDS and Behavior 10, no. 1 (2006): 95-104.

[48] Nattabi, Barbara, Jianghong Li, Sandra C. Thompson, Christopher Garimoi Orach, and Jaya Earnest. "A systematic review of factors influencing fertility desires and intentions among people living with HIV/AIDS: implications for policy and service delivery." AIDS and Behavior 13, no. 5 (2009): 949-968.

[49] Oster, Emily. "HIV and sexual behavior change: Why not Africa?." Journal of Health Economics 31, no. 1 (2012): 35-49.

[50] Oster, Emily. "Sexually transmitted infections, sexual behavior, and the HIV/AIDS epidemic." The Quarterly Journal of Economics 120, no. 2 (2005): 467-515.

[51] Oster, Emily. "Routes of Infection: Exports and HIV Incidence in Sub-Saharan Africa." Journal of the European Economic Association 10, no. 5 (2012): 1025-1058.

[52] Paiva, Vera, Elvira Ventura Filipe, Naila Santos, Tiago Novaes Lima, and Aluisio Segurado. "The right to love: the desire for parenthood among men living with HIV." Reproductive health matters 11, no. 22 (2003): 91-100.

[53] Peltzer, Karl, Li-Wei Chao, and Pelisa Dana. "Family planning among HIV positive and negative prevention of mother to child transmission (PMTCT) clients in a resource poor setting in South Africa." AIDS and Behavior 13, no. 5 (2009): 973-979.

[54] Pongou, Roland, and Roberto Serrano. "Fidelity networks and long-run trends in HIV/AIDS gender gaps." The American Economic Review 103, no. 3 (2013): 298-302.

[55] Sembuya, Rita. "Mother or nothing: the agony of infertility." Bull World Health Organ 88 (2010): 881-882.

[56] Smith, Daniel Jordan, and Benjamin C. Mbakwem. "Life projects and therapeutic itineraries: marriage, fertility, and antiretroviral therapy in Nigeria." AIDS 21 (2007): S37-S41.

[57] Thornton, Rebecca L. "The demand for, and impact of, learning HIV status." The American economic review 98, no. 5 (2008): 1829.

[58] Tsay, Wen-Jen, and CY Cyrus Chu. "The pattern of birth spacing during Taiwan's demographic transition." Journal of Population Economics 18, no. 2 (2005): 323-336.

[59] UNAIDS. UNAIDS GAP Report (2014).

[60] Wang, Ping, Chong K. Yip, and Carol A. Scotese. "Fertility choice and economic growth: Theory and evidence." The Review of Economics and Statistics (1994): 255-266.

[61] Winter, Carolyn. "Women Workers in South Africa: Participation, Pay and Prejudice in the Formal Labor Market." (1999).

[62] World Health Organization. "Progress report in Reproductive Health Research." no. 63 (2003a).

[63] World Health Organization. "Scaling up antiretroviral therapy in resource-limited settings: Treatment guidelines for a public health approach." (2003b).

[64] Young, Alwyn. "The gift of the dying: the tragedy of AIDS and the welfare of future African generations." Quarterly Journal of Economics 120, no(2) (2005): 423-466.

[65] Young, Alwyn. "In sorrow to bring forth children: fertility amidst the plague of HIV." Journal of economic growth 12, no. 4 (2007): 283-327.

Appendix

This appendix provides more details about the Evidence section.

Data is taken from Kenyan DHS surveys, which cover five years: 1989, 1993, 1998, 2003, 2008. The last two surveys contain results of HIV tests of a randomly selected subsample of the normal surveys. For regressions I use data from 2003 and 2008 surveys if HIV test results are needed and all years' data otherwise. Below I explain the regression model and data.

Premarital births and HIV

The hypothesis is that premarital births increase HIV infection probability by increasing unprotected premarital sexual activity. I use the logistic probability model for regression.

$$HIV_i = \beta_0 + \beta_1 prembth_i + \beta_2 X_i + \varepsilon_i$$

where the dependent variable HIV_i is a dummy which equals one if a woman i's HIV test result is positive, and zero if it is negative. $Prembth_i$ is a dummy computed by the author which equals one if a woman i's first child was born before her first marriage, and zero otherwise. X_i is a set of control variables including women's age and cohort, living area and region, ethnicity, religion, age at the first marriage, education, knowledge about contraception and HIV prevention, number of partners, occupation and wealth. A detailed description of variables is as follows.

HIV: is the dependent variable. It is the HIV test result of a randomly selected subsample of the normal sample of the DHS. The corresponding DHS variable is $hiv03$.

$Prembth$: is a dummy variable which equals one if a woman's first birth was before the first union (marriage). It is computed using the DHS variable $v212$ (age at the first birth) and $v511$ (age at the first union).

Edu: is a dummy variable of a woman's education level, including no education, primary school, secondary school, and higher. I add this variable since education level may affect both premarital fertility and HIV infection through its effect on sexual

and fertility behavior. The corresponding DHS variable is $v106$ (highest education attended)

Urban: is a dummy variable that equals one if a woman lives in an urban area and zero if in she lives in a rural area. The corresponding DHS variable $v102$ (type of place of residence where the respondent was interviewed as either urban or rural).

Region: is a dummy variable of a woman's living region that includes eight provinces in Kenya. The corresponding DHS variable is $v101$ (region in which the respondent was interviewed).

I use *urban* and *region* variables because sexual and fertility behavior may vary across urban/rural areas and regions due to variations of culture, customs, ethnicity and religion.

Year: is a dummy variable of the year of the survey.

Birthyear: is a dummy variable of a woman's year of birth. This variable is used to capture women's cohort effects. The corresponding DHS variable is $v010$ (year of birth of the respondent).

Age: is the age of the respondent at the time of survey. Age^2 is also included in regressions. The corresponding DHS variable is $v012$.

I use *year*, *birthyear*, *age* and age^2 since both fertility behavior and HIV prevalence may change over the years and cohorts, and may follow a quadratic pattern over the lifecycle.

Ethnicity: is a woman's ethnicity. The corresponding DHS variable is $v131$.

Religion: is a woman's religion. The corresponding DHS variable is $v130$.

I include *ethnicity* and *religion* variables since these cultural factors have great influences on women's premarital sexual and fertility behavior and HIV status (see the background section).

Age at the first union: is a woman's age at her first union (marriage). It is included because the age at marriage may be associated with premarital fertility. A woman who is married later may have more premarital births due to the lengthened spacing between menarche and marriage, but may also have fewer premarital children since premarital births facilitate the arrival of marriage. The age at marriage may also affect HIV status as studies show that delayed marriage is linked to higher HIV

infection rate through an increase of premarital sex. The corresponding DHS variable is $v511$.

Contraception knowledge: is a dummy variable of a woman's knowledge about any contraception method classified into modern, traditional and folkloric methods. The corresponding DHS variable is $v301$.

HIV prevention knowledge: is a dummy of whether a woman knows always using condoms during sex would reduce their chance of getting AIDS. The corresponding DHS variable is $v754CP$.

I include *Contraception knowledge* and *HIV prevention knowledge* since they affect both HIV transmission and fertility outcome.

In some regressions, I also add the following control variables.

Occupation: is a dummy variable of the standardized respondent's occupation groups in ten categories. I add this variable because women's sexual behavior and fertility outcome may be influenced by their occupation. The corresponding DHS variable is $v717$.

$ln\_NoPartners$: is a woman's lifetime number of sexual partners in logs. I use this variable to control sex style of a woman (more partners implies more casual sex conditional on other variables) which is correlated with both HIV status and premarital fertility. The corresponding DHS variable is $v836$.

$WealthID$: is a dummy variable of household wealth index ranged from one to five representing wealth status from poorest to richest. The corresponding DHS variable is $hv270$.

No_wife: is a dummy variable of the number of other wives a woman's husband currently has. The corresponding DHS variable is $v505$.

$Wife\_rank$: is a dummy variable of the rank of a woman among the partner's wives. The corresponding DHS variable is $v506$.

I use $wealthID$, No_wife, $wife\_rank$ variables to proxy a woman's wealth before marriage. By assuming assortative matching, wealth status of a married woman's household implies her wealth status before marriage. But this implication may not be correct if the woman enters a polygyny, i.e., she married a rich man but her rank among the wives is low. Hence I also include the number of wives and one's

rank among the wives to better proxy her wealth status before marriage.

For subsample regressions, I use the age of the first sexual intercourse (the DHS variable $v525$) to identify premarital sex, and use the variable that asked with whom the woman had sex ($v767a$, $v767b$, $v767c$) to identify those who had commercial sex.

Education, wealth and premarital births

I again use logistic model with the following specification.

$$prembth_i = \beta_0 + \beta_1 edu_i + \beta_2 wealth_i + \beta_3 X_i + \varepsilon_i$$

where *prembth* is defined in the same way as in the previous section. For *edu* I use two measures. One is education level, and the other is the number of schooling years (in logs). For *wealth* I use wealth index, the number of wives of one's husband and her rank among the wives to proxy her wealth status before marriage. X_i represents control variables, including women's age and cohort, living area and region, ethnicity, religion, age at the first marriage, age at the first sex, knowledge about contraception and HIV prevention and occupation. A detailed description of variables is as follows.
Prembth: is the dependent variable and is defined in the same way as in the previous regression.
Edu: there are two measures of education. One is education level which is the same as in the previous regression. The other is years of schooling in logs ($ln\_eduyear$), using the DHS variable $v133$ (education in single years).
$WealthID$, No_wife, $wife\_rank$: these are used to proxy a woman's premarital wealth status.
$Urban$, $region$, $year$, $birthyear$, age, age^2, $ethnicity$ and $religion$: these variables are included since a woman's living area, age, cohort, and ethnic and religious background may affect both her education/wealth and premarital fertility behavior. The definition of these variables is the same as in the previous regression.
$Age\ at\ the\ first\ union$, $Contraception\ knowledge$ and $HIV\ prevention\ knowledge$: are included since they may be correlated to both education/wealth and premarital fertility. They are defined in the same way as in the previous regression.

Sexage: is a dummy variable of a woman's age at her first sex. Literature shows that a woman's age at her sex onset is associated with her education level, and it also affects her premarital fertility since earlier sex onset may result in earlier pregnancy. The corresponding DHS variable is $v525$ (age at first sexual intercourse).