HIV/AIDS-related Expectations and Risky Sexual Behavior in Malawi

Adeline Delavande

University of Essex and Nova School of Business and Economics

Hans-Peter Kohler

University of Pennsylvania

May 2015

Abstract

We use probabilistic expectations data elicited from survey respondents in rural Malawi to investigate how risky sexual behavior may be influenced by individuals' expectations about survival and future HIV status, which in turn depend on the perceived impact of HIV/AIDS on survival, expectations about own and partner's current HIV status, and expectations about HIV transmission rates. Subjective expectations, in particular about mortality risk but not the risk of living with HIV, play an important role in determining the decision to have multiple sexual partners. Using our estimated parameters, we simulate the impact of various policies that would influence expectations. An information campaign on mortality risk would decrease risky sexual behavior on average, while an information campaign on HIV transmission risks, which tend to be overestimated by respondents, would actually increase risky behavior. Also, the expansion of anti-retroviral therapy (ART) treatments to all individuals infected with HIV would increase risky sexual behavior for a quarter of the HIV-negative individuals or those who have not been tested because they are aware that ART increases life expectancy, and thus reduces the cost of becoming HIV-positive.

We would like to thank Charles Bellemare, Iliyan Georgev, Alberto Holly, Erik Meijer, Chuck Manski, Steve Pudney, Joao Santos Silva, Olivier Sterck and Luis Vasconcelos for comments. We thank seminar participants at the 2009 Choice Symposium, the 2011 Dauphine Workshop on "Recent Advances on the Role of Beliefs in Decision Theory," the 2014 European meeting of the Econometric Society, DIW Berlin, the Faculdade de Economia da Universidade do Porto, the Federal Reserve Bank of New York, RAND Corporation, the School of Economics and Management of the Technical University of Lisbon, the department of Economics of the University of Essex, and the Institute for Social & Economic Research of the University of Essex for comments. Support for this research through the National Institute of Child Health and Development (grant number R03HD058976, R21 HD050652, R01 HD044228, R01 HD053781), Nova Forum and the ESRC Research Centre on Micro-social Change is gratefully acknowledged. This research has also benefitted from pilot grant funding by Population Aging Research Center (NIA P30 AG12836), the Boettner Center for Pensions and Retirement Security, and the Population Studies Center (NICHD R24 HD-044964), all at the University of Pennsylvania.

1. Introduction

Sub-Saharan Africa (SSA) is the region of the world most affected by the HIV/AIDS epidemic. It accounts for over two thirds of all people infected worldwide, with adult prevalence rates reaching above 25% in some countries (UNAIDS, 2012). Heterosexual intercourse is the most common pathway of infection in the region. While there is a growing emphasis on biomedical interventions to prevent HIV infections or improve the health of HIV-positive individuals, interventions targeted at behavioral change remain an essential part of the HIV/AIDS prevention strategy, especially in poor countries (UNAIDS, 2012). To effectively change risky sexual behavior, it is important to understand why individuals who live in high-HIV-prevalence environments engage in those behaviors. There are several non-mutually exclusive explanations that can be put forward. One explanation may be related to differential preferences (e.g., individuals in SSA may have higher utility from having multiple sexual partners). Another reason may be the high burden of disease and the resulting high levels of non-HIV-related mortality, lowering the return of safe sex strategies. For example, Oster (2012) and Philipson and Posner (1993) argue that individuals have little motivation to adopt risk-prevention strategies, as these strategies provide only limited gains in terms of longer life expectancy, while they are "costly" in monetary terms (e.g., purchasing condoms) or foregone pleasures (e.g., lower levels of satisfaction as a result of giving up extramarital partners). Another reason may be that individuals have misperceptions about their current HIV status, transmission rates or the HIV prevalence in their community. With data on behavior only, it is not possible to distinguish explanations based on preferences from those based on beliefs, thereby limiting the ability to devise effective behavioral interventions.

In this paper, we use very rich data on probabilistic beliefs that we collected directly from rural Malawi survey respondents to investigate the role of HIV/AIDS-related expectations on the decision to engage in risky sex (in this case, having multiple sexual partners), and simulate the impact of various policies on risky behavior. We develop a simple, though quite general, two-period theoretical framework which highlights the role of expectations in the decision to engage in risky sexual behavior. Our framework shows that the subjective survival probabilities associated with having risky sex and with having safe sex are crucial for decision-making. It also shows that these probabilities depend in turn on a set of six subjective expectations: (i) expectations of survival conditional on being healthy, (ii) expectations of survival conditional on being infected, (iii) expectations about own current HIV status, (iv) expectations about partners' current HIV status, (v) expectations about HIV transmission rate associated with safe sex and (vi) that associated with risky sex. A unique feature of this paper is that we have data on *all* those expectations that are potentially relevant for behavior. We combine those expectations data with data on sexual behavior to estimate our model. We find that the difference in subjective survival probability associated with having multiple sexual partners versus having one partner plays an important role in determining the decision to have multiple sexual partners. We also find that the estimate for the disutility parameter associated with being infected with HIV (compared to being healthy) is not statistically significantly different from zero. This suggests that (i) individuals in rural Malawi are forward-looking and take into account mortality risk when making health-related choices, (ii) the threat of reduced survival associated with HIV, rather than the threat of living with HIV, influences sexual behavior in the SSA context, (iii) expectations data are useful in understanding risk-taking behavior in the context of developing countries, and (iv) changing individual beliefs may be one way to change risky sexual behavior in high-HIV-prevalence environments.

Using our estimated preference parameters, we simulate the impact of various policies that would influence individual expectations. We find that an information campaign on HIV transmission risks, leading people to revise their subjective beliefs to statistics from medical studies, would have a counter-productive effect and *increase* the average probability of having multiple partners from 21.9% to 26.9% for men and from 2.6% to 3.8% for women. This is because respondents widely overestimate the relative impact of having multiple partners on the average probability of becoming infected with HIV compared to having one partner. However, providing information on the mortality risk of someone healthy and of someone infected with HIV, leading people to revise their beliefs to available statistics from life tables, would have a positive impact and *decrease* the average probability of having multiple partners to 15.6% for men and 2.3% for women. This is because individuals *under*-estimate the magnitude of the negative impact of HIV/AIDS on survival. To our knowledge, providing information on mortality risk has never been implemented and may be a new avenue for campaign. Also, we investigate the impact of the expansion of anti-retroviral therapy treatments (ART) to all individuals infected with HIV. While this is a medical intervention targeted to infected people, it may have repercussion on everyone's sexual behavior.[1] In particular, it could change beliefs about mortality risk if infected. We find that such an expansion would increase the probability of having multiple sexual partners for about a quarter of the HIV-negative individuals or those who have not been tested. Because those individuals are aware that ART increases life expectancy, universal treatment reduces the perceived cost of becoming HIV-positive. Note that because we do not know how respondents would revise their beliefs in light of new information or new treatment options, these results can be seen as an upper bound of the anticipated behavioral change.

While survey respondents in developed countries have increasingly been asked about subjective expectations in the last 20 years, only recently has this occurred in developing countries. Delavande et al. (2011) and Delavande (2014) review the existing evidence and conclude that collecting expectations data in developing countries is both feasible and valuable.[2] In this paper, we use data on probabilistic expectations about a wide range of events that we have collected as part of the 2006 wave of the Malawi Longitudinal Study of Families and Health (MLSFH, Kohler et al, 2014) covering more than 3,000 adult respondents in rural Malawi.[3] In Delavande and Kohler (2009), we find that respondents from Malawi provide meaningful expectations in probabilistic format according to various criteria: most respondents provide probabilities that are consistent with basic properties of probability theory, the subjective expectations are systematically correlated with observable characteristics (such as gender, age, education, and region of residence) in the same way that actual outcomes vary with these variables, and expectations about future events vary across individuals in the same way as individuals' past experience does. Yet, respondents exhibit a lot of heterogeneity in expectations.

The advantage of using expectations data in empirical work is that it mitigates a basic identification problem that researchers face when using data on choices only: observed choices may be consistent with many combinations of expectations and preferences. Although data on expectations are becoming available, only a limited number of studies have until now employed them to draw inferences on behavior. Recent studies incorporating expectations into econometric

[1] Among HIV-positive individuals, Lakdawalla et al. (2006) find that the introduction of ART increases sexual activity in the U.S. while Thirumurthy et al. (2012) find that enrollment in AIDS treatment programs increases the frequency of sex but also condom use in Kenya.

[2] See Manski (2004) and Hurd (2009) for a review of the literature on expectations in developed countries.

[3] The MLSFH has previously also been known as the Malawi Diffusion and Ideational Change Project (MDICP).

models have addressed various decisions, such as contraception choice (Delavande, 2008a), portfolio allocation (Delavande and Rohwedder, 2011, Kezdi and Willis, 2011), fertility and sexual behavior (Shapira, 2010, de Paula et al., 2014), education (Zafar, 2013, Arcidiacono et al., 2012), teacher career (van der Klaauw, 2012), committing a crime (Lochner, 2007), migration (McKenzie et al., 2013), strategies in games (Nyarko and Schotter, 2002, Bellemare et al., 2008), and the timing of Social Security claiming and retirement (van der Klaauw and Wolpin, 2008, Hurd et al., 2004). We contribute to this line of work that combines choice data with data on subjective expectations to draw inferences on preferences. We do this in the context of the HIV/AIDS epidemic in SSA.

Due to potential endogeneity issues, it is challenging in many empirical applications to evaluate how expectations of events over which individuals have some control *causally* affect their decisions: Unobservable characteristics may influence both the formation of expectations and decision-making. Few papers using expectations data have addressed the endogeneity issue directly (see discussion in van der Klaauw, 2012).[4] We deal with the potential endogeneity arising from the dependence of expectations on past behavior by estimating a system of equations: beliefs about current HIV status depend on past sexual behavior and observable characteristics; the decision to get tested for HIV is explicitly estimated; and the decision to engage in risky sexual behavior depends on individual HIV/AIDS-related expectations.

Our paper builds on several studies that have analyzed the impact of the belief about *one single* HIV/AIDS-related event on sexual behavior in SSA. De Paula et al. (2014), for example, use data from the MLSFH to evaluate the impact of beliefs about one's own HIV status on the likelihood of engaging in risky behavior. They find that downward revisions in beliefs of being HIV-positive increase risky behavior. Unlike them, Gong (2015) finds that, in the Tanzanian and Kenyan context, individuals who believed themselves to be at low risk of infection have an increased likelihood of contracting a sexually transmitted disease (STD), indicating riskier sexual behavior, after a positive HIV test, while the reverse is true for those who are surprised by a

[4] De Paula et al. (2014) use a panel data estimator which accommodates unobserved heterogeneity as well as belief endogeneity arising from the dependence of current beliefs on lagged behaviors, Lochner (2007) uses fixed-effect instrumental-variable estimates, and Bellemare et al. (2008) model preferences and beliefs jointly to address the endogeneity issue.

negative HIV test.[5] Without using changes in beliefs about one's own HIV status directly, several papers have also looked at the causal impact of learning one's HIV status on subsequent sexual behavior (e.g., Thornton, 2008; Delavande and Kohler, 2012).

Few papers have assessed the role of beliefs about transmission risks on sexual behavior (Meekers and Klein, 2002; Lammers et al., 2013). Sterck (2013) documents a large over-estimation of HIV transmission risks among students in Burundi and develops a behavioral model yielding a U-shaped relationship between risky behavior and expected transmission risk. We are not aware of any study looking directly at the impact of beliefs about mortality risk on sexual behavior. Indirectly but related to this, Oster (2012) finds that a high rate of non-HIV mortality decreases the change in sexual behavior due to an increase in HIV prevalence in SSA.

We argue that beliefs about all those events are likely to be crucial to the decision to engage in risky sex and therefore ought to be considered jointly in empirical work. We therefore complement and improve on the current literature by evaluating the impact of sexual behavior on beliefs about own and partner's HIV status, transmission risks (conditional on having a safe and a risky behavior) and mortality risk (conditional on being infected and on being healthy). This allows us to evaluate how a wide range of policies such as information campaigns on mortality risks and on transmission risks, and the roll-out of ART affect risky sexual behavior. Understanding the potential magnitude of the behavioral response to change in expectations is particularly relevant in the current context in which AIDS-related mortality and HIV transmission risks are rapidly changing, mostly as a result of ART (Bor et al., 2013, Cohen, 2011).[6] It is important to determine potential trends in behavior if individuals start to internalize those reduced probabilities. Note, however, that our analysis is restricted to how changes in beliefs change risky sexual behavior, but we do not model how this change in risky behavior may influence HIV prevalence. General equilibrium models are required to make that link (see

[5] In the US context, Boozer and Philipson (2000) find that individuals who are surprised by a HIV test results change their behavior more.

[6] Bor et al. (2013) report that in a cohort of people living in rural South Africa, adult life expectancy was 49.2 years in 2003, the year before ART became available in the public-sector health system. By 2011, adult life expectancy had increased to 60.5 years—an 11.3-year gain. Similar reductions due to the expansion of ART have been documented in Malawi and other SSA countries (Floyd et al., 2012). Cohen (2011) reports findings from a clinical trial conducted by the HIV Prevention Trials Network showing that ART reduces the risk of heterosexual HIV transmission by 96%.

for example Greenwood et al., 2013).

The paper proceeds as follows. Section 2 presents the theoretical framework that motivates the empirical analysis. Section 3 describes the data. Section 4 presents the econometric specification and Section 5 the analytical sample. Section 6 analyzes the role of subjective expectations in the decision to engage in risky sex and considers a series of robustness checks such as misreporting of sexual behavior or HIV testing outside the MLSFH surveys. Section 7 presents the policy simulation results.

2. Theoretical framework

Consider a sexually active individual i who has two periods left to live (period 1 and period 2). In period 1, she can choose between 2 different actions: $a = 0$, having sex with one partner only; $a = 1$, having sex with multiple partners. Her period 1 utility depends on the immediate utility from sex $V_i(a)$ associated with action a. Her period 2 utility is HIV-status-dependent, and equals ${U_i}^+$ if she is HIV-positive in period 2 and ${U_i}^-$ if she is HIV-negative in period 2. Individual i can enjoy period 2 utility only if she survives to period 2. The subjective probability of surviving to period 2 with a given HIV status depends on whether the individual believes that she will be infected at the end of period 1. She may believe that she has already been infected with HIV before period 1 or that she can contract HIV during period 1. Her subjective probability of surviving to period 2 with a given HIV status is therefore a function of the action taken in period 1 (since period 1 action may influence HIV status) and her subjective beliefs f_i^1 of being infected with HIV at the beginning of period 1. For example, if we consider someone who believes there is no chance she is infected with HIV at the beginning of period 1 (i.e., $f_i^1 = 0$), her subjective probability of surviving to period 2 with a HIV-positive status if she engages in action a is $p_i(a)S_i^+$, and her subjective probability of surviving to period 2 with a HIV-negative status is $\left(1 - p_i(a)\right)S_i^-$, where $p_i(a)$ is individual i's subjective probability of becoming HIV-positive if she engages in action a, S_i^+ is i's subjective probability of surviving to period 2 conditional on being HIV-positive, and S_i^- is i's subjective probability of surviving to period 2 conditional on being HIV-negative. We further assume that the utility function depends on a random term ε_{ia} that is unobservable to the econometrician and captures heterogeneity in tastes. Individual i chooses the action a that maximizes her lifetime subjective expected utility, i.e., she solves the following problem:

$$max_{a\in\{0,1\}}\{V_i(a)+f_i^1S_i^+U_i{}^+ + (1-f_i^1)\big(p_i(a)S_i^+U_i{}^+ + \big(1-p_i(a)\big)S_i^-U_i{}^-\big)+\varepsilon_{ia}\}.$$

Overall, a riskier sexual behavior may increase the direct pleasure from sex in period 1 but, by potentially increasing the (subjective) risk of becoming HIV-positive, it may also decrease the (subjective) probability of surviving to period 2, and therefore of enjoying period 2 utility at all, while also decreasing the probability of enjoying $U_i{}^-$ rather than $U_i{}^+$.[7]

3. The data: Malawi Longitudinal Study of Families and Health (MLSFH)

The analyses in this paper are based on the 2006 and 2008 waves of the Malawi Longitudinal Study of Families and Health (MLSFH).[8] The MLSFH is based in three regions of rural Malawi: Balaka, Mchinji and Rumphi. Balaka district is located in the Southern Region of Malawi, primarily inhabited by Yao-speaking individuals and is predominantly Muslim. Mchinji district is located in the Central Region near the border with Zambia. It is primarily inhabited by Chewa-speaking individuals, with almost equal proportions of Catholics and Protestants. Rumphi district in the Northern Region of the country is inhabited primarily by Tumbuka-speaking individuals who are predominantly Protestant. A "Cohort Profile" of the MLSFH, providing detailed discussion of MLSFH sampling procedures, survey methods, survey instruments and biomarkers, and analyses of attrition is available in Kohler et al. (2014). The MLSFH cohorts were selected to represent the rural population, where the majority of Malawians (85%) live in poor health conditions similar to those prevailing in other rural SSA low-income countries (over 60% of total SSA population lives in rural areas): high morbidity/mortality, over-burdened health facilities, and frequently unmet nutritional needs. The rural population predominantly engages in home production of crops (mostly maize), complemented by small-scale market activities.

[7] Note that the specification of the utility function does not allow for consideration of altruism. One possibility is to assume that the decision-maker gets disutility $-d_i$ if she infects her spouse, which could happen with probability p^s if she is HIV+ (which we assume is independent of action a, i.e. whether she has sex with her spouse only or with multiple partners). Without loss of generality, let c_i be a constant such that $U_i{}^+ = U_i{}^- - c_i$ (as in Section 4, it measures the cost of living with HIV). The maximization problem presented would become: $max_{a\in\{0,1\}}\{V_i(a)+f_i^1S_i^+(U_i{}^- - c_i - p^sd_i)+(1-f_i^1)\big(p_i(a)S_i^+(U_i{}^- - c_i - p^sd_i)+\big(1-p_i(a)\big)S_i^-U_i{}^-\big)+\varepsilon_{ia}\}$. This would mean that, if d_i >0, we would over-estimate (in absolute value) the cost c_i of living with the HIV.

[8] Additional information about the MLSFH is available on the project website at http://www.malawi.pop.upenn.edu.

HIV/AIDS is widespread, including in the MLSFH study population. Comparisons with the Malawi Demographic and Health Survey showed that the MLSFH sample populations are reasonably representative of the rural Malawi population (Kohler et al., 2014; Anglewicz et al., 2007).

In 2006, the MLSFH included more than 3,200 male and female respondents aged 17 to 60, who were asked about a wide range of demographic, health, and socio-economic characteristics. In 2008, slightly more than 4,000 respondents were interviewed, with the additional respondents resulting primarily from a new parent sample that extended the age range from 17 to 92 years by also interviewing parents of earlier MLSFH respondents.

An innovation of the 2006 and 2008 waves was the inclusion of an *interactive elicitation technique for subjective expectations* that was based on asking respondents to allocate up to ten beans on a plate to express the likelihood that an event will be realized (Delavande and Kohler, 2009). Interviewers introduced the interactive elicitation technique with a short introduction (see Appendix B). After any clarifying questions, respondents were first asked a training question about the probability of winning in a local board game (Bawo), followed by a series of expectations questions related to economic and health outcomes. They were in particular asked about the probability that they are currently infected with HIV, and that their spouse/partner is currently infected with HIV. The questionnaire also included several questions about the one-year, 5-year, and 10-year mortality of the following hypothetical individuals: (*i*) a woman/man of the respondent's age who is healthy and does not have HIV; (*ii*) a woman/man of the respondent's age who is infected with HIV; (*iii*) a woman/man of the respondent's age who is sick with AIDS and is treated with ART. The gender used in the scenarios was the same as that of the respondent. Respondents were also asked the probability that someone of the same gender who was currently healthy would become infected with HIV in the next 12 months if she (a) is married to an HIV-positive spouse, or (b) has several sexual partners in addition to her spouse. Respondents were also asked their perception of the village HIV prevalence (from 0 to 10).

Delavande and Kohler (2009) provide a detailed analysis and evaluation of the probabilistic expectations collected using the above interactive elicitation technique. Key findings from the 2006 data include these: (a) About 99% of the respondents are found to provide beliefs consistent with basic properties of probability theory when asked about nested events; (b) in basically all the considered domains, subjective beliefs vary considerably across individuals; (c) subjective expectations are systematically correlated with observable characteristics – such as gender, age, education, and region of residence – in the same way that actual outcomes vary with these variables (e.g., expectations about infant mortality exhibit regional differences that are similar to actual outcomes, and expectations about economic outcomes vary with socio-economic status in the expected directions); and (d) expectations about future events vary across individuals in the same way as individuals' past experience does.

Another innovative aspect of the MLSFH is the collection of HIV status that incorporated in 2004 an experimental design that created exogenous variation in which respondents learned the result of the HIV test. Specifically, as part of a randomized experiment to study the determinants of HIV testing uptake, respondents were offered a free HIV test at the end of the 2004 interview (Thornton, 2008). At the time of testing, respondents were given randomly assigned vouchers redeemable for a sum of money upon picking up their HIV test results at local clinics a couple of months after testing. A quarter of the participants did not receive a positive incentive. The remaining three quarters received a positive incentive ranging from 10 to 300 Kwachas. The average incentive was about 100 Kwachas, equivalent to a day's wage (agricultural labor). Thornton (2008) finds that learning one's HIV results was highly responsive to receiving a positive financial incentive.[9] In 2006 and 2008, respondents were re-visited by nurses shortly after completing the interview and were offered a free at-home HIV test with immediate results. There was no financial incentive provided in 2006 and 2008 for learning one's HIV status. In 2006, 93% of the respondents agreed to be tested and 98% of those who were tested learned their HIV status. Overall, 5.1% of the tested were HIV-positive. Nevertheless, 14% of the sample was not found by the team of nurses at the second visit, and were therefore not offered a test (see Table 1).

Finally, the questionnaire asks several questions about sexual behavior. Of particular interest to this paper is the number of sexual partners in the last 12 months, asked in 2008.[10] We use self-reported behavior, which may suffer from reporting biases. In a subsample of the

[9] In 2004, 91% of the respondents agreed to be tested and among those, 69% went to pick up their test result.

[10] We focus on the number of partners, abstracting from condom use, for several reasons. First, respondents were not asked about condom use in 2008. Second, condoms are relatively infrequently used in Malawi, especially in regular relationships (Chimbiri, 2007) so the number of sexual partners is likely to be the most important margin of behavioral adjustment. Third, women may have limited decision power regarding the use of condom.

MLSFH adolescents, sexually transmitted infection (STI) status (which was collected in 2004) and self-reported behavior have been found to be positively correlated (Mensch et al., 2008). We however do not have STI biomarker information in 2008 to complement our analysis based on self-reported sexual behavior.[11] We discuss the robustness of our results to potential misreporting in Section 6.3.2.

Figure 1 shows the important aspects of the timeline of the data collection.

4. Econometric specification

Based on section 2, the probability of choosing multiple sexual partners is the probability that the action $a_i = 1$ yields higher subjective expected utility than the action $a_i = 0$, i.e.:

$$P(a_i = 1) = P\begin{pmatrix} V_i(1) + f_i^1 S_i^+ U_i{}^+ + (1 - f_i^1)\big(p_i(1) S_i^+ U_i{}^+ + \big(1 - p_i(1)\big) S_i^- U_i{}^-\big) + \varepsilon_{i1} \geq \\ V_i(0) + f_i^1 S_i^+ U_i{}^+ + (1 - f_i^1)\big(p_i(0) S_i^+ U_i{}^+ + \big(1 - p_i(0)\big) S_i^- U_i{}^-\big) + \varepsilon_{i0} \end{pmatrix}$$

$$= P\left(\varepsilon_{i0} - \varepsilon_{i1} \leq V_i(1) - V_i(0) + (1 - f_i^1)\big(p_i(1) - p_i(0)\big)\big(S_i^+ U_i{}^+ - S_i^- U_i{}^-\big)\right).$$

Without loss of generality, let c_i be a constant such that $U_i{}^+ = U_i{}^- - c_i$, then we have

$$P(a_i = 1) = P\begin{pmatrix} \varepsilon_{i0} - \varepsilon_{i1} \leq V_i(1) - V_i(0) + \\ (1 - f_i^1)\big(p_i(1) - p_i(0)\big)(S_i^+ - S_i^-)U_i{}^- - (1 - f_i^1)\big(p_i(1) - p_i(0)\big) S_i^+ c_i \end{pmatrix} \quad (1)$$

The term $p_i(1) - p_i(0)$ is the difference in probability of becoming infected with HIV between having multiple partners (action 1) versus having one partner (action 0), while the term $\big(p_i(1) - p_i(0)\big)(S_i^+ - S_i^-)$ is the difference in survival probability between having multiple partners versus one partner for those who believe they are currently not infected with HIV. We seek to draw inferences on the structural preference parameters $U_i{}^-$ and c_i to evaluate whether

[11] We view self-reported behavior and biomarker as complements because STI status does not have a one-to-one relationship with frequency of unprotected sexual behavior (e.g., Fishbein et al., 2000). In the MLSFH context, tests of gonorrhoea, chlamydia, and trichomoniasis were conducted in 2004 but were not repeated given their low prevalence in the study population (Kohler et al., 2014). The low prevalence suggests that those specific STIs were not very common and/or, the STI management for these curable STIs was relatively good, making it a poor proxy for risky sexual behavior. Using a standard epidemiological model, Corno and de Paula (2014) find that STI biomarkers have a higher probability of misclassification than self-reported behavior in populations with low STI prevalence.

subjective beliefs are important for decision-making. In particular, we want to evaluate whether individuals are forward-looking, and take into account relative survival risk and future disutility associated with being infected with HIV when making decisions. U_i^- and c_i is relevant to the decision-making only of those who believe there is a non-zero chance that there are HIV-negative, i.e. for whom $1 - f_i^1$ is different from zero. In our empirical analysis, we will allow U_i^- and c_i to vary by gender, age and wealth.

We use beliefs elicited *in 2006* to explain sexual behavior that occurs *between 2006 and 2008* (see Figure 1). The timing is important because sexual behavior may lead individuals to revise their beliefs (in particular about current HIV status) subsequently. Therefore, it is critical to use beliefs elicited *prior* to the decision to engage in risky behavior to avoid issues of reverse causality. Yet, there may still be some issues in the estimation of equation (1) due to potential endogeneity of beliefs arising from the dependence of *current* beliefs on *past* behaviors. In particular, unobserved heterogeneity capturing time-invariant preferences for the number of partners or the search cost of having multiple partners may be correlated with the beliefs f_i^1 , if beliefs at the beginning of period 1 depend on the prior number of partners, or if this unobserved heterogeneity also influences the decision to get tested for HIV.

To deal with this concern, we estimate a four-equation model where all random terms are allowed to be correlated. The first equation models the 2006 subjective probability f_i^0 of being infected with HIV prior to the 2006 HIV test. It is a reduced-form equation which depends on demographic characteristics and lagged sexual behavior x_{i4}. Note that because f_i^0 is confined to the range [0,1], we estimate this equation as a truncated linear regression, as if there was a latent variable f_i^{0*} which is observed only when it is between 0 and 1. The second and third equations model the process by which someone gets tested for HIV in 2006. This process is important to take into consideration in this context if unobserved heterogeneity influencing HIV testing (and therefore beliefs about HIV status) also influences the decision to have multiple sex partners. In order to get tested for HIV, a respondent has to be found by the team of nurses on the day of the HIV test, and she has to agree to get tested. The second equation in the system therefore deals with the propensity to be found on the day of the HIV test, $Found_i^*$, which depends on demographic characteristics and past sexual behavior x_{i3}, and on the probability f_i^0 of being infected with HIV prior to the 2006 HIV test. The third equation models the decision to get tested for HIV, $Test^*$, conditional on being found on the day of the test, which depends on demographic characteristics and past sexual behavior x_{i2}, and on the probability f_i^0 of being infected with HIV prior to the 2006 HIV test. Finally, the last equation, which is the one of substantive interest, models the propensity to have multiple partners $a_i{}^*$. As defined in equation (1), it will depend on the difference in immediate utility from sex associated with having multiple partners and one partner, which we assume depends on demographic characteristics and lagged sexual behavior (i.e., $V_i(1) - V_i(0) = \beta x_{i1}$)), and on the two subjective beliefs $(1 - f_i^1)\big(p_i(1) - p_i(0)\big)(S_i^+ - S_i^-)$ and $(1 - f_i^1)\big(p_i(1) - p_i(0)\big)S_i^+$ that characterize how subjective survival probabilities affect decisions about risky sexual behaviors. Note that those beliefs include the potentially endogenous post 2006 HIV test belief f_i^1, which is equal to the pre-test belief f_i^0 if a respondent did not get tested for HIV in 2006 or to the actual 2006 HIV status if the respondent got tested for HIV in 2006. The system of equations (2) (along with the law of motion of the probability f_i) is formally given by the following set of equations:[12]

2006 pre-test probability f_i^0 of being currently infected with HIV:

$$f_i^{0*} = \beta_4 x_{i4} + \varepsilon_{i4}$$
$$\text{with } f_i^0 = f_i^{0*} \text{ if } 0 \le f_i^{0*} \le 1 \qquad (2.4)$$
$$\text{and } f_i^0 \text{ unobserved otherwise.}$$

Found by MLSFH nurse team for HIV test in 2006:

$$\begin{aligned} Found_i{}^* &= \beta_3 x_{i3} + \gamma_3 f_i^0 + \varepsilon_{i3} \\ Found_i &= 1 \; if \; Found_i{}^* > 0 \\ Found_i &= 0 \text{ otherwise} \end{aligned} \qquad (2.3)$$

Participation in MLSFH HIV test in 2006:

$$\begin{aligned} Test_i{}^* &= \beta_2 x_{i2} + \gamma_2 f_i^0 + \varepsilon_{i2} \\ Test_i &= 1 \; if \; Test_i{}^* > 0 \\ Test_i &= 0 \text{ otherwise} \\ Test_i &\text{ is observed only if } Found_i = 1 \end{aligned} \qquad (2.2)$$

2006 post-test probability of being currently infected with HIV, f_i^1 :

$$f_i^1 = 0 \text{ if } Test_i = 1 \text{ and test result is negative}$$
$$f_i^1 = 1 \text{ if } Test_i = 1 \text{ and test result is positive}$$

[12] We number equations in a decreasing order to reflect the fact that equation (2.1) is the one of substantive research interest, that we therefore present in the first column of our results tables.

$$f_i^1 = f_i^0 \text{ if } Test_i = 0 \text{ or } Found_i = 0$$

Having multiple partners in 2008, a_i:

$$a_i^* = \beta_1 x_{i1} + U_i^{-} P_u - c_i P_c + \varepsilon_{i1}$$
$$a_i = 1 \; if \; a_i^* > 0 \qquad (2.1)$$
$$a_i = 0 \text{ otherwise.}$$

The correlation of the residuals across equations 2.4-2.1 is given by

$$\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \end{pmatrix} = N\left(\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho_{12} & \rho_{13} & \rho_{14} \\ \rho_{12} & 1 & \rho_{23} & \rho_{24} \\ \rho_{13} & \rho_{23} & 1 & \rho_{34} \\ \rho_{14} & \rho_{24} & \rho_{34} & \rho_{44} \end{pmatrix}\right),$$

and $P_u = (1 - f_i^1)\big(p_i(1) - p_i(0)\big)(S_i^+ - S_i^-)$ and $P_c = (1 - f_i^1)\big(p_i(1) - p_i(0)\big)S_i^+$ are subjective probabilities, discussed in more detail below, that affect the decision to have multiple sexual partners.

We seek to estimate the parameters $\{\beta_4, \beta_3, \gamma_3, \beta_2, \gamma_2, \beta_1, U_i^{-}, c_i\}$. Equation (2.4) is a truncated linear regression (without mass point at 0 or 1) while equations (2.3), (2.2) and (2.4) are probit regressions, for which the variance of the random term is normalized to 1. The system is partially recursive in the sense that f_i^0, the dependent variable in (2.4), is an independent variable in equations (2.2) and (2.3) and enters indirectly in (2.1) through f_i^1 (because $f_i^1 = f_i^0$ for respondents who have not been tested for HIV). But note that the dependent variables $Found_i^*$ in (2.3) or $Test^*$ in (2.2) are not independent variables in (2.1) (Equations (2.2), (2.3) and (2.1) are in fact seemingly unrelated). Identification requires therefore at least one variable in x_{i4} not included in x_{i3}, x_{i2} and x_{i1} (Maddala, 1983, Chapter 5.7, model 4) -- but does not require a variable in x_{i3}, x_{i2} not included in x_{i1} since the dependent variables of equation (2.3) and (2.2) do not enter equation (2.1). Note also that Equations (2.2) and (2.3) consist of a classic probit model of sample selection, with (2.3) being the selection equation. For the identification not to rely purely on functional form, the selection equation should have at least one variable that is not in the probit equation, i.e. we need a variable in x_{i3} not in x_{i2} (Van de Ven and Van Pragg, 1981). We discuss our exclusion restrictions in details in Section 5.3. We will estimate the system of equations (2) by maximum likelihood.[13] We cluster standard errors at the couple level.

5. Analytic sample and definition of the variables

5.1. Analytic sample

Different samples are used to estimate the various equations of the system (2). We use all 2006 MLSFH respondents to estimate equations (2.4) and (2.3), and the 2006 respondents who were found by the team of nurses for testing for equation (2.2). We use the 2008 MLSFH respondents who have been sexually active in the 12 months prior to the survey for our analysis of sexual behavior.[14] We also exclude 202 men who are in a polygamous marriage and have therefore multiple sexual partners within marriage. Our empirical strategy relies on estimating whether sexual behavior in the last 12 months reported in 2008 is influenced by elicited beliefs reported in 2006. However, 23% of the 2006 respondents could not be resurveyed in 2008, so their behavior cannot be used to make inference on U_i^- and c_i. We investigate in Section 6.3.4 the robustness of our results when we take attrition into consideration. Table 1 presents basic characteristics of the 2006 respondents, and of the 2008 respondents who answered the 2006 probabilistic expectations.[15]

5.2. Definition of variables

We start by describing the dependent variables in the system of equations (2). Table 2 presents their descriptive statistics.

- *2006 pre-test probability of being currently infected with HIV*. The variable f_i^0 is the respondent's answer about the likelihood of being currently infected with HIV, elicited in 2006 (and re-scaled from zero to 1 by dividing the number of beans by 10). The average belief is 0.11. However, the distribution is skewed: 66 percent of the respondents report a probability of zero of being currently infected with HIV. We plot in Figure 2 the distribution of subjective probability

[13] We use the cmp Stata command developed by Roodman (2011). The likelihood function for the fully recursive model is given on page 176 of Roodman (2011) while the likelihood function for the truncated model is given by equation (5).

[14] We establish the robustness of our results to the inclusion of non-sexually active respondents in Section 6.3.3.

[15] Our analytical sample includes respondents who have been tested for HIV and found out that they were HIV+ in 2004. Whether we include them or not does not change our results. Note also that we also include respondents who learned they were HIV+ in 2006 in the estimation of equation (2.4). For those, the beliefs P_u and P_c are equal to zero. Whether we exclude them or not does not change our results.

of being currently infected with HIV elicited in 2006, according to their actual 2006 HIV status (2006 test result not known by respondents at the time of the survey). Subjective beliefs range between zero and one, though a large proportion is concentrated at zero. However, the concentration at zero varies by HIV status, with 40 percent of the HIV-positive individuals reporting a zero chance, compared to 68 percent of the HIV-negative individuals. We also note that 11 percent of the HIV-positive report a probability one of being infected, compared to less than 1 percent of the HIV-negatives. The distribution of those who tested negative and those who did not get tested is very similar.

- *Found for HIV test*. We define the variable $Found_i = 1$ if the respondent was found by the team of nurses conducting HIV test in 2006, and 0 otherwise. Overall, 86 percent of respondents interviewed as part of the main survey in 2006 were found by the team of nurses.
- *HIV testing*. We define the variable $test_i = 1$ if the respondent agreed to be tested and learned his HIV status at the end of the 2006 interview, and 0 otherwise (conditional on being found). Overall, 90 percent of the respondents who were found by the team of nurses learned their HIV status in 2006.
- *Having multiple partners in 2008.* The variable a_i is equal to 1 if the respondent reports in 2008 having had more than 1 sexual partner in the last 12 months, and zero if the respondent reports having had only one sexual partner in the last 12 months. Table 2 shows that among respondents who were sexually active in 2008, 13% had more than one sexual partner. Note however that there is a large difference in self-reports about multiple partnerships by gender: this percentage is 22% among males compared to 2% among females.

We now describe how we construct the two individual-specific subjective beliefs $P_u = (1 - f_i^1)\big(p_i(1) - p_i(0)\big)(S_i^+ - S_i^-)$ and $P_c = (1 - f_i^1)\big(p_i(1) - p_i(0)\big)S_i^+$ that are critical for the decision to have multiple sexual partners. The former reflects the difference in survival probability between having multiple partners versus one partner for those who believe they are currently not infected with HIV, multiplied by the probability of being not infected with HIV; the latter reflects the difference in the probability of surviving to period 2 with HIV between having multiple partners versus one partner for those who believe they are currently not infected with HIV, multiplied by the probability of being not infected with HIV. Table 3 presents the descriptive statistics of the various relevant expectations for respondents who were sexually active in 2008.

- *2006 post-test probability of being currently infected with HIV,* f_i^1. We define the variable f_i^1 as follows:

$$f_i^1 = \begin{cases} 0 & \text{if the respondent learned that s/he was HIV-negative in 2006 HIV test} \\ 1 & \text{if the respondent learned that s/he was HIV-positive in 2006 HIV test} \\ f_i^0 & \text{if the respondent did not learn his/her HIV status in 2006} \end{cases}$$

The underlying assumption is that individuals revised their belief upon learning their HIV status according to the test result.[16] We evaluate the robustness of our results to this assumption in Section 6.3.5. Table 3 shows that the average post-test probability is equal to 0.05 and is thus lower than the pre-test probability, due to the fact that a large number of respondents who were tested found out that they were HIV-negative.

- *Survival expectations,* S_i^+ *and* S_i^-. We use the 2006 elicited 10-year mortality rate of a hypothetical individual of respondent's gender, age and village currently being infected with HIV to determine S_i^+ and the 2006 elicited 10-year mortality rate of a hypothetical individual of respondent's gender and age currently healthy to determine S_i^-.[17] Table 3 shows that on average respondents think that there is a 41 percent chance of being alive in 10 years conditional on

---

[16] There is no empirical evidence on how people revise their beliefs just after learning their test results. In Delavande and Kohler (2012), we report that about two-thirds of respondents who learn they were HIV-negative in 2004 allocated zero beans when asked their beliefs of infection in 2006. Because individuals may have subsequently engaged in risky behavior, this pattern in the data is consistent with our current updating assumption. More surprisingly, only 10% of the respondents who were told they were HIV-positive in 2004 provided 10 beans in response to the question about their subjective probability of being infected with HIV in 2006. Delavande and Kohler (2012) consider various explanations: (i) respondents may have believed the test result at first, but have "forgotten" about it or reassessed their status as time elapsed, specifically if they had continued to feel fairly healthy; and (ii) respondents may actually believe that they are HIV-positive, but were embarrassed to acknowledge it *vis-à-vis* the interviewer during the 2006 survey. Several results lead us to believe that embarrassment may be an issue: respondents are more likely to report zero beans in places with more HIV-related stigma, and learning a HIV-positive status lead to a change in behavior that reduces the HIV infection risk to others, suggesting an underlying change in beliefs about HIV status.

[17] With respect to the model, the definition of the survival probability variables implies that our empirical analysis focuses on the following trade-off: direct utility from sex now versus higher chance of survival in 10 years. Given the delay between HIV infection through sexual behaviours and mortality, relating sexual behaviour within a 1-year horizon to mortality over a 10-year horizon is conceptually consistent with the basic epidemiology of HIV. For example, Morgan et al. (2002) report a median time from seroconversion to death of 9.8 years in rural Uganda.

being healthy, compared to only a 12 percent chance of being alive in 10 years conditional on being infected with HIV. This shows that individuals are aware that being infected with HIV reduce life expectancy in the long run. Panel A of Table 4 shows the average survival beliefs by age group for someone healthy, infected with HIV, and sick with AIDS, and sick with AIDS but treated with anti-retroviral therapy (ART). It shows a gradient of survival expectations by age: younger people expect to live longer. It also shows that for all age groups, individuals are aware that being infected with HIV will shorten life expectancy substantially, and that being on ART will mitigate this. Comparisons with survival rates from other studies further highlights that individuals are pessimistic about survival.

- *Subjective probability of infection associated with having multiple partners,* $p_i(1)$. We use the 2006 elicited expectations about the likelihood that a hypothetical individual of the respondent's gender and village would become infected with HIV in the next 12 months if s/he was having several sexual partners in addition to a spouse. Respondents believe on average that there is an 77 percent chance of becoming infected with HIV conditional on having multiple partners (Table 3).

- *Subjective probability of infection associated with having one partner,* $p_i(0)$. This probability is again individual-specific and depends on respondents' belief about the status of their main partner. It is defined as $p_i(0) = \Pi_i \times {f_{pi}}^1$, where Π_i is person *i*'s perceived likelihood of becoming infected with HIV during the next 12 months for someone who is married to an HIV-positive individual and ${f_{pi}}^1$ is *i*'s subjective beliefs about her partner's HIV status after the 2006 HIV test and before engaging in action a_i. Note that a respondent may not know the test result of her spouse if the latter did not share the results. We assume that a respondent learned the status of her spouse after the 2006 HIV test if, in 2008, she reports that the last time her spouse got tested, he shared his test results with her, and if the last time occurred during the 2006 MLSFH data collection period. So we define ${f_{pi}}^1$ as follows:

$$f_{pi}^{\ 1} = \begin{cases} \end{cases}$$

0	if the spouse learned that s/he was HIV-negative in the 2006 HIV test and the respondent reports that the spouse shared test results
1	if the spouse learned that s/he was HIV-positive in the 2006 HIV test and the respondent reports that the spouse shared test results
2006 elicited beliefs about spouse's status	if the spouse did not learn his/her HIV status in 2006 or the respondent reports that the spouse did not share
Beliefs about village prevalence	if the respondent did not report having a main partner in 2006

Respondents believe on average that there is a 93 percent chance of becoming infected with HIV within 12 months if one is married to an HIV-positive spouse, while the average partner's probability of being infected with HIV is 8 percent. Overall, the average subjective probability of becoming infected with HIV within 12 months, conditional on having sex with one partner, is 8 percent, which is about one-tenth the perceived chance of becoming infected conditional on having multiple partners.[18] So, overall, respondents believe that having multiple partners puts them at a substantially greater risk of becoming infected with HIV.

Panel B of Table 4 shows the 2006 average subjective probabilities of becoming infected with HIV by 2006 HIV status (not known to the respondents at the time of the survey). We find that the subjective probability of infection associated with having multiple partners, $p_i(1)$, and the probability of infection if married to an HIV-positive individual, Π_i, are similar by HIV status. However, the subjective probability of infection associated with having one partner, $p_i(0)$, is larger for respondents who are HIV-positive compared to the HIV-negative and the non-tested. This is driven by the higher post-test subjective probability of a partner being

[18] Note that our analysis used beliefs regarding the transmission of HIV $p_i(0)$ and Π_i that were elicited in 2006 *before* respondents had the opportunity to get tested and learned their HIV status. Potentially, upon learning their status, respondents could have updated not only their beliefs about their own HIV status but also their beliefs about the transmission of HIV associated with various behaviors. This would be problematic only for respondents who are HIV-negative (as those transmission risk expectations would not enter the decision problem of individuals who found out that they are HIV-positive). In Delavande and Kohler (2012), we investigate the causal impact of learning HIV status in 2004 on elicited 2006 HIV/AIDS-related expectations using an instrumental-variable approach. We find that, among HIV-negative individuals, learning one's status had no impact on expectations about transmission risk. This suggests that 2006 beliefs about transmission risk are unlikely to have been revised after the 2006 HIV test by respondents who found out that they were HIV-negative.

infected with HIV.

Figure 3 shows the distribution of the difference in the probability of being alive in 10 years by HIV status $(S_i^+ - S_i^-)$ according to the number of sexual partners (one versus 2 or more). It shows that respondents who believe that HIV substantially reduce their survival (i.e., those for whom $(S_i^+ - S_i^-)$ ranges between -0.6 to -0.4) are more likely to have only one partner while respondents who believe that HIV does not have a large impact of survival (i.e., those for whom $(S_i^+ - S_i^-)$ ranges between -0.3 to 0) are more likely to have multiple sex partners.[19] This is consistent with the idea that those who expect HIV to have a large negative impact on survival choose a safer sexual behavior. Figure 2 also shows the distribution of the difference in the probability of becoming infected with HIV, $p_i(1) - p_i(0)$, rounded to the first decimal. The patterns are less clear than for the survival probability but it shows that among those who believe the difference to be the largest (i.e., equal to one), there are a more individuals having one sex partner than multiple. Again, this would be consistent with the idea that those who view the difference in transmission risk to be large choose a safer sexual behavior.

Finally, we include in all equations basic demographic characteristics (e.g., age, marital status, education, land ownership, region) and an indicator for whether the respondents report 0, 1, or more than 1 sexual partners in the 2006 interview, as those are thought to influence the 2008 sexual behavior, the propensity to be found for the HIV test, the testing decision, and pre-test beliefs about HIV status. We also include indicators for religion, as religion may influence risky behavior and risk perceptions (e.g., Trinitapoli and Regnerus, 2006).

5.3. Exclusion restrictions

As pointed out in Section 4, identification requires at least one variable in x_{i4} not included in x_{i3}, x_{i2} and x_{i1}. We use two variables for this purpose. First, we take advantage of the randomized experiment that was conducted in 2004 in which participants were provided randomized financial incentives (equal on average to 102 Kwacha, which corresponds approximately to a day's agricultural labor wage at the time) for learning one's HIV status and use an indicator for whether a respondent received a positive financial incentive as excluded

[19] The difference in mortality risk is between -0.6 and 0 for 94% of the respondents.

variable. The idea is that individuals provided with a positive financial incentive were more likely to learn their HIV status in 2004 (Thornton, 2008), which would influence their 2006 beliefs about whether they are infected with HIV. Seventy-five percent of the participants received a positive financial incentive. Thornton (2008) notes that 34 percent of the participants receiving a zero incentive learned their HIV results, while even the smallest incentive doubled that share. However, receiving a positive financial incentive in 2004 should not influence the propensity to be found by the team of nurses conducting HIV test in 2006, which is distinct from the participation in HIV testing and is largely determined by fieldwork logistics that affect the time and day of the visit attempts, nor should it directly influence the decision to have multiple sex partners in 2008.[20] Note that it could influence the decision to get tested in 2006 (if for example, people become less scared of getting tested for HIV once they have already been tested) so we do not exclude it from equation (2.3).

About 30 percent of the respondents did not participate in the 2004 HIV test (Table 1). Among those who were offered the test, only 9 percent refused to get tested. The remaining share included individuals who were not eligible, not found at the time of the 2004 HIV test (short-term migration/mobility is relatively common in rural Malawi) or not interviewed in 2004. We therefore use another variable that will be excluded from equations (2.1), (2.2) and (2.3). This is the 2006 elicited probability that a hypothetical healthy individual of the respondent's gender and village would become infected with HIV in the next 12 months with normal sexual behavior. The motivation is that this subjective probability should be related to a respondent's own probability of being infected. For example, consider a respondent who believed she was not infected with HIV 12 months prior to the survey and who also considers that she has had normal sexual behavior in that period. Her 2006 subjective probability of being infected with HIV should be equal to that reported for the hypothetical individual.[21] However, once we condition on

[20] Those who received a positive financial incentive are as likely to be found by the team of nurses in 2006 as those who received a zero incentive (92.19% vs. 91.05%, P-value for t-test of equality of mean is 0.400). Similarly, those who received a positive incentive are as likely to have multiple partners as those who received a zero incentive (12.20% vs. 12.34%, P-value for t-test of equality of mean is 0.948).

[21] More generally, let f_i^{-1} be i's subjective beliefs of being infected with HIV 12 months prior to the 2006 survey. We have $f_i^0 = (1 - f_i^{-1}) \times P_i(infection\ in\ 12\ months) + f_i^{-1}$. Let denote by T_i the 2006 elicited probability that a hypothetical healthy individual of the respondent's gender would become infected with HIV in the next 12 months with normal sexual behavior. If the respondent had normal sexual behavior, we have

the individual subjective probability of being infected with HIV and the subjective probability of becoming infected when having one and when having multiple sexual partners, the subjective probability that a hypothetical individual would become infected with HIV in the next 12 months with normal sexual behavior should be irrelevant to the HIV testing decision, the propensity to be found by the team of nurses, or the propensity to have multiple partners.[22]

In order not to rely solely on functional form assumption, the selection equation (2.3) should also have at least one variable that is not in the probit equation for HIV testing (2.2), i.e. we need a variable in x_{i3} not in x_{i2}. We use an indicator for whether the first attempted visit of the team of nurses conducting the 2006 HIV test was within one week of the first attempted visit of the 2006 survey team. A short time span between attempted visits, which is purely due to logistical considerations in the field, is likely to increase the propensity to be found, without affecting directly the testing decision. Both the survey team and the team of nurses made three attempts to find a respondent. Note that we used the first *attempted* visit date rather than actual survey date because (i) it is defined even for respondents who were not found and (ii) the interview date may not be fully exogenous (e.g., if respondents who have been interviewed on the third visit are systematically different than respondents who have been interviewed on the first visit). Our results are unchanged if we use 10 days or two weeks rather than a one week gap between visits.

6. Empirical Results

6.1. Baseline Results

We start by assuming that U^{-} and c are identical for all respondents. Table 5 presents the average marginal effects of the maximum-likelihood estimation results of equation (2).[23] In the first column of Table 5, where the indicator for having multiple partners is used as a dependent

---

$P_i(infection\ in\ 12\ months) = T_i$. If respondent i had a sexual behavior different from "normal," then $P_i(infection\ in\ 12\ months) = T_i + d_i$, where d_i may be positive if i had sexual activity above "normal" or negative if i had sexual activity below "normal." The magnitude of d may depend on past sexual activity and other demographic characteristics. Note that we do not observe f_i^{-1}. Yet, based on this framework, we expect T_i to be a strong predictor of f_i^0.

[22] For each of the variables ($Found_i$, $test_i, a_i$), we cannot reject the hypothesis that the mean belief of becoming infected with HIV in the next 12 months with normal sexual behavior is different for those whose variable equals 1 and for those whose variable equals zero, using a t-test.

[23] We present the coefficients for the truncated regression (2.4), i.e. the marginal effects for f_i^{0*}.

variable (equation 2.1), we find that the average marginal effect of the subjective belief $P_u = (1 - f_i^1)\big(p_i(1) - p_i(0)\big)(S_i^+ - S_i^-)$ is positive (equal to 0.097) and statistically significant at 5%. That is, an increase from the 25th and 75th percentile of P_u increases the probability of having multiple sex partners by 2 percentage points (which corresponds to a 15% increase in the probability of having multiple partners). Suppose that $(1 - f_i^1)$ and $(S_i^+ - S_i^-)$ are evaluated at their mean, then an increase of $\big(p_i(1) - p_i(0)\big)$ from zero to 1 would reduce the propensity to have multiple sex partners by 2.7 percentage points. Similarly, suppose that $(1 - f_i^1)$ and $\big(p_i(1) - p_i(0)\big)$ are evaluated at their means, then a decrease in $(S_i^+ - S_i^-)$ from zero to minus 1 would decrease the propensity to have multiple sex partners by 6.4 percentage points. Similarly, if $\big(p_i(1) - p_i(0)\big)$ and $(S_i^+ - S_i^-)$ are evaluated at their means, then an increase in f_i^1 from zero to 1 would reduce the propensity to have multiple sex partners by 2.0 percentage points. This provides evidence that individuals are forward-looking and take into consideration subjective expectations about relative mortality risk, HIV status, and transmission rates when making decisions related to sexual behavior.

We find however that the marginal effect of the belief $P_c = (1 - f_i^1)\big(p_i(1) - p_i(0)\big)S_i^+$, which relates to the disutility $-c$ associated with being HIV-positive rather HIV-negative, is negative as expected but much smaller in absolute value than the marginal effect of P_u, and is not statistically significantly different from zero. This suggests that the stage of sickness is not taken into consideration in the decision to have multiple sexual partners. Rather than the threat of being sick with HIV/AIDS, it is the threat of dying early that motivates safe sex strategies.

Looking at the other marginal effects that are precisely estimated, we find, as already noted in Table 2, that women are much less likely to have multiple partners than men. Also, married respondents are much less likely to have multiple partners. Respondents who have had multiple sexual partners in the past are much more likely to continue this behavior: the marginal effect is of similar magnitude (in absolute) as the marginal effect of the female dummy, and precisely estimated (statistically significant at 1%).

The second column of Table 5 presents the average marginal effects for the propensity to be found by the teams of nurses conducting HIV testing in 2006. We note that the excluded variable, i.e. an indicator for whether the time difference between the survey team's first attempted visit and the nurses' first attempted visit is less than a week, is positive and precisely estimated (P-value=0.001). Older respondents are also more likely to be found, along with non-married individuals and those having a higher subjective probability of being infected with HIV. A possible explanation is that that those who are sick (and thus with a higher belief of being infected) are more likely to be home during the nurses' visit.

The third column of Table 5 presents the average marginal effects for the HIV testing decision, conditional on being found by the team of nurses. Married respondents are less likely to agree to testing, while those who have had multiple partners in the past and those who have received a positive financial incentive in the 2004 HIV testing experiment are more likely to agree.

Finally, the last column of Tables 5 presents the average marginal effects for the equation using the pre-HIV-test subjective probability of being HIV-positive as dependent variable. Importantly, our excluded variables are strong predictor of this pre-HIV-test probability (P-value=0.048 for having received a positive financial incentive and P-value=0.059 for the belief of becoming infected with HIV in the next 12 months with normal sexual behavior, and P-value=0.032 for a test of joint significance). We see that other characteristics influence beliefs in the same way as they influence actual HIV status: female and those who had multiple partners in the past report higher chances of being infected with HIV.

The bottom panel of Table 5 shows estimates of the variances and covariances of the random terms. We find that ρ_{24} is statistically significantly different from zero at 5%, suggesting that there is a (negative) correlation between the random terms of the beliefs about HIV status equation and the HIV testing equation. However, the covariance between the random term ε_1 of the equation estimating the propensity to have multiple sex partners and the other equations is not statistically significantly different from zero, suggesting that, in this context, there is no endogeneity issue when estimating the impact of HIV-related expectations on having multiple sexual partners.[24]

Because the credibility of the results relies on the credibility of the exclusion restrictions, we also conducted additional regressions to obtain estimates for U^{-} and c when the residual

[24] Indeed, we do find estimates and standard errors very similar to those reported in the first column of table 5 if (i) we estimate a simple probit model for the decision to have multiple sexual partner, or (ii) we estimate the system of equations (2) without any exclusion restrictions.

correlation ρ_{14} is fixed to certain levels in the system of equations (2). Table A1 shows that results are qualitatively similar when ρ_{14}varies between -0.5 to 0.3, with the magnitude of the average marginal effect of P_u being actually larger when the correlation is negative. Only when the ρ_{14} is equal to 0.5 do we get an average marginal effect of P_u that is about 2/3 of the value reported in Table 5 and is imprecisely estimated. This suggests that the endogeneity issue would have to be quite severe for our results to indicate substantively different conclusions.

Note that while we estimate (2.4) as a truncated linear regression, our results are essentially unchanged if we estimate this relation as a simple linear regression. One may also wonder whether our second excluded variable, the "probability that a hypothetical healthy individual of the respondent's gender and village would become infected with HIV in the next 12 months with normal sexual behavior," is distinct enough from the two other transmission probabilities that are present in equation (2.1) to be an adequate excluded variable. Specifically, these two other transmission probabilities included in equation (2.1) are "the subjective probability that a hypothetical healthy individual of the respondent's gender and village would become infected with HIV in the next 12 months if married to an HIV-positive individual," Π_i, and "the subjective probability that a hypothetical healthy individual of the respondent's gender and village would become infected with HIV in the next 12 months if has several sexual partners in addition to spouse," $p_i(1)$. First, we note that there is considerable variation across the three subjective probabilities. For example, the sample average transmission probability with normal sexual behavior is 0.22, while, as shown in Table 3, it is 0.77 for the transmission probability with multiple partners and 0.93 for the probability if married to a HIV-positive spouse. Second, we re-estimate the system of equations using all three transmission probabilities in equation (2.4). The precision of the excluded variable "having received a positive financial incentive" is unaffected, while the one of the transmission risk with normal sexual behavior diminishes somewhat (p-value=0.132). We also evaluated the robustness of the precision of the excluded variables in this augmented equation (2.4) to other functional forms assumptions. We find that the two excluded variables are still precisely estimated, and with similar magnitude in terms of marginal effect, when using a linear regression or a generalized linear model with a logit link and binomial distribution, sometimes called fractional logit, developed by Papke and Wooldridge (1996) –estimated outside of the system.

6.2. Heterogeneity in preferences

In the previous section, we allowed heterogeneity in preferences for having multiple partners but not in the parameters U^{-} and c. We now relax this and first allow the parameters to vary by gender. Because men and women do report quite different levels of multiple partnerships, we also allow men and women to have different preferences for having multiple partners by interacting marital status, lagged sexual behavior, and age with female. The motivation for those interactions is that the psychological or social cost for having extra-marital affairs may differ for married men and married women; the level of lagged sexual behavior is quite different by gender, and women's peak of HIV infection is at an earlier age than men's, suggesting different sexual behavior by age for men and women. We have also experimented interacting our excluded variables in the other equations of the system (2) by gender but never found a statistically significant coefficient for the interacted terms at conventional level. The other equations are therefore unchanged compared to table 5. Because of this, we only present the estimation results for equation (2.1).

The first column of Table 6 shows the average marginal effect of P_u and P_c for males and females respectively. For ease of interpretation of the results, we present the average marginal effect of P_u for males and the average marginal effect for P_u for females (rather than showing the main marginal effect and the marginal effect for the interaction with the female dummy). We do the same for the estimates of the marginal effect of P_c. The average marginal effects of the male and female P_u are positive and statistically significant at 10%. While the estimate for females is twice as large as the one for males, we cannot reject the hypothesis that they are equal (P-value=0.463). As in Table 5, the average marginal effect for c are small in magnitude and not statistically significantly different from zero. None of the interactions with female are statistically significantly different from zero at conventional levels (Table A2), suggesting that the effect of marital status, age and lagged sexual behavior on having multiple sexual partners is identical for men and women.[25]

Utility in 10 years may be valued differently by different age groups. We therefore now allow for additional heterogeneity in preferences by having the parameters U^{-} and c differing by

[25] Note that there are very few women aged 50+ who have multiple sexual partners, so we cannot in practice identify the interaction female with age above 50. We therefore just have a dummy for female aged 40+.

both gender and age. We maintain the interaction of marital status, age and lagged sexual behavior with female, as in the previous specification. The second column of Table 6 shows the average marginal effect of P_u and P_c for males and females above and below 30 years old, respectively. The average marginal effect of P_u is positive and statistically significant at 5% for females who are less than 29 and males who are above 30, suggesting that survival expectations are more relevant for younger women and older men in the decision to engage in risky sex. Note that these are the age groups at which HIV prevalence tends to peak in Malawi and other SSA countries. We do reject the equality of the estimates of $U^-$for the four age-group/gender categories at 10% (P-value=0.055). As earlier, the marginal effects of P_c are imprecisely estimated.

Future utility may also vary depending on individual wealth. We investigate this by allowing the parameters U^- and c to differ by gender and wealth. We proxy wealth by land ownership. We define high wealth as the highest land ownership tertile. The last column of Table 6 shows that average marginal effects of P_u are more precisely estimated for high wealth females and males (P-value=0.058 and 0.115 respectively), which is consistent with the idea that wealthier individuals may value future utility more. We however cannot reject the equality of the 4 estimates of U^- by wealth (P-value=0.673). Again, the marginal effects of P_c are imprecisely estimated.

6.3. Robustness checks

6.3.1. HIV testing between 2006 and 2008

So far, we have assumed that the only way to get tested for HIV was through the MLSFH survey. However, HIV testing is becoming more common in Malawi, and some respondents reported that they had been tested between the 2006 and the 2008 waves. Overall, 16.5% of respondents reported in 2008 that they had been tested for HIV between the 2006 interview and December 2007, and this proportion is 12.1% for respondents who did not learn their HIV status as part of the 2006 MLSFH HIV testing. We may therefore have measurement error in beliefs about current infection for those who got tested outside of the MLSFH. To deal with this, we assume that the HIV status that they learned in between the 2006 MLSFH testing and December

2007 is the same as the results of the 2008 MLSFH HIV test.[26]

We re-estimate the system of equation (2) based on those assumptions for the case of homogenous U^{-} and c and when they vary by gender and age. The first two columns of Table A3 presents the results of the equation with having multiple sexual partners as the dependent variable and shows that the results are very similar to those of Tables 5 and 6 (column 2).

6.3.2. Misreporting of sexual behavior

Table 2 shows a large difference in reported sexual behavior by gender. This difference by gender is typical of many surveys done in SSA. Dinkelman and Lam (2009) for example indicate that in 9 recent African Demographic and Health Surveys, men report between 10% and 80% more sexual partners than women. They point out that in a closed population without misreporting and with everyone sampled, the average number of partners of men and women should balance. In our analytical sample, the average number of partners in the last 12 months is 1.41 and 1.11 for men and women respectively. However, Dinkelman and Lam (2009) also highlight that a disequilibrium of partner reports may occur without misreporting when there is undersampling of sex workers. Yet, because sexual behavior is a sensitive topic, misreporting is a legitimate concern. To evaluate the robustness of our results to misreporting, we follow Hausman et al. (1998), who correct for misclassification and estimate its prevalence. A similar measurement error strategy was introduced in this context by de Paula et al. (2014). We assume that individuals report truthfully when they do not have multiple partners and misreport about having multiple partners. This probability of misreporting is assumed to depend on the true sexual behavior value $\tilde{a}_i$, and on other observable characteristics z_i. In particular, it is given by:

$$\alpha_1(z_i) = P(a_i = 0|\tilde{a}_i = 1, z_i)$$

With this misreporting probability, we have:

[26] This is accurate for those who tested negative in 2008, but may potentially be a strong assumption for those who tested positive in 2008, as some may have sero-converted after their latest test. But HIV-incidence in the MLSFH and other population-based studies in generally low (in the MLSFH, incidence during 2004—08 was .63 per 100 person years; Kohler et al, 2014), and thus sero-conversions during a 1-2 year period are relatively rare. Among those who were sexually active in 2008 and did learn their HIV status in 2006 and report being tested after the 2006 interview, 4.9% tested positive in 2008 and 9.4% refused to get tested. For the latter group (12 observations), we use the 2006 elicited beliefs about current infected status as the beliefs used for decision-making. Among those who got tested as part of the 2006 MLSFH and got re-tested before December 2007, less than 1% (corresponding to two respondents) changed HIV status and tested positive in the 2008 MLSFH.

$$E(a_i|x_{i1}, P_u, P_c, z_i) = \big(1 - \alpha_1(z_i)\big)\Phi(\beta_1 x_{i1} + U^- P_u - cP_c)\,(3)$$

Identification requires that there exists a covariate that affects the true sexual behavior but does not affect the probability of misclassification (Lewbel, 2000). We assume that the probability of misclassification depends on gender, age, marital status, education, religion and wealth only and that the excluded variables are the subjective beliefs. We estimate equation (3), a simple probit regression allowing for misreporting of sexual behavior in Table A4.[27] With misreporting and homogenous U^- and c, the average marginal effect of P_u is 0.146 (therefore slightly larger than the one reported in Table 5, suggesting that misreporting leads to a downward bias) and is statistically significant at 5% (column 1). When we allow U^- to differ by gender and age (column 3), we also find that the marginal effects are of larger in magnitude than those of Table 6.

Columns (2) and (4) of Table A4 show that female, less educated and wealthier respondents are more likely to misreport their sexual behavior. Also, Indigenous Christian /African Independent Churches are more likely to misreport than Catholic. We compute the predicted probability of misreporting using coefficients from Column (4) and find that females are more likely to misreport than men: the median misreporting probability is 24.8% and 3.9% respectively.

6.3.3. Non-sexually active respondents

As explained in Section 5.1, we use the 2008 MLSFH respondents who have been sexually active in the 12 months prior to the survey for our analysis of risky sex. 13.1% of the 2008 respondents who answered the 2006 expectations questions had no sex in the past 12 months.[28] We investigate in columns 3 and 4 of Table A3 the robustness of our results when we

---

[27] We estimate a simple probit rather than the whole system of equation (2). This is to simplify the estimation procedure and is motivated by the fact that the covariance between the random term ε_1 of the equation estimating the propensity to have multiple sex partners and the other equations is not statistically significantly different from zero in Table 5.

[28] Of those, 34% are married, but we do not find any statistically significant difference in their subjective probability that their spouse is infected with HIV. This suggests that abstinence to avoid infection does not seem to be playing a critical role. Among those who have not had sex in the past 12 months, 52% are separated, divorced or widow and may therefore have other psychological consideration for abstinence beyond preventing HIV infection. Among the never married, 47% have never had sex and the decision of sexual activity initiation may be different from the one we focus on.

include non-sexually active respondents. We assume that their decision is between having zero partner and multiple partners. The subjective probability of becoming infected with HIV conditional on having zero partners is assumed to be zero for those respondents. All other subjective probabilities are defined as in Section 5.2. Again, results are very similar to those of Tables 5 and 6 (column 2).

6.3.4. Attrition between 2006 and 2008

As mentioned in Section 5.1, 23% of the 2006 respondents could not be resurveyed in 2008, so their behavior cannot be used to make inferences on U_i^- and c_i. We investigate whether the probability of attrition between 2006 and 2008 is associated with the relevant 2006 expectations and 2006 sexual behavior (Table A5). We find that past sexual behavior, expectations about transmission risk and survival expectations are not correlated with the probability of attrition. However, respondents who report a higher 2006 subjective probability of being infected with HIV are more likely to attrit. In order to evaluate whether this has an impact on our coefficients of interest, we augment the system of equation (2) with the following probit attrition equation:

$$\begin{aligned} Stay0608_i^* &= \beta_5 x_{i5} + \gamma_5 f_i^0 + \varepsilon_5 \\ Stay0608_i &= 1 \; if \; Stay0608_i^* > 0 \\ Stay0608_i &= 0 \text{ otherwise} \end{aligned} \qquad (2.5)$$

Where $Stay0608 = 1$ if the respondent did not attrit between 2006 and 2008. Again, we allow ε_5 to be correlated with the random terms $\varepsilon_j, j = 1, \ldots, 4$. Identification requires therefore at least one variable in x_{i4} not included in x_{i3}, x_{i2}, x_{i1} and x_{i5}. We will use the same exclusion restrictions as those discussed in Section 5.3. Equation (2.5) can be seen as a selection equation for the sexual behavior equation (2.1). In order not to rely solely on functional form assumption, the selection equation (2.5) should have at least one variable that is not in the probit equation for having multiple sexual partner (2.1), i.e. we need a variable in x_{i5} not in x_{i1}. We use the 2008 month of the first attempted visit by the interviewers' team as this exogenous variable. Interviews took place between May and August 2008. The team of interviewers is more likely to find a respondent in earlier than later during the field period. The reason is likely to be that May and June are harvest months so respondents tend to be close to their home. By August fewer respondents are tending to crops and they may therefore be more likely to migrate temporarily.[29]

We present the estimation results for all 5 equations in Table A6. Our main results are robust to the consideration of attrition: the marginal effect of P_u is very similar to those of Table 5 and precisely estimated. Moreover, the 2008 month of first attempted interview is a strong predictor of attrition. In addition, our excluded variables (having received a positive financial incentive and the belief of becoming infected with HIV in the next 12 months with normal sexual behavior) are still strong predictor of the 2006 pre-HIV test subjective probability of being currently infected with HIV. We obtain results similar to those of Table 6, column 2, if we allow U^- and c to vary by gender and age (not shown).

6.3.5. Non-updating of subjective probability of being infected with HIV after testing

Our model assumes that respondents who got tested for HIV in 2006 revise their beliefs fully based on the test result. There is however no conclusive empirical evidence on how people revise their beliefs just after obtaining their HIV test results. In Delavande and Kohler (2012), we find that only 10% of the respondents who were told they were HIV-positive in 2004 provided 10 beans in response to the question about their subjective probability of being infected with HIV in 2006 (see further discussion in footnote 16). If those who are told they are HIV-positive do not revise their beliefs upward to a probability of one, we may mistakenly exclude them for identifying the parameters U^- and c. We may also wrongly assume that some respondents who tested negative revised their beliefs to a probability of zero. It is unclear how this may bias our results. As a robustness check, we re-estimate the system of equations (2) under the assumption that individuals do not revise at all their beliefs based on their HIV test results. Under the assumption that individuals' posterior subjective probability of being infected with HIV is in the interval made by their prior probability and the actual test result (i.e., ${f_i}^1$ belongs to $[{f_i}^0, 1]$ for

[29] The month of interview is correlated with the region but we control for region in all our regressions. One concern is whether this pattern is due to how remote the village where respondent live is, which could also influence sexual behavior. In several sites, the interviewers start with the more remote villages and work their way closer to the main roads, from which people are more likely to migrate than the more rural areas. As a robustness check, we also included in equation (2.5) a variable equal to the difference between the date of the first attempted visit and the date of the beginning of fieldwork for that region. The coefficient of this time difference is negative and statistically significant at 5%. However, it does not change the month effects we find, nor the estimate of U^- and c.

those who tested positive and $[0, f_i{}^0]$ for those who tested negative), the "true" estimate for U^- and c will be between the bounds created by the estimates based on the full-updating assumption (table 5) and the ones based on the no-updating assumption (presented in table A3, column 5). Table A3 shows that, under the no-updating assumption, the marginal effect of P_u is still precisely estimated, though slightly larger in magnitude. The resulting bounds for marginal effect of P_u are tight and equal to [0.097, 0.115]. The marginal effect of P_c is again statistically not significantly different from zero. We find similarly tight bounds if we allow U^- and c to vary by age and gender (Table A3, column 6).

6.3.6. Potential endogeneity of transmission and survival expectations

Equation (2.1) uses two beliefs as independent variables: $P_u = (1 - f_i^1)\big(p_i(1) - p_i(0)\big)(S_i^+ - S_i^-)$ and $P_c = (1 - f_i^1)\big(p_i(1) - p_i(0)\big)S_i^+$. Our estimation strategy so far has dealt with the potential endogeneity of the subjective probability of being infected with HIV, f_i^1, arising from the dependence of current beliefs on past behaviors (lagged sexual behavior or HIV testing). Another concern might be that unobservable characteristics also influence the relative subjective HIV transmission risks, $p_i(1) - p_i(0)$, as well as the decision to engage in risky sex. For example, the relative subjective HIV transmission risk, $p_i(1) - p_i(0)$ depends on beliefs about the main partner's HIV status and may also depend on other behaviors, such as condom use or frequency of intercourse, which may be related to unobservable characteristics that also influence the decision of having multiple partners. This concern is importantly mitigated by the fact that the subjective beliefs about the chance of becoming infected if married to someone is infected with HIV/AIDS or if having multiple partners in addition to spouses are asked about hypothetical individuals of the respondent's age and gender (see Appendix B). Similarly, one could worry that unobservable characteristics also influence the subjective survival expectations, S_i^+ and $(S_i^+ - S_i^-)$, and the decision to have multiple sex partners. Again, this concern is mitigated by the fact that those expectations are asked about hypothetical individuals, (see Appendix B), though this could add measurement error. In order to address this potential influence of unobservable characteristics, we augment the system of equations (2) with the following two equations:

$$S_i^+ = \beta_6 x_{i6} + \varepsilon_{i6} \qquad (2.6)$$
$$\big(p_i(1) - p_i(0)\big) = \beta_7 x_{i7} + \varepsilon_{i7} \qquad (2.7)$$

We continue to allow the random terms of all equations to be correlated.[30] Similar to the earlier discussion, identification of the new system requires in addition one variable in x_{i6} and x_{i7} not included x_{i1}. We use the subjective probability that a baby born in the respondents' community will die within one year, which is likely to be influenced by the respondents' perception of the burden of disease but should have no direct effect on the propensity to have multiple partners, as an excluded variable in the equation modeling elicited 10-year survival expectations conditional on being HIV-positive.[31] We use as excluded variables in x_{i8} whether the first attempted visit of the team of nurses in 2006 was within one week of the first attempted visit of the survey team in 2006, whether the first attempted spouse's visit of the team of nurses in 2006 was within one week of the first attempted spouse's visit of the survey team in 2006 and their interaction. Those create exogenous variation in the difference in transmission rates $p_i(1) - p_i(0)$ by influencing whether the respondent and his/her main partner have been tested for HIV in 2006, and therefore the respondent's beliefs about own and the main partner's HIV status.

Table A7 reports a subset of the average marginal effects for all the equations of this augmented system. We note three important facts: (i) the marginal effect of P_u is positive, statistically significant at 5% as before and of similar magnitude as in Table 5 while the marginal effect of P_c is still imprecisely estimated; (ii) the variables from x_{i6} and x_{i7}, excluded in x_{i1} are statistically significant at at least 10%, (iii) none of the estimated correlation between ε_{i1} and the random terms of the remaining 6 equations is statistically significantly different from zero at conventional level. This is consistent again with the lack of endogeneity of the HIV-related expectations in this context.

[30] Note that this set-up does not address the potential endogeneity of the difference in survival expectation for someone healthy and someone infected with HIV, $(S_i^+ - S_i^-)$. A main source of potential endogeneity is an unobserved trait η_i that promotes health investment. A higher η_i could be associated with a higher survival expectations and a lower propensity for risky sex. If this trait is additive and influences both survivals expectations in the same way, it cancels out in the when looking at the difference $(S_i^+ - S_i^-)$. The fact that we do not find a statistically significant correlation between ε_{i6} and ε_{i1} in Table A7 suggests that the endogeneity of survival expectations is not an empirically important issue.

[31] We do not find the child mortality expectations to be correlated with the propensity to have multiple partners. The average expectations is 2.45 for those who had one sexual partner and 2.89 for those who had multiple partners (P-value for t-test of equality is 0.378).

7. Simulation of hypothetical policies

Our results show that individuals consider their subjective beliefs about HIV status, transmission risk and survival when deciding to have multiple sexual partners. We now investigate how sexual behavior would change in response to various policy experiments that would change beliefs. Because we do not know how respondents would revise their beliefs when provided with new information, the simulation results can be seen as an upper bound of the behavioral change. Moreover, it is likely that the policies we envisioned, by changing behavior, would change transmission risk in the long run, and therefore perceived transmission risk as well. The counterfactuals we conduct do not take this possibility into consideration.

The first two policies we consider are information campaigns. For those, we explore two alternative outcomes. First, we assume that the information campaign is fully successful at educating people who fully revise their beliefs by aligning them to the information provided (fully updated beliefs scenario). Second, we assume that individuals take into account both their prior beliefs and the information provided by the campaign to revise their beliefs, and that the resulting beliefs are a simple weighted average of the two (partially updated beliefs scenario).[32] We consider the four following policies:

(i) *Information campaign on mortality risk*. We assume that individuals would be provided the life table estimates of males and females uninfected with HIV and of males and females infected with HIV. For example, for the fully updated belief scenario, we set the subjective probability of survival of a healthy individual equal to the appropriate probability from the UN life table estimates for Malawi without AIDS (United Nations, 2008), and we set the survival probability of an individual infected with HIV to those provided by Todd et al. (2007). Todd et al. (2007) measure survival since sero-conversion based on 4 East African population cohort studies before the availability of ART (two studies in Uganda, one in Tanzania, and one in Rwanda). See Table 4 for those mortality rates.

(ii) *Information campaign on transmission risk*. We assume that individuals would be provided accurate information about transmission risk. For example, for the fully updated beliefs

[32] It is unclear how individuals would process this information. Delavande (2008b) shows that educated women in the U.S. exhibit considerable heterogeneity in their revision of beliefs when provided with statistical information. For simplification, we just therefore consider the weighted average.

scenario, we set the probability of becoming infected with HIV within one year if married to someone who is HIV-positive to 5.17% for men and 10.55% for women (Carpenter et al. 1999), and we set the probability of becoming infected within one year if one has multiple partners to 0.38% for men and 2.08% for women, which is half the two-year sero-conversion rate that we observe in the MLSFH among respondents who had multiple sex partners and when pooling data from 2004, 2006 and 2008 (see Panel B of Table 4).[33] Finally, we also set the village prevalence equal to the MLSFH regional prevalence (this is relevant for respondents who did not report beliefs about a main partner's HIV status).

(iii) *Extension of ART to all infected people*. Respondents were asked the mortality risk for an individual who is sick with AIDS and an individual who is sick with AIDS and on ART.[34] Panel A of Table 4 shows that, on average, people believe that being on ART will give an individual sick with AIDS the same odd of survival as someone infected with HIV. For our simulation, we would like to know respondents' belief about the survival for an individual infected with HIV and on ART, a belief we did not elicit. We assume that respondents believe that ART will be as protective, in relative term, for an HIV-positive individual as for someone infected with AIDS, i.e., $\frac{S_{HIV}^{ART}}{S_{HIV}^{NO\ ART}} = \frac{S_{AIDS}^{ART}}{S_{AIDS}^{NO\ ART}}$, where S_h^a is the survival of an individual infected with $h, \{h = HIV, AIDS\}$, and with treatment status $t, \{t = ART, NO\ ART\}$. Under this assumption, the average 10 year probability of survival for someone infected with HIV and on ART is 20.25% (compared to an average of 11.78% for someone infected with HIV and of 40.95% for someone healthy). For this policy experiment, we therefore replace the survival probability of an individual infected with HIV with the individual-specific subjective survival probability for an individual infected with HIV who is on ART. We focus on the effect of such a campaign on individuals who know they are HIV-negative or who have not been tested for HIV.

(iv) *Extension of ART to all infected people and information of the effect of ART on transmission risk*. Cohen et al. (2013) report that being on ART reduces the transmission risky by 96%. We now assume that respondents' beliefs about survival are like in point (iii) and

[33] We assume that HIV incidence was the same among those who refused to get tested and those who agreed.

[34] While ART is now becoming more prevalent in Malawi, it was essentially not available in the MLSFH study regions until shortly before the 2008 MLSFH survey.

further assume that respondents are aware that being on ART reduces transmission risk by 96%, i.e. following the notation of Section 5.2, we replace for each respondent the subjective probability of infection associated with having multiple partners $p_i(1)$ by $0.04p_i(1)$ and the perceived probability of becoming infected with HIV during the next 12 months for someone who is married to an HIV-positive individual Π_i by $0.04\Pi_i$. This dramatically reduces the difference in transmission risk between having multiple partners versus one partner $\left(p_i(1) - p_i(0)\right)$, which now becomes 0.028 on average (compared to a baseline average of 0.691).

We compute the predicted probabilities of having multiple sex partners using the second column of Table 6, in which U^- and c vary by gender and age. Table 7 shows the mean and the 25th, 50th and 75th percentiles of the predicted probabilities of having multiple partners, along with the proportion of individuals for whom the predicted probability increases, decreases or remain the same, for all the policy scenarios.

An information campaign on mortality that would lead to full revision of beliefs would be beneficial and decrease the average predicted probability of having multiple partners from 21.94% to 15.55% for men and from 2.59% to 2.17% for women. Respondents are very pessimistic regarding their survival conditional on being healthy and conditional on being infected with HIV, compared with available statistics (see Table 4). Note however that the difference in perceived survival probabilities with and without HIV infection $(S_i^+ - S_i^-)$ is key to decision-making. Despite being pessimistic regarding their survival rates, respondents on average underestimate the impact of HIV on survival. The average $(S_i^+ - S_i^-)$ based on subjective beliefs is -0.29 in the sample (see figure 2 for the whole distribution), compared to an average of -0.47 based on available statistics. This is why providing information on mortality risk reduces risky sexual behavior on average. When looking at distributional impact, Table 7 shows overwhelmingly positive impact for men: 93% would reduce their risky behavior following this information campaign, while only 3% would increase it. About 4% would not change their behavior: those with inelastic behavior to information on mortality risk are individuals who believe/know they are HIV-positive or those for whom the difference in transmission risk is zero. For women, 63% (30%) would reduce (increase) their risky sexual behavior.

The effect would still be beneficial, though smaller, if individuals only partially update beliefs in light of the new information. Under the assumption that revised beliefs would be the weighted average of prior beliefs and the provided information, the average predicted probability of having multiple partners would decrease to 18.40% for men and to 2.29% for women.

Table 7 shows that an information campaign on transmission risk would actually have an undesirable effect: the average predicted probability of having multiple partners would actually increase to 26.93% for men and 3.75% for women under this policy and the fully updated beliefs scenario. This is because respondents over-estimate the relative impact that having multiple partners has on the probability of becoming infected with HIV. Respondents are overall very pessimistic regarding HIV transmission risks. While the yearly incidence in sero-discordant couples is estimated to be 5.17% for men and 10.55% for women (Carpenter et al., 1999), the average subjective probability of becoming infected with HIV conditional on being married to an HIV-positive spouse is 93.9% for men and 92.2% for women (Panel B of Table 4). Similarly, while the probability of becoming infected within 1 year if one has multiple partners is 0.38% for males and 2.08% for females in the MLSFH, the average subjective probabilities are 76.3% and 78.2% respectively. What matters for decision-making is the relative subjective risk of becoming infected under these conditions, i.e., $p_i(1) - \Pi_i \times {f_{pi}}^1$. This average subjective difference in risk is 0.709 for men and 0.680 for women. If we use statistics from existing studies (but still use the beliefs about the partner's HIV status ${f_{pi}}^1$), the difference in risk is much smaller: 0.004 for men and 0.100 for women. This explains why providing information on transmission risk actually increases risky behavior on average. Under this scenario, most of the men (89%) and women (69%) would have a higher probability of having multiple partners. Under the partially update beliefs scenario, risky behavior would also increase, though less than under the fully updated beliefs scenario.[35]

Finally, we consider the effect on sexual behavior of extending ART to all infected people. As the survey shows, people are aware that being on ART decreases the "cost" of becoming infected with HIV by increasing survival probability. As a result, we find an increase in risky behavior on average, though the effect is rather small, when taking into consideration

[35] In order to evaluate how elastic the sexual behavior is to change in beliefs, we also compute the average predicted probabilities of having multiple partners in the "best case scenario" regarding the beliefs, i.e. under the assumption that the probability of being infected is zero, difference in transmission risk is one, the probability of survival conditional on being healthy is one and conditional on being HIV+ is zero. The average predicted probability is 13.71% for men and 1.14% for women, substantially lower than under the current beliefs.

only how the roll-out of ART would influence mortality risk. Excluding individuals who know they are HIV-positive, the average predicted probability of having multiple partners increases slightly from 22.02% to 22.05% for men and from 2.43% to 3.27% for women. The small magnitude is due to the heterogeneous effect of such a policy. More than half of the respondents would not change their behavior because they do not see an improvement in 10-year survival by being on ART. Yet, about a quarter of respondents (22% of men and 28% of women) would have a higher predicted probability of having multiple partners. However, we also find that, if people become aware in addition that being on ART would dramatically reduce HIV transmission risk, risky behavior would increase more substantially. The average predicted probability of having multiple partners increases to 26.92% for men and to 3.67% for women. In addition, we find that 91% of the men and 73% of the women would increase the probability of having multiple partners.

Conclusion

Behavioral changes related to the transmission of HIV are likely to depend on the information and knowledge of individuals, in particular their expectations about their HIV status and that of their partners, survival risks, and transmission risks associated with behaviors such as having multiple sexual partners or not using condoms. Yet, little is known about health-related subjective expectations in high-HIV-prevalence environments and how they influence decision-making related to the spread of the disease. In this paper, we fill this knowledge gap by using rich data on probabilistic beliefs elicited directly from rural Malawi survey respondents to investigate the role of HIV/AIDS-related expectations on the decision to engage in risky sex. We deal with the potential endogeneity of expectations by estimating a system of equations. We find that the difference in probability of survival associated with having multiple partners versus having one partner, which in turn depends on a large set of expectations all observed in our data, plays an important role in the decision to have multiple sexual partners. Moreover, our results suggest that the threat of reduced survival associated with HIV, rather than the threat of living with HIV, influences sexual behavior in the SSA context.

Using our estimated preference parameters, we simulate the impact of policies that would influence expectations. Our results suggest that information campaigns focused on disease transmission are likely to have limited impact on behavior. Actually, providing information on transmission risk may have a perverse effect and increase the likelihood of risky sex. Rather, we suggest a new focus for information campaign that would decrease the prevalence of risky behavior: survival rates for healthy and infected individuals or on relative survival rates.

We also find that the expansion of ART to all individuals infected with HIV would increase risky sexual behavior for a quarter of the HIV-negative respondents (or those who have not been tested). The effect would be much larger if people also become aware that ART reduces transmission risk dramatically. This suggests that expansion of ART should not be done in isolation, but rather combined with behavioral interventions to mitigate the effects of the roll-out of ART on HIV-negative individuals.

References

Anglewicz, P., Adams J., Obare-Onyango, F., Kohler, H-P. and Watkins, S. (2009). "The Malawi Diffusion and Ideational Change Project 2004–06: Data collection, data quality and analyses of attrition." *Demographic Research*, 20(21), 503-540.

Arcidiacono, P., Hotz, J. and Kang, S. (2012). "Modeling College Major Choice using Elicited Measures of Expectations and Counterfactuals," *Journal of Econometrics*, 166(1), 3-16.

Bellemare, C., Kröger, S. and van Soest, A. (2008). "Measuring Inequity Aversion in a Heterogeneous Population Using Experimental Decisions and Subjective Probabilities," *Econometrica*, 76(4), 815-839.

Boozer, M. A. and Philipson, T. (2000). "The Impact of Public Testing for Human Immunodeficiency Virus." *Journal of Human Resources*, 35, 419–46.

Bor, J., Herbst, A., Newell, M-L., Bärnighausen1, T. (2013). "Increases in Adult Life Expectancy in Rural South Africa: Valuing the Scale-Up of HIV Treatment," *Science*, 339(6122), 961-965.

Carpenter, L.M., Kamali, A., Ruberantwari, A., Malamba, S.S. and Whitworth, J.A. (1999). "Rates of HIV-1 Transmission within Marriage in Rural Uganda in Relation to the HIV Sero-status of the Partners." *AIDS*, 13(9): 1083–1089.

Chimbiri, A. (2007). "The condom is an intruder in marriage: evidence from rural Malawi." *Social Science and Medicine*, 64, 1102-1115.

Cohen, J. (2011). "HIV Treatment as Prevention." *Science*, 334(6063), 1628.

Corno, L. and A. de Paula. (2014). "Measuring risky sexual behavior: biological markers and self-reported data." Working Paper, University of College London.

Delavande, A. (2008a). "Pill, Patch or Shot? Subjective Expectations and Birth Control Choice," *International Economic Review*, 49(3).

Delavande, A. (2008b). "Measuring Revisions to Subjective Expectations," *Journal of Risk and Uncertainty*, 36(1).

Delavande, A. (2014). "Probabilistic Expectations in Developing Countries," *Annual Review of Economics*.

Delavande, A. and Kohler, H-P. (2009). "Subjective Expectations in the Context of HIV/AIDS in Malawi." *Demographic Research*, 20(31).

Delavande, A. and Kohler, H-P. (2012). "The impact of HIV Testing on Subjective Expectations and Risky Behavior in Malawi," *Demography*, 49(3):1011-36.

Delavande, A., Xavier, G. and McKenzie, D. (2011). "Measuring Subjective Expectations in Developing Countries: A Critical Review and New Evidence" *Journal of Development Economics*, 94, 151–163

Delavande, A. and Rohwedder, S. (2011). "Individuals' Uncertainty about Future Social Security Benefits and Portfolio Choice," *Journal of Applied Econometrics*, 26(3), 498–519.

Dinkelman, T. and Lam, D. (2009). "A Model for Understanding Gender Discrepancy in Sexual Behavior Report." Population Studies Center, University of Michigan, Working Paper 09-669.

Fishbein, M. and W. Pequegnat. (2000). "Evaluating AIDS prevention interventions using behavioral and biological outcome measures." *Sexually Transmitted Diseases*, 27(2): 101–110.

Floyd, S.; Marston, M.; Baisley, K.; Wringe, A.; Herbst, K. et al (2012). "The effect of antiretroviral therapy provision on all-cause, AIDS and non-AIDS mortality at the population level---a comparative analysis of data from four settings in Southern and East Africa." *Tropical Medicine & International Health*, 17(8), e84-e93

Gong, E. (2015). "HIV Testing and Risky Behavior." *The Economic Journal*, 125(582): 32–60.

Greenwood J., Kircher, P., Santos, C. , and Tertilt, M. (2013). "An Equilibrium Model of the African HIV/AIDS Epidemic." Unpublished manuscript, University of Mannheim.

Hausman, J.A., Abrevaya, J. and Scott-Morton, F.M. (1998). "Misclassification of the dependent variable in a discrete-response setting," *Journal of Econometrics*, 87 239-269.

Hurd, M. (2009). "Subjective Probabilities in Household Surveys," *Annual Review of Economics*, 1, 543-564.

Kezdi, G. and Willis, R. (2011). "Household stock market beliefs and learning," NBER working paper No. 17614.

Kohler, H.-P., Watkins, S. C, Behrman, J. R., Anglewicz, P., Kohler, I. V., et al. (2014). "Cohort profile: The Malawi Longitudinal Study of Families and Health (MLSFH)." *International Journal of Epidemiology*. First published online: March 16, 2014. DOI: 10.1093/ije/dyu049. Earlier version is available as Population Studies Center, University of Pennsylvania, Working Paper #2013-06. http://repository.upenn.edu/psc_working_papers/46.

Lakdawalla, D., Sood, N. and Goldman, D. (2006). "HIV breakthroughs and risky sexual behavior." *Quarterly Journal of Economics*, 121 1063–1102.

Lammers, J., S.J.G Van Wijnbergen, D .Willebrands. (2013). "Condom use, risk perception, and HIV knowledge: a comparison across sexes in Nigeria." *HIV/AIDS - Research and Palliative Care*. DOI: 10.2147/HIV.S31687.

Lewbel, A. (2000). "Identification of the Binary Choice Model with Misclassification." *Econometric Theory*, 16, 603-609.

Lochner, L. (2007). "Individual Perceptions of the Criminal Justice System," *American Economic Review*, 97, 444-460.

Long, J.S. (1997). *Regression Models for Categorical and Limited Dependent Variables*. Sage Publishing.

Maddala, G.S. (1983). *Limited Dependent Variable and Qualitative Variables in Econometrics.* Cambridge University Press.

Manski, C. (2004). “Measuring Expectations.” *Econometrica* 72, 1329-1376.

McKenzie, D, Gibson, J. and Stillman, S. (2013). “A land of milk and honey with streets paved with gold: Do emigrants have over-optimistic expectations about incomes abroad?” *Journal of Development Economics*, 102:116–127.

Mensch, B.S., P.C. Hewett, R. Gregory, and S. Helleringer. (2008). “Sexual Behavior and STI/HIV Status among Adolescents in Rural Malawi: An Evaluation of the Effect of Interview Mode on Reporting.” *Studies in Family Planning* 39(4):321-334.

Morgan, D, C. Mahe, B. Mayanja, J.M. Okongo, R. Lubega, J. Whitworth. (2002). “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(4):597-603.

Nyarko, Y. and Schotter, A.. (2002). “An Experimental Study of Belief Learning Using Elicited Beliefs,” *Econometrica* 70(3), 971–1005.

Oster, E. (2012). “HIV and sexual behavior change: Why not Africa?” *Journal of Health Economics* 31(1):35-49.

Papke, L.E. and J.M. Wooldridge. (1996). “Econometric Methods for Fractional Response Variables with an Application to 401(k) Plan Participation Rates.” *Journal of Applied Econometrics* 11:619-632.

de Paula, A., Shapira, G. and Todd, P. (2014). “How Beliefs About HIV Status Affect Risky Behaviors: Evidence From Malawi.” *Journal of Applied Econometrics* 29(6): 944–964.

Philipson, T. J. and Posner, R. A. (1993). *Private Choices and Public Health: The AIDS Epidemic in an Economic Perspective*. Cambridge: Harvard University Press.

Roodman, D. (2011). “Fitting fully observed recursive mixed-process models with cmp.” *The Stata Journal* 11(2): 159–206.

Shapira, G. (2009). “How Subjective Beliefs about HIV Infection Affect Life-Cycle Fertility: Evidence from Rural Malawi,” University of Pennsylvania Working Paper.

Sterck, 0. (2013). “HIV/AIDS and Fatalism: Should Prevention Campaigns Disclose the Transmission Rate of HIV.” *Journal of African Economics* doi: 10.1093/jae/ejt018.

Thirumurthy, H., C. Pop-Eleches, J. Habyarimana, M. Goldstein, and J. G. Zivin. (2012). “Behavioral Responses of Patients in AIDS Treatment Programs: Sexual Behavior in Kenya.” *Forum for Health Economics & Policy* 15(2): 1.

Thornton, R. (2008). “The Demand for and Impact of Learning HIV Status: Evidence from a Field Experiment.” *American Economic Review*, 98, 1829–1863.

Todd, J., Glynn, J.R., Marston, M., Lutalo, T., Biraro, S., Mwita, W., Suriyanon, V., Rangsin, R., Nelson, K.E., Sonnenberg, P., Fitzgerald, D., Karita, E., Zaba, B. (2007). “Time from HIV seroconversion to death: a collaborative analysis of eight studies in six low and middle-income countries before highly active antiretroviral therapy,” *AIDS*, 21 Suppl 6:S55-63.

UNAIDS (2012). Global report: UNAIDS report on the global AIDS epidemic 2012. New York:World Health Organization and UNAIDS URL: http://www.unaids.org/globalreport/.

van der Klaauw, W. (2012). “On the Use of Expectations Data in Estimating Structural Dynamic Choice Models,” *Journal of Labor Economics* 30(3): 521 – 554.

van der Klaauw, W. and Wolpin, K. (2008). “Social security and the retirement and savings behavior of low-income households,” *Journal of Econometrics*, 145(1-2), 21-42.

van de Ven, W. and van Praag, B. (1981). “The demand for deductibles in private health insurance : A probit model with sample selection,” *Journal of Econometrics*, 17(2), 229-252, November.

Zafar, B. (2013). “College Major Choice and the Gender Gap,” *Journal of Human Resources* 48(3), 545-595.

Figure 1: Timeline of data collection

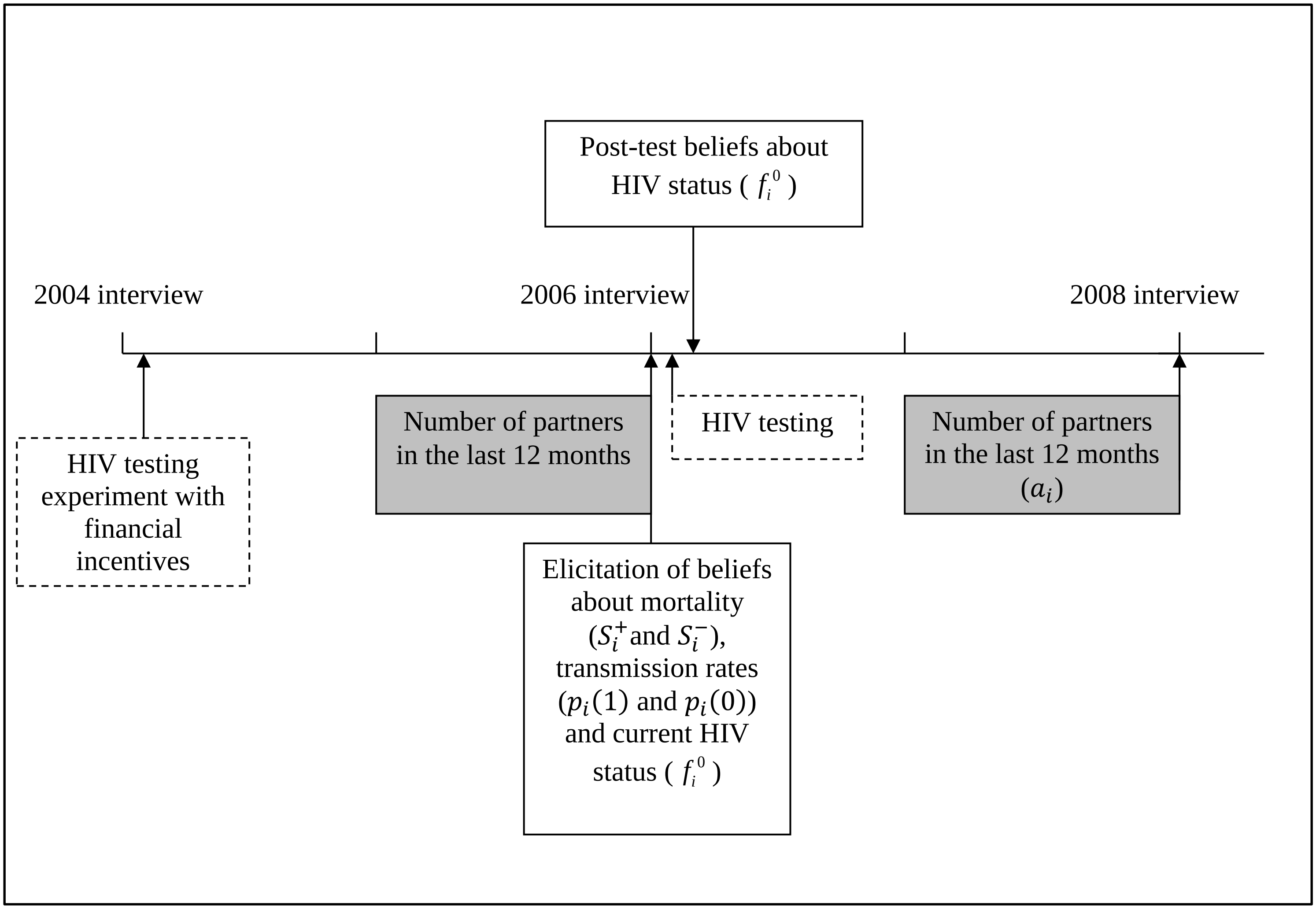


Figure 2: Distribution (in %) of the subjective probability of being infected with HIV, by HIV status (HIV test done after elicitation of probabilities)

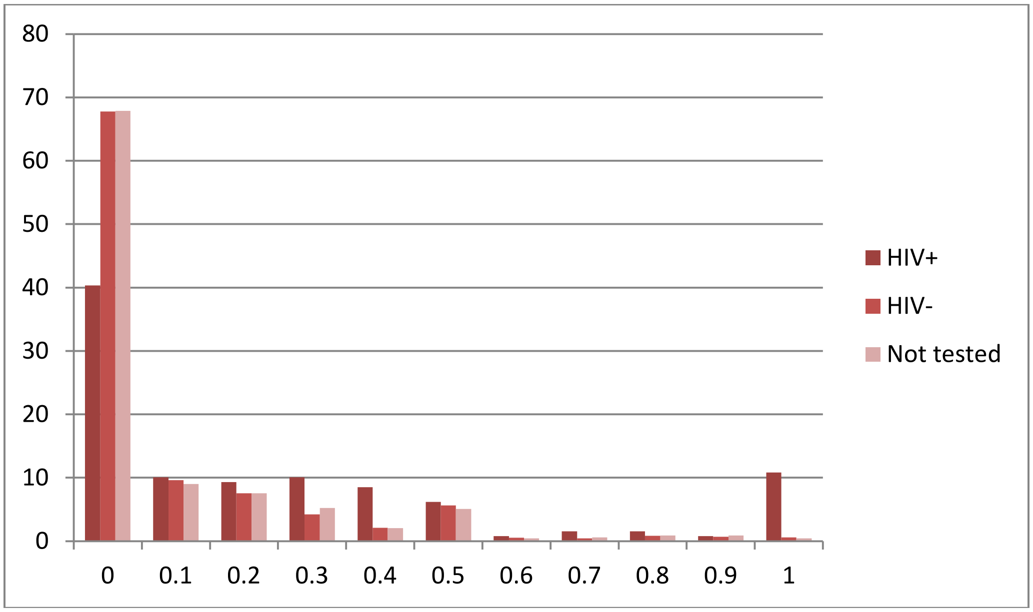
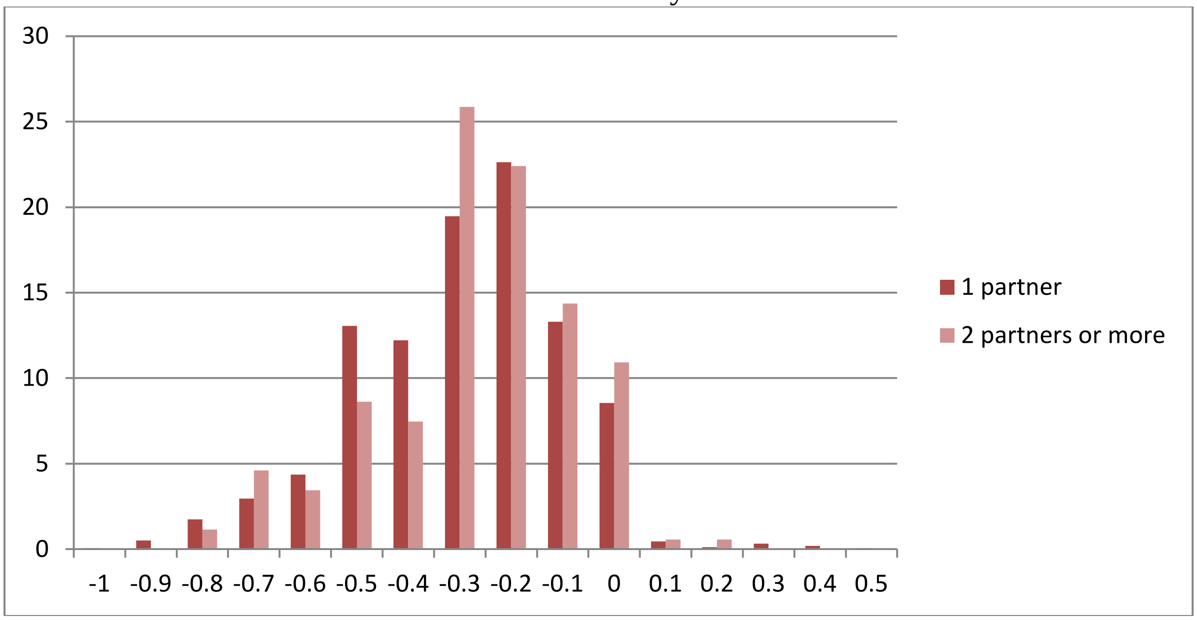


Figure 3: Distribution (in %) of beliefs by number of sexual partners

Probability of being alive in 10 years if HIV-positive minus Probability of being alive in 10 years if healthy



Probability of becoming infected if sex with multiple partners minus Probability of becoming infected if sex with one partner

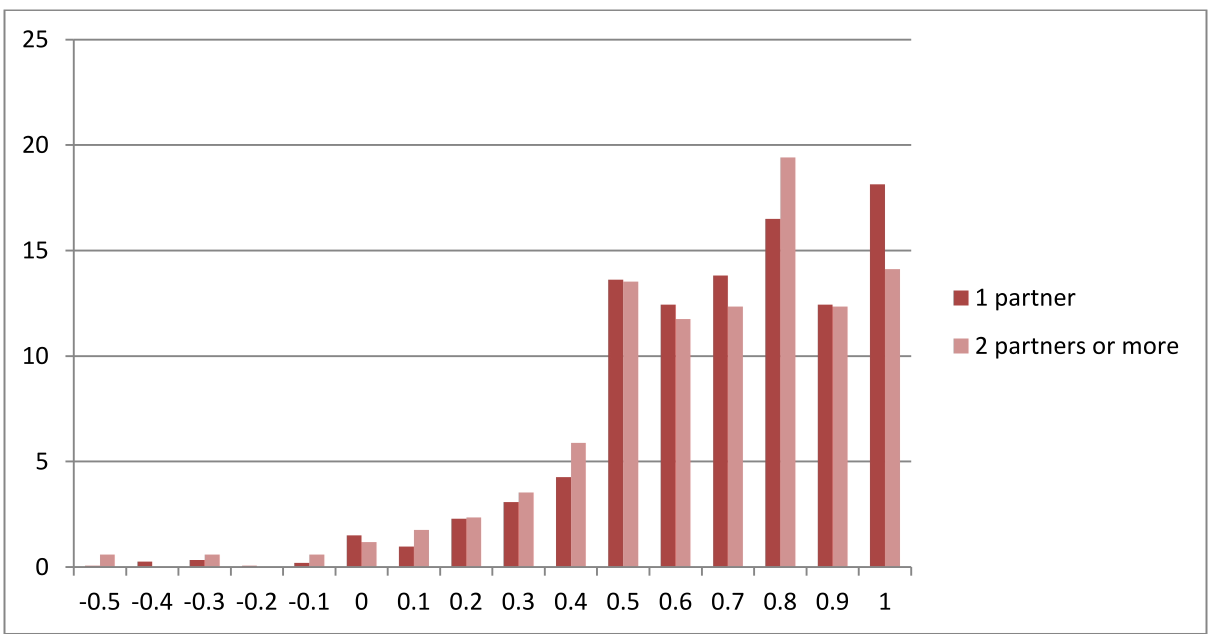


Table 1: Characteristics of 2006 respondents and 2008 respondents who answered the 2006 probabilistic expectations

	2006	2008
N	3081	2230
Female	57.7	60.5
Age		
less than 29	44.1	32.7
30 to 39	22.6	26.3
40 to 49	17.8	22.0
50+	14.9	19.1
Missing	0.6	0.0
Education		
No School	20.8	23.2
Primary level	64.8	64.3
Secondary level +	13.8	12.5
Missing	0.7	0.0
married/living together	80.0	86.7
Religion		
Catholic	17.3	16.5
Muslim	24.6	24.5
Indigenous Christian /African Independent Churches	14.9	13.6
Other Christian	36.8	33.4
Other religions	5.7	8.4
No religion	0.5	0.9
Missing	0.8	4.0
Land ownership		
2 acres of less	43.2	40.6
Between 2 acres and 4 acres	32.1	31.2
More than 4 acres	23.7	24.6
Missing	1.1	3.6
Region		
Central region	32.5	31.5
Southern region	35.2	36.2
Northern region	32.3	32.2
Number of sex partners in past 12 months		
0	14.9	13.1
1	77.0	78.2
2 or more	8.2	8.7

HIV test results		
Negative	74.7	74.0
Positive	4.0	3.7
Indeterminate	0.2	0.0
Refuse test	6.8	5.8
Not found on day of test	14.3	16.5
2004 HIV testing experiment		
Received a positive financial incentive (amont participants)	74.9	75.6
Financial incentive among participants (in Kwacha)	102.2	103.2
Did not participate in the 2004 HIV test	32.5	29.1
Probability of becoming infected with HIV within the next 12 months with normal sexual behavior for someone healthy and same gender and village as respondent	0.2	0.3
First attempted visit of the 2006 team of nurses is within one week of the first attempted visit of the survey team	13.4	
Month of first attempted 2008 interview		
May		13.6
June		22.6
July		32.9
August		22.7

Table 2: Descriptive statistics of independent variables: mean and standard deviation in parenthesis

	All	Males	Females
2006 pre-test probability of being currently infected with HIV	0.11	0.08	0.14
	(0.20)	(0.17)	(0.22)
Found by team of nurses conducting HIV test in 2006	85.96	84.24	87.23
Learned HIV status in 2006 conditional on being found	90.33	89.80	90.67
Had multiple partners in the last 12 month in 2008 conditional on being sexually active in 2008 and having non-missing beliefs	13.00	22.20	2.08

Table 3: Descriptive statistics of subjective probabilities for respondents who were sexually active in the past 12 months in 2008

	Notation in model from Section 2	Mean	SD
Subjective probability (from 0 to 1) that:			
Respondent is currently infected with HIV (post 2006 HIV test)	${f_i}^1$	0.046	0.188
Someone of respondent's age and gender who is currently healthy will survive within 10 years	${S_i}^-$	0.409	0.219
Someone of respondent's age and gender who is currently infected with HIV will survive within 10 years	${S_i}^+$	0.118	0.162
Someone healthy of respondent's gender become infected with HIV in the next 12 months if several sexual partners in addition to her spouse	$p_i(1)$	0.774	0.194
Someone healthy of respondent's gender become infected with HIV in the next 12 months if married to someone who is infected with HIV/AIDS	Π_i	0.930	0.139
Spouse/romantic partner is currently infected with HIV status (post 2006 HIV test)	${f_{pi}}^1$	0.078	0.183
One becomes infected with HIV in the next 12 months if having sex with spouse only	$p_i(0) = \Pi_i \times {f_{pi}}^1$	0.082	0.180
Beliefs relevant for decision to have multiple partners			
	$(1 - f_i^1)\big(p_i(1) - p_i(0)\big)(S_i^+ - S_i^-)$	-0.196	0.167
	$(1 - f_i^1)\big(p_i(1) - p_i(0)\big)S_i^+$	0.076	0.118

Table 4: Average probabilities in 2006 among respondents who were sexually active in 2008, with standard deviations in parentheses

Panel A	Subjective probability that hypothetical individual will be alive in 10 years by age group				Ten-year survival rates	
	Someone your age healthy	Someone your age sick infected with HIV	Someone your age sick with AIDS	Someone your age sick with AIDS and treated with ART	Malawi life tables (no AIDS scenario) †	Since year of sero-conversion observed in East Africa cohort population ††
<29	0.443	0.125	0.008	0.109	0.980	0.607
	(0.219)	*(0.172)*	*(0.040)*	*(0.150)*		
30-39	0.400	0.124	0.006	0.110	0.962	0.429
	(0.217)	*(0.159)*	*(0.028)*	*(0.151)*		
40-49	0.393	0.105	0.004	0.094	0.898	0.279
	(0.211)	*(0.150)*	*(0.028)*	*(0.137)*		
50+	0.349	0.102	0.011	0.113	0.682	0.175
	(0.349)	*(0.150)*	*(0.060)*	*(0.164)*		
*F-test****	*0.000*	*0.089*	*0.216*	*0.349*		

Panel B	Subjective probability of becoming infected in the next 12 months by 2006 HIV status and gender			One-year transmission risk		
	If several sexual partners in addition to spouse	If married to someone who is infected with HIV/AIDS	If has sex with main partner only	If several sexual partners in addition to spouse*	If married to someone who is infected with HIV/AIDS**	If has sex with one partner only*
Male				0.004	0.052	0.003
HIV-	0.769	0.938	0.051			
	(0.197)	*(0.129)*	*(0.135)*			
HIV+	0.764	0.971	0.133			
	(0.210)	*(0.073)*	*(0.281)*			
Not tested	0.739	0.938	0.058			
	(0.217)	*(0.171)*	*(0.175)*			
F-test	*0.273*	*0.674*	*0.120*			
Female				0.021	0.106	0.006
HIV-	0.783	0.924	0.093			
	(0.189)	*(0.138)*	*(0.180)*			
HIV+	0.757	0.886	0.246			
	(0.174)	*(0.165)*	*(0.345)*			
Not tested	0.787	0.922	0.110			
	(0.187)	*(0.132)*	*(0.210)*			
*F-test****	*0.641*	*0.209*	*0.000*			

† Source: United Nations (2008); †† Source: Todd et al. (2007) based on 4 East African Population cohorts; * Source: 2004, 2006 and 2008 MLSFH data; ** Source: Carpenter et al. (1999); *** P-value for F-test of equality of means

Table 5: The impact of subjective beliefs on sexual behavior (average marginal effects)

	Had multiple partners in the last 12 month in 2008 conditional on being sexually active	Found by team of nurses conducting HIV test in 2006	Learned HIV status in 2006 conditional on being found	2006 pre-test probability of being currently infected with HIV
P_u	0.097**			
	[0.042]			
P_c	-0.068			
	[0.074]			
Female	-0.149***	0.010	0.009	0.089**
	[0.019]	[0.016]	[0.012]	[0.030]
Had no partner in the last 12 months in 2006				
Had 1 partner in the last 12 months in 2006	0.033	0.003	0.033	0.009
	[0.038]	[0.023]	[0.021]	[0.048]
Had multiple partners in the last 12 months in 2006	0.171***	-0.047	0.049*	0.179***
	[0.040]	[0.031]	[0.028]	[0.062]
Less than 29 years old				
30-39	0.017	0.099***	0.004	0.037
	[0.020]	[0.019]	[0.016]	[0.032]
40-49	0.002	0.088***	0.010	-0.008
	[0.021]	[0.021]	[0.019]	[0.041]
50+	-0.011	0.085***	-0.006	-0.029
	[0.021]	[0.021]	[0.019]	[0.045]
No school in 2008				
Primary school	0.018	-0.016	0.011	-0.044
	[0.020]	[0.019]	[0.015]	[0.032]
Secondary school or more	0.038	-0.016	-0.026	0.008
	[0.029]	[0.029]	[0.025]	[0.054]
Married	-0.099***	-0.055**	-0.038**	0.032
	[0.023]	[0.020]	[0.019]	[0.038]
Catholic				
Muslim	0.038	0.002	-0.012	-0.054
	[0.026]	[0.024]	[0.023]	[0.040]
Indigenous Christian /African Independent Churches	0.029	-0.013	0.018	0.001
	[0.027]	[0.024]	[0.022]	[0.049]
Other Christian	-0.054**	0.021	0.006	-0.044
	[0.022]	[0.020]	[0.017]	[0.040]
Other or missing or no religions	-0.027	-0.013	0.073**	-0.098
	[0.033]	[0.032]	[0.037]	[0.062]
Own 2 acres or less of land				
Between 2 and 4 acres	0.020	-0.016	-0.005	-0.012
	[0.017]	[0.016]	[0.015]	[0.030]
More than 4 acres	0.008	-0.033*	-0.021	-0.025
	[0.019]	[0.018]	[0.016]	[0.035]

Continuation of Table 5

2006 pre-test probability of being currently infected with HIV		0.152***	-0.062	
		[0.054]	[0.044]	
First attempted visit of the 2006 team of nurses is within one week of the first attempted visit of the survey team		0.068***		
		[0.020]		
Receive positive financial incentive in 2004 testing experiment			0.032*	0.083**
			[0.018]	[0.042]
Did not participate in 2004 testing experiment			-0.053***	0.021
			[0.018]	[0.047]
Probability that someone healthy of respondent's gender become infected with HIV in the next 12 months with normal sexual behavior				0.129*
				[0.068]
r_44	0.271***			
	[0.020]			
r_12	-0.009			
	[0.087]			
r_13	-0.004			
	[0.101]			
r_14	-0.011			
	[0.129]			
r_23	-0.442*			
	[0.246]			
r_24	-0.421**			
	[0.128]			
r_34	0.206*			
	[0.111]			
N	2,576			

Table shows marginal effects. In column 4, we report the coefficients, i.e. the marginal effect for f_i^{0*}. Robust standard errors clustered at the couple level in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%. All regressions include regions dummies and a missing indicator for age, education and land ownership.

Table 6: The impact of subjective beliefs on sexual behavior: heterogeneity in preferences (Dependent variable is "Had multiple partners in the last 12 month in 2008 conditional on being sexually active") (Average marginal effects)

	Gender	Gender and age	Gender and wealth
P_u - female	0.176*		
	[0.107]		
P_u - male	0.090*		
	[0.046]		
P_c - female	0.024		
	[0.111]		
P_c - male	-0.106		
	[0.085]		
P_u - female less than 29 years old		0.317**	
		[0.128]	
P_u - female more than 30 years old		-0.007	
		[0.166]	
P_u - male less than 29 years old		0.019	
		[0.059]	
P_u - male more than 30 years old		0.217***	
		[0.073]	
P_c - female less than 29 years old		0.070	
		[0.127]	
P_c - female more than 30 years old		-0.105	
		[0.205]	
P_c - male less than 29 years old		-0.098	
		[0.102]	
P_c - male more than 30 years old		-0.105	
		[0.139]	
P_u - female and low wealth			0.126
			[0.156]
P_u - female and high wealth			0.271*
			[0.141]
P_u - male and low wealth			0.086
			[0.070]
P_u - male and high wealth			0.092
			[0.058]
P_c - female and low wealth			-0.242
			[0.174]
P_c - female and high wealth			0.168
			[0.128]
P_c - male and low wealth			-0.056
			[0.130]
P_c - male and high wealth			-0.135
			[0.106]
P-value for test of equality of U-	0.464	0.055	0.673
P-value for test of equality of -c	0.344	0.726	0.175

Robust standard errors clustered at the couple level in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%. See all remaining coefficients in Appendix table A2.

Table 7: Impact of policies on the predicted probabilities of having multiple partners

		Predicted probabilities			
	All sample	Fully updated beliefs		Partially updated beliefs	
stats	Baseline	Mortality risk information	Transmission risk information	Mortality risk information	Transmission risk information
Males					
25th perc.	0.0942	0.0490	0.1398	0.0700	0.1149
Median	0.1622	0.0962	0.2173	0.1215	0.1829
75th perc.	0.2747	0.2041	0.3294	0.2333	0.2901
mean	0.2194	0.1555	0.2693	0.1840	0.2419
Percentage with reduced risk		93.32	1.09	93.32	1.36
Percentage with same risk		3.54	9.69	3.54	9.69
Percentage with increased risk		3.13	89.22	3.13	88.95
Females					
25th perc.	0.0070	0.0057	0.0147	0.0063	0.0114
Median	0.0160	0.0122	0.0273	0.0136	0.0229
75th perc.	0.0321	0.0248	0.0460	0.0267	0.0477
mean	0.0259	0.0217	0.0375	0.0229	0.0565
Percentage with reduced risk		63.36	17.38	63.36	17.86
Percentage with same risk		6.20	13.94	6.2	13.94
Percentage with increased risk		30.44	68.67	30.44	68.19

	Excludes respondents who know they are HIV+		
stats	Baseline	Roll-out of ART	Roll-out of ART and reduced transm. risk
Males			
25th perc.	0.0934	0.0973	0.1397
Median	0.1618	0.1628	0.2161
75th perc.	0.2774	0.2771	0.3339
mean	0.2202	0.2205	0.2692
Percentage with reduced risk		25.42	1.55
Percentage with same risk		52.68	7.63
Percentage with increased risk		21.89	90.82
Females			
25th perc.	0.0067	0.0074	0.0145
Median	0.0151	0.0176	0.0263
75th perc.	0.0299	0.0389	0.0449
mean	0.0243	0.0327	0.0367
Percentage with reduced risk		16.54	18.12
Percentage with same risk		54.97	8.9
Percentage with increased risk		28.48	72.98

Reduced (same/increased) risk means strictly smaller (same/strictly larger) predicted probability compared to baseline.

Appendix A

Table A1: Estimation of the system of equations (2) with fixed residual correlation (Average marginal effects)

	ρ_{14}=-0.5	ρ_{14}=-0.3	ρ_{14}=-0.1	ρ_{14}=0.1	ρ_{14}=0.3	ρ_{14}=0.5
P_u	0.158***	0.130***	0.107**	0.089**	0.074*	0.063
	[0.048]	[0.045]	[0.043]	[0.041]	[0.041]	[0.041]
P_c	-0.132	-0.102	-0.078	-0.059	-0.045	-0.035
	[0.081]	[0.078]	[0.075]	[0.074]	[0.072]	[0.071]

Robust standard errors clustered at the couple level in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%. Estimates based on system of equations (2) and same controls as in Table 5.

Table A2: The impact of subjective beliefs on sexual behavior: heterogeneity in preferences – other coefficients for results presented in table 6 (Dependent variable is "Had multiple partners in the last 12 month in 2008 conditional on being sexually active") (Average marginal effects)

	Gender	Gender and age	Gender and wealth
Female	-0.072	-0.045	-0.070
	[0.073]	[0.073]	[0.073]
Had no partner in the last 12 months in 2006			
Had 1 partner in the last 12 months in 2006	0.071	0.072	0.069
	[0.047]	[0.046]	[0.047]
Had multiple partners in the last 12 months in 2006	0.210***	0.211***	0.207***
	[0.049]	[0.047]	[0.048]
Had 1 partner in the last 12 months in 2006 x female	-0.094	-0.094	-0.096
	[0.069]	[0.068]	[0.068]
Had multiple partners in the last 12 months in 2006 x female	-0.134	-0.13	-0.137
	[0.104]	[0.103]	[0.105]
Less than 29 years old in 2008			
30-39	0.022	0.024	0.023
	[0.024]	[0.024]	[0.024]
40-49	0.002	0.042	0.003
	[0.023]	[0.033]	[0.023]
50+	-0.01	0.027	-0.011
	[0.023]	[0.032]	[0.023]
30-39 x female	-0.015	-0.019	-0.016
	[0.043]	[0.043]	[0.043]
40-49 x female	0.005	-0.070	0.008
	[0.040]	[0.062]	[0.040]
No school in 2008			
Primary school	0.020	0.020	0.020
	[0.020]	[0.020]	[0.020]
Secondary school or more	0.039	0.037	0.038
	[0.029]	[0.029]	[0.029]
Married in 2008	-0.105***	-0.107***	-0.106***
	[0.027]	[0.027]	[0.027]
Married in 2008 x female	0.023	0.028	0.026
	[0.053]	[0.053]	[0.053]
Catholic			
Muslim	0.037	0.035	0.039
	[0.026]	[0.026]	[0.026]
Indigenous Christian /African Independent Churches	0.029	0.028	0.029
	[0.027]	[0.027]	[0.027]
Other Christian	-0.054**	-0.056**	-0.054**
	[0.022]	[0.022]	[0.022]

Continuation of Table A2			
Other or missing or no religions	-0.027	-0.031	-0.026
	[0.033]	[0.033]	[0.033]
Own 2 acres or less of land			
Between 2 and 4 acres	0.019	0.020	0.023
	[0.017]	[0.017]	[0.025]
More than 4 acres	0.006	0.008	0.009
	[0.019]	[0.019]	[0.026]
N		2,576	

Robust standard errors clustered at the couple level in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%. All regressions include regions dummies and a missing indicator for age, education and land ownership. Estimation based on system of equations (2).

Table A3: The impact of subjective beliefs on sexual behavior: robustness checks (Dependent variable is “Had multiple partners in the last 12 month in 2008 conditional on being sexually active”) (Average marginal effects)

	HIV testing between 2006 and 2008		Pooling sexually active and non-sexually active respondents		Non-updating of subjective probability of being infected	
P_u	0.092**		0.096***		0.115***	
	[0.042]		[0.036]		[0.043]	
P_c	-0.063		-0.034		-0.084	
	[0.073]		[0.063]		[0.074]	
P_u - female less than 29 years old		0.298**		0.274***		0.280**
		[0.121]		[0.106]		[0.126]
P_u - female more than 30 years old		-0.013		0.002		0.084
		[0.163]		[0.135]		[0.148]
P_u - male less than 29 years old		0.011		0.035		0.057
		[0.059]		[0.049]		[0.058]
P_u - male more than 30 years old		0.216***		0.192***		0.205***
		[0.073]		[0.063]		[0.071]
P_c - female less than 29 years old		0.071		0.084		0.086
		[0.125]		[0.109]		[0.124]
P_c - female more than 30 years old		-0.098		-0.010		-0.190
		[0.203]		[0.166]		[0.196]
P_c - male less than 29 years old		-0.086		-0.054		-0.126
		[0.105]		[0.084]		[0.101]
P_c - male more than 30 years old		-0.100		-0.083		-0.099
		[0.139]		[0.121]		[0.137]
Total number of observations used in system of equations (2)	2576	2573	2581	2581	2571	2571

Robust standard errors clustered at the couple level in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%. Estimates based on system of equations (2). Regressions in columns 1,3 and 5 include same controls as in Table 5. Regressions in columns 2,4 and 6 include same controls as in Tables 6 and A2.

Table A4: The impact of subjective beliefs on sexual behavior: allowing for misreporting of sexual behavior (Average marginal effects)

	Homogenous U and c		Heterogeneity by gender and age	
	Had multiple partners in the last 12 month in 2008 conditional on being sexually active	Misreporting	Had multiple partners in the last 12 month in 2008 conditional on being sexually active	Misreporting
P_u	0.146**			
	[0.066]			
P_c	-0.149			
	[0.123]			
P_u - female less than 29 years old			0.670**	
			[0.294]	
P_u - female more than 30 years old			-0.085	
			[0.252]	
P_u - male less than 29 years old			-0.027	
			[0.124]	
P_u - male more than 30 years old			0.391***	
			[0.125]	
P_c - female less than 29 years old			0.078	
			[0.218]	
P_c - female more than 30 years old			-0.231	
			[0.352]	
P_c - male less than 29 years old			-0.171	
			[0.161]	
P_c - male more than 30 years old			-0.276	
			[0.235]	
Female	-0.177***	0.164**	0.086	0.193***
	[0.042]	[0.074]	[0.165]	[0.069]
Had no partner in the last 12 months in 2006				
Had 1 partner in the last 12 months in 2006	0.073		0.151*	
	[0.073]		[0.082]	
Had multiple partners in the last 12 months in 2006	0.288***		0.385***	
	[0.095]		[0.102]	
Had 1 partner in the last 12 months in 2006 x female			-0.164	
			[0.113]	
Had multiple partners in the last 12 months in 2006 x female			-0.215	
			[0.161]	

Continuation of Table A4				
Less than 29 years old				
30-39	0.043	0.022	0.046	-0.027
	[0.047]	[0.124]	[0.049]	[0.090]
40-49	-0.059	-0.146	0.009	-0.180*
	[0.069]	[0.096]	[0.065]	[0.093]
50+	-0.048	-0.059	0.019	-0.101
	[0.054]	[0.099]	[0.058]	[0.083]
30-39 x female			-0.054	
			[0.081]	
40+ x female			-0.175	
			[0.118]	
No school				
Primary school	-0.141*	-0.357***	-0.160*	-0.365***
	[0.081]	[0.135]	[0.089]	[0.096]
Secondary school or more	-0.091	-0.242**	-0.125	-0.275**
	[0.796]	[0.123]	[0.095]	[0.111]
Married	-0.283***	-0.180	-0.290***	-0.141
	[0.049]	[0.177]	[0.052]	[0.128]
Married x female			0.006	
			[0.104]	
Catholic				
Muslim	-0.052	-0.226	-0.059	-0.226
	[0.082]	[0.147]	[0.086]	[0.141]
Indigenous Christian /African Independent Churches	0.216**	0.474***	0.248***	0.475***
	[0.104]	[0.180]	[0.094]	[0.170]
Other Christian	-0.065*	0.121	-0.071*	0.105
	[0.039]	[0.098]	[0.041]	[0.095]
Other or missing or no religions	-0.064	-0.021	-0.081	-0.066
	[0.052]	[0.119]	[0.057]	[0.147]
Own 2 acres or less of land				
Between 2 and 4 acres	0.126*	0.219***	0.155***	0.242***
	[0.065]	0.084	[0.060]	[0.086]
More than 4 acres	0.091*	0.212*	0.110**	0.221**
	[0.049]	[0.116]	[0.049]	[0.112]
N		1386		

Robust standard errors clustered at the couple level in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%. All regressions include regions dummies and a missing indicator for age, education and land ownership.

Table A5: Probability of being re-interviewed in 2008 (probit specification – Average marginal effects)

	Re-interviewed in 2008
2006 pre-test probability of being currently infected with HIV	-0.106***
	[0.039]
Probability of being alive in 10 years if healthy	-0.022
	[0.043]
Probability of being alive in 10 years if infected with HIV	0.075
	[0.060]
Probability that someone healthy of respondent's gender become infected with HIV in the next 12 months if several sexual partners in addition to spouse	-0.036
	[0.042]
Probability that someone healthy of respondent's gender become infected with HIV in the next 12 months if married to someone who is infected with HIV/AIDS	-0.006
	[0.064]
Female	0.028*
	[0.017]
Had no partner in the last 12 months in 2006	
Had 1 partner in the last 12 months in 2006	0.020
	[0.027]
Had multiple partners in the last 12 months in 2006	-0.001
	[0.038]
Less than 29 years old in 2008	
30-39	0.048**
	[0.022]
40-49	0.071***
	[0.025]
50+	0.050*
	[0.027]
No school in 2008	
Primary school	-0.029
	[0.023]
Secondary school or more	-0.056*
	[0.034]
Married in 2008	0.055**
	[0.024]
Catholic	
Muslim	-0.023
	[0.035]

Continuation of Table A5	
Indigenous Christian /African Independent Churches	-0.024
	[0.031]
Other Christian	-0.018
	[0.025]
Other or missing or no religions	0.013
	[0.041]
Own 2 acres or less of land	
Between 2 and 4 acres	0.023
	[0.020]
More than 4 acres	0.018
	[0.023]
N	2403

Robust standard errors clustered at the couple level in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%. Regression includes regions dummies and a missing indicator for age, education and land ownership.

Table A6: The impact of subjective beliefs on sexual behavior taking into account attrition (Average marginal effects)

	Had multiple partners in the last 12 month in 2008 conditional on being sexually active	Reinterviewed in 2008	Found by team of nurses conducting HIV test in 2006	Learned HIV status in 2006 conditional on being found	2006 pre-test probability of being currently infected with HIV
P_u	0.084**				
	[0.040]				
P_c	-0.053				
	[0.069]				
Female	-0.135***	0.022	0.010	0.010	0.091***
	[0.022]	[0.018]	[0.015]	[0.013]	[0.032]
Had no partner in the last 12 months in 2006					
Had 1 partner in the last 12 months in 2006	0.028	0.024	0.000	0.033	0.007
	[0.034]	[0.029]	[0.023]	[0.021]	[0.050]
Had multiple partners in the last 12 months in 2006	0.153***	-0.004	-0.049	0.049*	0.176***
	[0.039]	[0.040]	[0.031]	[0.028]	[0.062]
Less than 29 years old					
30-39	0.021	0.045*	0.099***	0.004	0.040
	[0.019]	[0.024]	[0.019]	[0.016]	[0.031]
40-49	0.009	0.052*	0.087***	0.011	-0.004
	[0.020]	[0.027]	[0.021]	[0.019]	[0.040]
50+	-0.007	0.055*	0.084***	-0.006	-0.028
	[0.020]	[0.028]	[0.021]	[0.019]	[0.045]
No school in 2008					
Primary school	0.017	-0.032	-0.014	0.011	-0.043
	[0.019]	[0.025]	[0.019]	[0.016]	[0.031]
Secondary school or more	0.037	-0.063*	-0.016	-0.027	0.007
	[0.028]	[0.036]	[0.029]	[0.025]	[0.053]
Married	-0.090***	0.058**	-0.053***	-0.039**	0.026
	[0.024]	[0.026]	[0.020]	[0.020]	[0.038]
Catholic					
Muslim	0.033	-0.005	0.002	-0.013	-0.053
	[0.024]	[0.037]	[0.024]	[0.023]	[0.040]
Indigenous Christian /African Independent Churches	0.026	0.000	-0.011	0.018	0.001
	[0.026]	[0.033]	[0.024]	[0.023]	[0.049]
Other Christian	-0.051**	0.006	0.022	0.006	-0.042
	[0.020]	[0.026]	[0.020]	[0.017]	[0.040]
Other or missing or no religions	-0.023	0.032	-0.017	0.075**	-0.081
	[0.030]	[0.042]	[0.032]	[0.037]	[0.062]
Own 2 acres or less of land					

Continuation of Table A6					
Between 2 and 4 acres	0.020	0.032	-0.017	-0.005	-0.012
	[0.016]	[0.021]	[0.016]	[0.015]	[0.030]
More than 4 acres	0.009	0.024	-0.033*	-0.022	-0.025
	[0.017]	[0.024]	[0.018]	[0.016]	[0.035]
2008 interview attempted in May					
June		-0.032			
		[0.032]			
July		-0.090*			
		[0.046]			
August		-0.189***			
		[0.056]			
2006 pre-test probability of being currently infected with HIV		-0.063	0.150***	-0.068	
		[0.070]	[0.052]	[0.047]	
First attempted visit of the 2006 team of nurses is within one week of the first attempted visit of the survey team			0.065***		
			[0.020]		
Receive positive financial incentive in 2004 testing experiment				0.033*	0.083**
				[0.018]	[0.042]
Did not participate in 2004 testing experiment				-0.054***	0.021
				[0.018]	[0.046]
Probability that someone healthy of respondent's gender become infected with HIV in the next 12 months with normal sexual behavior					0.129*
					[0.067]
N	2569				
sig_4	0.271*** [0.020]				
r_12	0.026 [0.092]				
r_13	-0.007 [0.099]				
r_14	-0.026 [0.126]				
r_15	0.327 [0.402]				
r_23	-0.525* [0.227]				
r_24	-0.413** [0.120]				
r_25	0.299*** [0.045]				
r_34	0.227* [0.113]				
r_35	-0.030 [0.061]				
r_45	-0.096 [0.111]				

Table shows marginal effects. In column 5, we report the coefficients, i.e. the marginal effect for f_i^{0*}. Robust standard errors clustered at the couple level in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%. All regressions include regions dummies and a missing indicator for age, education and land ownership.

Table A7: Potential endogeneity of transmission and survival expectations (Average marginal effects)

	Had multiple partners in the last 12 month in 2008 conditional on being sexually active	Found by team of nurses conducting HIV test in 2006	Learned HIV status in 2006 conditional on being found	2006 pre-test probability of being currently infected with HIV	Probability of survival in the next 10 years conditional on being infected with HIV	Difference in transm. risk associated with multiple versus one partner
P_u	0.115**					
	[0.048]					
P_c	0.170					
	[0.188]					
2006 pre-test probability of being currently infected with HIV		0.178*	-0.003			
		[0.106]	[0.054]			
Own visit gap less than 1 week		0.063***				0.045*
		[0.019]				[0.024]
Spouse visit gap less than 1 week						0.026
						[0.032]
Interaction of own and spouse visit gap						-0.055
						[0.048]
Receive positive financial incentive in 2004 testing experiment			0.031*	0.051		
			[0.018]	[0.054]		
Did not participate in 2004 testing experiment						
Probability that someone healthy of respondent's gender become infected with HIV in the next 12 months with normal sexual behavior				0.288***		
				[0.092]		
Child mortality expectations (from 0 to 10 beans)					-0.009***	
					[0.002]	
r_12			-0.005 [0.084]			
r_13			-0.003 [0.099]			
r_14			0.018 [0.115]			
r_16			0.204 [0.128]			
r_17			-0.040 [0.085]			
Total number of observations used			2,592			

Robust standard errors clustered at the couple level in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%. All regressions include same controls as in Table 5. Visit gap is difference between the first attempted visit of the 2006 team of nurses and the first attempted visit of the survey team.

Appendix B: Mortality and Infection Expectations Questions (Woman questionnaire)

I will ask you several questions about the chance or likelihood that certain events are going to happen. There are 10 beans in the cup. I would like you to choose some beans out of these 10 beans and put them in the plate to express what you think the likelihood or chance is of a specific event happening. One bean represents one chance out of 10. If you do not put any beans in the plate, it means you are sure that the event will NOT happen. As you add beans, it means that you think the likelihood that the event happens increases. For example, if you put one or two beans, it means you think the event is not likely to happen but it is still possible. If you pick 5 beans, it means that it is just as likely it happens as it does not happen (fifty-fifty). If you pick 6 beans, it means the event is slightly more likely to happen than not to happen. If you put 10 beans in the plate, it means you are sure the event will happen. There is not right or wrong answer, I just want to know what you think.

Let me give you an example. Imagine that we are playing Bawo. Say, when asked about the chance that you will win, you put 7 beans in the plate. This means that you believe you would win 7 out of 10 games on average if we play for a long time.

X2	Pick the number of beans that reflects how likely you think it is that…
a)	you are infected with HIV/AIDS now
FOR MARRIED RESPONDENTS (INTERVIEWER: If respondent is not married → X2f)	
b)	your spouse is infected with HIV/AIDS now
FOR UNMARRIED RESPONDENTS	
c)	your romantic partner is infected with HIV/AIDS now (INTERVIEWER: If no romantic partner, write 99 and → X2h)
FOR BOTH MARRIED AND UNMARRIED RESPONDENTS	
X3	Consider a healthy woman in your village who currently does not have HIV. Pick the number of beans that reflects how likely you think it is that she will become infected with HIV …
a)	during a single intercourse without a condom with someone who has HIV/AIDS
b)	within the next 12 months (with normal sexual behavior)
c)	within the next 12 months if she is married to someone who is infected with HIV/AIDS
d)	within the next 12 months if she has several sexual partners in addition to her spouse
e)	What about if this woman we just spoke about [in X3d] uses a condom with all extra-marital partners? How many beans would you leave on the plate?

Next, I would like you to consider the likelihood that somebody dies as time goes by. This is an imaginary person, and I am going to describe her to you. The beans in the plate represent the chances out of 10 that the person dies within a certain time period. The person is alive today so we start with an empty plate. As time goes by, more unfortunate things can happen and the person has more chances of dying, so more beans will be added to the plate"

INTERVIEWER:

1. Ask questions X4 to X5b for the INDIVIDUAL described in Column A. After X4 and X5a, LEAVE beans in plate. After X5b, put beans back in the cup. RECORD the number of beans in the plate after each question.
2. COLUMN by COLUMN, REPEAT questions X4 to X5b for the INDIVIDUALS described in Columns B, C and D. For each individual, LEAVE the beans in the plate after X4 and X5a, and put beans back in the cup after X5b. RECORD the number of beans in the plate after each question.
3. If respondent says "I Don't Know", probe with examples: "someone might die because of old age, disease, car accident. How likely do you think it is any of those things happen within [*for X4*: 1 year; *for X5a:* 5 years; *for X5b:* 10 years]?

RECORD the number of beans in the plate for each question.	DESCRIPTION OF INDIVIDUAL A A woman your age who is healthy and does not have HIV	 B A woman your age who is infected with HIV
X4 Pick the number of beans that reflects how likely you think it is that [INDIVIDUAL] will die within a <u>one-year</u> period beginning today. (LEAVE BEANS ON PLATE)	[_____] Beans in plate If 10, → X4 for individual B	[_____] Beans in plate If 10, →X4 for individual C
X5 Add additional beans so that the number of beans in the plate reflects how likely you think it is that [INDIVIDUAL] a) will die within a <u>five-year</u> period beginning today (LEAVE BEANS ON PLATE; IT IS POSSIBLE TO ADD ZERO ADDITIONAL BEANS)	[_____] Beans in plate If 10, → X4 for individual B	[_____] Beans in plate If 10, →X4 for individual C
b) will die within a <u>ten-year</u> period beginning today. (IT IS POSSIBLE TO ADD ZERO ADDITIONAL BEANS. PUT BEANS BACK IN CUP AFTER RECORDING THE ANSWER)	[_____] Beans in plate → X4 for individual B	[_____] Beans in plate → X4 for individual C
	C A woman your age who is sick with AIDS and is treated with ARV If R does not know about ARV, skip and go to X6	
X4 Pick the number of beans that reflects how likely you think it is that [INDIVIDUAL] will die within a <u>one-year</u> period beginning today. (LEAVE BEANS ON PLATE)	[_____] Beans in plate If 10, → X4 for individual D	
X5 Add additional beans so that the number of beans in the plate reflects how likely you think it is that [INDIVIDUAL] a) will die within a <u>five-year</u> period beginning today. (LEAVE BEANS ON PLATE; IT IS POSSIBLE TO ADD ZERO ADDITIONAL BEANS)	[_____] Beans in plate If 10, → X4 for individual D	
b) will die within a <u>ten-year</u> period beginning today. (IT IS POSSIBLE TO ADD ZERO ADDITIONAL BEANS. PUT BEANS BACK IN CUP AFTER RECORDING THE ANSWER)	[_____] Beans in plate → X4 for individual D	