WPS6823

POLICY RESEARCH WORKING PAPER   6823

The Heterogeneous Effects of HIV Testing

Sarah Baird
Erick Gong
Craig McIntosh
Berk Özler

Abstract

An extensive multi-disciplinary literature examines the effects of learning one's HIV status on subsequent risky sexual behaviors. However, many of these studies rely on non-experimental designs; use self-reported outcome measures, or both. This study investigates the effects of a randomly assigned home based HIV testing and counseling (HTC) intervention on risky sexual behaviors and schooling investments among school-age females in Malawi. The study finds no overall effects on HIV, Herpes Simplex Virus (HSV-2), or achievement test scores at follow-up. However, among the small group of individuals who tested positive for HIV, a large increase in the probability of contracting HSV-2 is found, with this effect stronger among those surprised by their test results. Similarly, those surprised by HIV-negative test results see a significant improvement in achievement test scores, consistent with increased returns to investments in human capital. The finding of increased HSV-2 prevalence among HIV-positive individuals suggests that the conventional wisdom that those who learn they are HIV-positive will adopt safer sexual practices should be treated with caution.

The Heterogeneous Effects of HIV Testing[*]

Sarah Baird, University of Otago/George Washington University
Erick Gong, Middlebury College
Craig McIntosh, University of California San Diego
Berk Özler, The World Bank/University of Otago

I. Introduction

HIV counseling and testing[1] is one of the pillars of HIV prevention (Potts et al. 2008). Recently, the U.S. Preventive Services Task Force released a draft recommendation statement suggesting that nearly everyone between the ages of 15-65 be screened for HIV (U.S. Preventative Services Task Force 2012). One argument for universal HIV testing is due to recent findings indicating that HIV-positive individuals who immediately receive antiretroviral therapy (ART) significantly reduce their transmission risk to uninfected partners and experience lower morbidity and mortality rates than those receiving delayed treatment (Cohen et al. 2011). A distinct argument suggests that HIV-positive individuals reduce risky sexual behaviors after learning their serostatus (Cassell and Surdo 2007; Gersovitz 2010; Fonner et al. 2012), indicating that testing itself would be a means of decreasing subsequent sexual transmission of disease.

Despite the belief that increasing awareness of HIV status has prevention benefits, there are no studies that have shown population level reductions in the incidence of HIV or other sexually transmitted infections (STIs) as a result of HIV testing (Potts et al. 2008).[2] Several studies in public health and economics have presented VCT impacts by HIV serostatus. These studies generally show no effects of testing on sexual behavior or seroconversion for HIV-negative individuals (Allen et al. 1992a; Weinhardt et al. 1999; Corbett et al. 2007; Thornton 2008; Fonner et al. 2012), but available evidence suggests that HIV-positive individuals reduce risky behaviors after testing (Allen et al. 1992b; Weinhardt et al. 1999; Boozer and Philipson 2000; Thornton 2008; Delavande and Kohler 2012; Fonner et al. 2012).[3]

---

[1] HIV counseling and testing encompasses both voluntary counseling and testing (VCT) and home based HIV testing and counseling (HTC).

[2] Furthermore, there are no randomized controlled trials (RCT) that directly assess subsequent HIV transmission for VCT vs. non-VCT. Kamb et al. (1998), using an RCT, assessed the effects of different modes of counseling for HIV-negative individuals on STI incidence, but everyone received an HIV test and learned their serostatus. More recently, Project Accept (HPTN 043) is assessing the effectiveness of community-based VCT relative to standard clinic-based VCT on community-level HIV incidence (http://www.cbvct.med.ucla.edu/overview.html). The Voluntary HIV-1 Counseling and Testing Efficacy Study Group (2000) found significant decreases in risky activity with non-primary partners for individuals and couples who received VCT.

[3] de Paula et al. (2010) reports that an increase in the beliefs of being HIV-positive reduced extra-marital affairs among married men in Malawi. Beegle, Poulin and Shapira (2012), using randomly assigned VCT in Malawi, find no effects on either sexual behavior or economic outcomes. Haile (2011) examines randomly assigned testing in Ethiopia and also finds limited impacts on sexual behavior and if anything finds that getting a negative test actually leads to more risky behavior.

The interdisciplinary literature summarized above is mainly based on self-reported data, which may be subject to social desirability bias (Allen et al. 2003; Minnis et al. 2009).[4] The possibility of bias may be particularly relevant in the case of HIV testing interventions. In a prominent study where offers of VCT were randomly assigned, those offered HIV tests self-report less unprotected sex with casual partners 6 months after the intervention, however there was no corresponding change in sexually transmitted infections (VCT Efficacy Study Group 2000). Re-examining the data, Gong (2013) finds suggestive evidence that the information provided in the pre- and post-test counseling sessions appeared to alter the stigma of reporting risky sexual behaviors at follow-up, leading to a potentially biased estimate of the actual behavior change. Even objectively observed behavior such as the purchase of condoms in Thornton (2008) may be subject to this kind of bias.[5]

In addition to these measurement limitations, the effects of testing are likely to display considerable heterogeneity across individuals. One strand of the literature in economics has focused primarily on the informational signal sent, suggesting that HIV tests should only have an effect on those who are 'surprised' by their HIV serostatus. In this case, the greatest effect will be found among those who believed they were HIV-negative (HIV-positive) but received a positive (negative) test result (Boozer and Philipson 2000). In fact, these two groups may have offsetting behavioral responses to testing, leading to small or no effects of testing on the entire population: in the extreme, the average treatment effect may be relevant for no one.

While this hypothesis indicates that people will change behavior only as a result of receiving new information, it does not tell us the direction in which behavior will change. The concept of 'fatalism' suggests that responses may be non-monotonic: individuals may increase their risky activity in response to increased HIV prevalence around them above a certain threshold (Kremer 1996; Kerwin 2012). Learning one's own HIV-positive status can be seen as an extreme case,

---

[4] In economics, Gong (2013), which examines the effects of a randomly assigned VCT intervention on incident gonorrhea and chlamydia, is a notable exception. In public health, two studies of women presenting at prenatal and pediatric outpatient clinics in Rwanda provide evidence on STI incidence for HIV-positive women (Allen et al. 1992a) and HIV seroconversion for the uninfected partner in cohabiting couples with serodiscordant results (Allen et al. 1992b), but both of these are prospective cohort studies comparing outcomes before and after VCT with no randomly assigned comparison group. The Voluntary HIV-1 Counseling and Testing Efficacy Study Group (2000) states that their RCT was not powered to detect effects on STIs, while another RCT was designed to assess HIV incidence among initially HIV-negative individuals and cannot comment on onward transmission from individuals who were HIV-positive at baseline (Corbett et al. 2007).

[5] Self-reported data in arguably less sensitive subjects, such as school enrollment (Baird and Özler 2012) and hand washing (Halder et al. 2010; Ram et al. 2010), have also been shown to be subject to social desirability bias.

where the cost of risky sex becomes zero for the infected individual who may become fatalistic because she has 'nothing to lose'. Gong (2013) presents a simple model to suggest that the effects of HIV testing are, a priori, ambiguous: the response of a utility maximizing individual to new information about HIV infection will depend on her prior beliefs of HIV infection and the degree to which altruism affects her behavior. Given the widely documented 'transactional' nature of risky sexual activity in Malawi (Poulin 2007; Swidler and Watkins 2007), the setting for this study, one might expect the direct benefits of risky activity to be high and altruism towards casual partners to be low for some individuals. In such circumstances, it is possible that individuals who find out that they are infected with HIV may increase their risky sexual activity.

Economic theory also suggests the possibility of changes for a broader set of behaviors in response to testing. For instance, if learning one's HIV status leads to changes in subjective life expectancy, this may lead to changes in consumption, savings, and investments in human capital (Thornton 2012) – changes that are also likely to be moderated by people's beliefs about HIV infection prior to testing (Yeatman 2009, Goldstein et al. 2013).

In this paper, we utilize objective outcome measures within a home based HIV testing and counseling (HTC) experiment to test these hypotheses among young females in Malawi. Specifically, we are interested in the effects of HTC on two outcomes: risky sexual behavior and investments in human capital. We first examine the overall effect of HIV testing on the population, and then motivated by theory, examine both the effects of testing by HIV-status (i.e. the effects of an HIV-negative/positive test), as well as the effects of being surprised by the test result.

In order to minimize the risk of bias from self-reports, we use HSV-2 as our primary outcome measure for sexual behavior and achievement tests in mathematics, English reading comprehension, and Raven's Colored Progressive Matrices as our primary measure of investments in human capital.[6] We also collected data on the participants' prior beliefs regarding HIV infection.

Overall, we find no significant effects of testing on HIV, HSV-2, or test scores. The lack of an overall effect is consistent with the existing literature that finds no effects among those who were

---

[6] In our sample of young females, we consider marginal investments in schooling to be the most important long-term decision, and focus on changes in achievement test scores to capture this effect.

HIV-negative at baseline – the group that constitutes 95% of our sample. However, the HSV-2 prevalence at follow-up among the small group of young women who received HIV-positive test results was 23 percentage points higher than HIV-positive individuals who did not receive HIV tests. While this result warrants some caution given the small sample size (N=73), the finding is at odds with most previous studies, which suggest that those who learn they are HIV-positive adopt safer sex practices.[7]

Consistent with economic theory, the perverse effect of testing is stronger among those surprised by HIV-positive tests, i.e. those who reported no chance of being infected with HIV at baseline. Also in line with theory, we find that testing can increase investments in one's own human capital. Those surprised by HIV-negative tests, i.e. those who report some chance of HIV infection at baseline but discover that they are HIV-negative, perform significantly better in achievement tests – a finding supported by an increase in self-reported school enrollment among the same subgroup.

Our results suggest that home based HIV testing and counseling can lead to important changes in human capital investments by providing vital health information and add to a larger literature on the effects of information-interventions (Nguyen 2008; Dupas 2010; Jensen 2010; Goldstein et al. 2013; Oster, Shoulson, and Dorsey 2013). Furthermore, our finding of increased STI risk among HIV-positive individuals who receive HIV tests serves as a caution to the assumption that such individuals will adopt safer sexual practices upon learning their serostatus. Finally, our study points to the importance of using objective outcome measures rather than self-reported behavior change when studying treatments that may differentially alter social desirability bias.

In the next section, we describe the setting of our study. Sections III and IV describe the methods and our findings. Section V provides a concluding discussion that includes the limitations of our study.

---

[7] The lone exception is a recent study by Gong (2013), who used biomarker data from a VCT experiment in a high-risk urban sample in Kenya and Tanzania and found a similar increase in STI incidence among those infected with HIV.

II. Setting

Malawi, which is a small and poor country in southern Africa, is the setting for this study. 81% of its population of 15.3 million lived in rural areas in 2009, with most people relying on subsistence farming. The country is poor even by African standards: Malawi's 2008 GNI per capita figure of $760 (PPP, current international $) is less than 40 percent of the Sub-Saharan African average of $1,973 (World Development Indicators 2010). As of 2007, the prevalence of HIV was estimated to be between 11.0% and 12.9% among adults aged 15-49 – the ninth-highest HIV prevalence in the world (UNAIDS 2008). The gender gap in HIV prevalence among young adults, aged 15-24, is startling: prevalence was more than *four* times higher for females than males in 2004 (9.1% vs. 2.1%) (National Statistical Office (Malawi) and ORC Macro 2004).

The study took place in Zomba District, in Southern Malawi. Zomba District is divided into 550 enumeration areas (EAs), which are defined by the National Statistical Office of Malawi and contain an average of 250 households spanning several villages. In 2007, 176 EAs in urban and rural areas of Zomba were selected to form the sample of a cash transfer experiment. In each of these EAs, all dwellings were visited to obtain a full listing of never-married females, aged 13-22. From this sampling frame, 3,796 school-age girls aged 13-22 were selected for inclusion to form the impact evaluation sample of the cash transfer experiment. This sample of approximately 22 girls per EA included two strata: those who were enrolled in school at time of listing and those who were out of school. Of the 176 EAs in the overall study, 88 were assigned to the treatment group where program participants were offered cash transfers. The other 88 were assigned to the comparison group, which received nothing.[8] This study assesses the impact of a randomly assigned home-based HIV testing and counseling (HTC) intervention (discussed in more detail in the Methods section below) within the 88 EAs in the comparison group.

---

[8] For more details on the cash transfer experiment and its impact evaluation, please refer to Baird, McIntosh and Özler (2011).

III. Methods

52 of the 88 EAs in the comparison group of the larger experiment were randomly assigned to receive HTC between the months of June and September in 2009.[9] Then, between the months of March and August 2010, i.e. after an average of approximately 10 months, all 88 EAs were offered HTC. The sample used in this paper therefore consists of 1948 females in 88 comparison EAs, of which 1122 in 52 EAs were offered HTC in 2009, with everyone offered HTC in 2010. In the remainder of this paper we refer to the group that was offered testing and counseling in 2009 as the *HTC* group, and those who were offered HTC only in 2010 as the *control* group.

HTC took place at the homes of the core respondents where they were invited to receive counseling and rapid testing for HIV, HSV-2, and syphilis by a trained counselor.[10] Overall acceptance of HTC was high at 98%, with no differences between the two study arms. Malawian nurses and counselors certified in conducting rapid HIV tests through the Ministry of Health HIV Unit HCT Counselor Certification Program conducted HTC. Testing and counseling was performed in a private location at or near the participant's house. Whole blood samples were obtained using a finger-stick. The teams conducting HTC at follow-up were blinded to the testing treatment status of the subjects and were independent of the survey teams that conducted household surveys and administered achievement tests (discussed below).[11] Ethics review committees at the National Health Sciences Research Council (Malawi), the University of California at San Diego (USA), and George Washington University (USA) approved the study design. Further details of the HTC procedures can be found in Appendix A.

---

[9] Details of the randomization for the HTC experiment can be found in Baird et al. (2012).

[10] Rapid tests for syphilis were only performed during the first round of HTC as its prevalence in this sample was found to be less than 1%. The effects of this intervention should be considered noting that the subjects were invited to receive counseling and testing not only for HIV but for HSV-2 and syphilis as well.

[11] HTC counselors could deduce the HTC treatment status of the subjects who had tested positive for HIV or HSV-2 at baseline. To avoid unnecessary burdens on the subjects, re-testing for the previously positive test was avoided in such circumstances, meaning that the HTC counselor could deduce that the individual had been tested before for the foregone STI test. Furthermore, to protect the privacy of all individuals, every individual in both the HTC and control groups were visited at follow-up: those who had tested positive for only one STI were tested for the other, while those who had tested positive for both were given a short survey. The counselors were blinded to the treatment status of everyone else, i.e. those in the HTC group who tested negative at baseline or those in the control group who were not tested at baseline.

The household survey data used in this paper were collected in two rounds, the second and third wave of the overall study.[12] The first household survey used in this study was conducted approximately 6 months prior to the first HTC data collection between October 2008 and February 2009. The second follow-up household survey was conducted between February and June 2010, just prior to the second round of the HTC testing. In addition to collecting data on the household, more pertinent for this study were the detailed information collected on the self-reported schooling and sexual activity of the respondents, as well as their subjective probabilities of being infected with HIV and their subjective life expectancies to ascertain prior beliefs and changes in those beliefs as a result of receiving HTC.[13]

Economic theory suggests that changes in subjective life expectancy may lead to changes in human capital investments. Perhaps the most critical investment that adolescents make is the laborious work of building their own human capital. Because the returns from investments in schooling are largely realized in the distant future they appear particularly informative as to inter-temporal tradeoffs. To measure improvements in student achievement, mathematics and English reading comprehension tests were developed and administered to all study participants at their homes as part of the follow-up data collection effort in 2010. The tests were developed by a team of experts at the Human Sciences Research Council according to the Malawian curricula for these subjects for Standards 5-8 and Forms 1-2.[14] In addition, to measure cognitive skills, we utilized a version of Raven's Colored Progressive Matrices that was used in the Indonesia Family Life Survey (IFLS-2).

Objective measures of behavior change minimize the risk of social desirability bias in the estimated treatment effects. In addition, the intervention itself might change the self-reporting bias of those receiving HTC. Thus, compared with self-reported sexual behaviors, biological markers of unprotected sex are likely to be preferable.[15] In this paper, we utilize biomarker data on both HIV and HSV-2, collected during HTC. For educational outcomes, we utilize test score data on mathematics, English reading comprehension, and cognitive ability, as they are likely to provide a more accurate measure of human capital investments than self-reported attendance or

---

[12] There was also a household survey conducted between October 2007 and January 2008 which served as a baseline for the cash transfer intervention, but is not used for this analysis.

[13] These surveys, as well as the achievement tests described below, are available from the authors upon request.

[14] In Malawi, there are eight grades in primary school (Standard 1-8) and four in secondary school (Form 1-4).

[15] However, such proxies are also imperfect surrogates for HIV transmission (Padian et al. 2011).

educational expenditures. The percentages of correct answers across these three tests were averaged to create an overall percentage of correct answers and standardized to serve as our primary outcome.

We supplement these objective measures of behavior with self-reported data from the household survey. We use the subjective likelihood of HIV infection at baseline to examine the heterogeneity of HTC impacts by prior beliefs.[16] Finally, we examine impacts on the following self-reported outcomes: subjective probability of HIV infection, perceived life expectancy, sexual behaviors, and school enrollment at follow-up. Given the caveats regarding the potential of bias in impact estimates on self-reported behaviors, we use these secondary outcomes to complement our findings of HTC effects on STIs and test scores.

In a randomized experiment with multiple rounds of data collection, the internal validity of the findings depends crucially on whether the randomization was conducted properly and whether attrition from the study sample was independent of treatment status. Table 1 presents a summary of baseline demographic characteristics, and measures of sexual behavior and education. Overall, the samples in the two arms are similar, with no difference between them that is statistically significant at the 5% level out of 16 comparisons, and two differences (female-headed household and never had sex) significant at the 10% level. The average age at baseline is just under 17 and more than three-quarters of the sample reported being enrolled in school – with the mean highest grade attended being Standard 8, i.e. the last year of primary school in Malawi. Less than a quarter of the sample reported being sexually active during the 12 months prior to baseline data collection, and thus, not surprisingly, less than 10% report any chance of being infected with HIV. Appendix Table B1 examines whether there is any differential attrition between the two study arms. The analysis indicates that the attrition in the control group is around 15% and treatment status is orthogonal to attrition from the sample.

---

[16] Self-reports at baseline may still suffer from social desirability bias, but such misreporting should be orthogonal to treatment status as HTC was randomly assigned and both groups were untreated at baseline.

Table 1: Baseline Balance

	Control	HTC	p-value (control-HTC)
Age	16.650	16.709	0.687
	(2.215)	(2.198)	
=1 if Live Inside 16km	0.604	0.575	0.841
	(0.489)	(0.495)	
=1 if Live Outside 16km	0.092	0.097	0.930
	(0.289)	(0.295)	
=1 if Live in Urban Area	0.305	0.328	0.874
	(0.461)	(0.470)	
Asset Index	0.179	0.421	0.535
	(2.429)	(2.595)	
=1 if Female Headed Household	0.432	0.367	0.053
	(0.496)	(0.482)	
=1 if in School	0.750	0.802	0.103
	(0.434)	(0.398)	
Highest Grade	8.188	8.407	0.199
	(1.993)	(1.981)	
=1 if Ever Married	0.090	0.075	0.412
	(0.287)	(0.264)	
=1 if Ever Pregnant	0.180	0.169	0.652
	(0.384)	(0.375)	
=1 if Never Had Sex	0.576	0.639	0.097
	(0.495)	(0.480)	
=1 if Sexually Active in Past 12 months	0.233	0.218	0.583
	(0.423)	(0.413)	
Number of Partners in Past 12 Months	0.240	0.231	0.760
	(0.442)	(0.458)	
=1 if Engage in Risky Sex	0.186	0.171	0.551
	(0.389)	(0.377)	
=1 if Partner Over 25	0.035	0.034	0.943
	(0.184)	(0.183)	
=1 if Any Chance Infected with HIV Now	0.079	0.088	0.611
	(0.271)	(0.283)	
Number of observations	720	961	

Notes: Column (1) shows the baseline means for the control group while column (2) shows the means for the tested group. Column (3) presents the p-value on the difference between the two means. All variables are weighted to make them representative of the target population in the study EAs with robust standard errors clustered at the EA level to calculate the p-value.

We estimate the overall effects of HTC on HIV and HSV-2, educational achievement, and some selected self-reported outcomes using an intention-to-treat (ITT) estimator. We use a linear probability model:[17]

$$(1)\ Y_{ij} = \alpha + \beta HTC_j + \gamma \left(X_{ijk} - \bar{X}_k \right) + \delta \left(HTC_j * \left(X_{ijk} - \bar{X}_k \right) \right) + \varepsilon_{ij}$$

---

[17] Probit specifications generate similar estimates when compared to our main results.

where Y_{ij} is the outcome of interest for individual i in treatment cluster j, HTC_j is an indicator of whether the cluster is randomly assigned to receive HTC in 2009, and X_{ijk} is a set of K individual-level baseline characteristics. This small set of controls, X_{ijk}, includes age, an indicator for never having had sex, and school enrollment status at baseline, all of which are demeaned and fully interacted with the treatment indicator. These baseline controls are prognostic of STI status at follow-up and orthogonal to treatment assignment, and thus they substantially improve the precision of our estimates.[18] The error terms, ε_{ij}, are clustered at the EA level, which accounts both for the design effect of our EA level treatment and the heteroskedasticity inherent in the linear probability model. Age- and stratum-specific sampling weights are used to make the results representative of the target population in the study area.

We also estimate the effect of HTC separately for those receiving HIV-positive and HIV-negative test results, as we expect behavioral responses to differ by HIV status.[19] Under the ideal study design to address this question, we would compare individuals who are HIV-positive (negative) at baseline and receive HTC to control individuals who have the same HIV-status at baseline but are not informed of their status. However, the design of this study forewent collecting samples from individuals without informing them of the test results, meaning that baseline HIV status is observed only in the HTC group.[20] Since we do not observe baseline HIV

---

[18] While our primary analysis includes adjustments using covariates measured at baseline, we also present unadjusted estimates of HTC effects for full transparency. To select the baseline covariates for adjustment, we ran a stepwise regression using 13 explanatory variables that theory suggests should be predictive of HSV-2 status at follow-up and retained those variables that were significant at the 10% level. In the absence of a pre-analysis plan, this procedure largely removes the potential for ad hoc specification searching. The interacted adjustment produces, asymptotically and for finite samples, the most precise average treatment effect (Lin 2013).

[19] Appendix Tables B2 and B3 show baseline balance for these two sub-groups.

[20] Both types of study designs can be found in the existing literature. Corbett et al. (2007) reports collecting venous or finger-prick blood or oral mucosal transudate for anonymous HIV testing from all employees at the 22 businesses that were recruited for their study. Similarly, an ongoing study assessing the impacts of HIV testing by Duflo, Dupas, and Sharma also utilized anonymous linked testing of all study participants at baseline (http://www.povertyactionlab.org/node/5649). On the other hand, The Voluntary HIV-1 Counseling and Testing Efficacy Study Group (*2000*) thought neither collecting serum and testing for HIV without giving the individuals the results at baseline, nor testing for HIV but not providing counseling to be ethical. Thornton (2008) overcomes this identification problem by collecting blood samples from the entire sample and providing randomly varying levels of incentives for individuals to come back and learn their test results two to four months after sample collection and by randomizing the location of the temporary centers where the study participants can obtain their results. However, it is important to note that such estimates are local average treatment effects on those individuals who would have only learned their results due to the incentive and would not have otherwise. The introduction of HIV incidence assays, which can detect recent infections, may also provide further improvements in this area. As described in detail below,

status for individuals in the control group, we cannot form the true counterfactual comparison group. Instead, we present analysis by HIV status using the observed status at baseline for the treatment group, and at follow-up for the comparison group. Because the difference between the estimator we *wish to* construct and the one we *can* construct is unobserved but still well defined, its extreme values can be calculated with no assumptions beyond simple random assignment. Below, we provide an informal description of how we estimate HTC effects on HSV-2 by baseline HIV-serostatus. Appendix C formally defines our estimator and describes the calculation of the bounds within which the true ITT lies.

Individuals in the control group who tested HIV-positive at follow-up fall into two sub-groups: those who were HIV-positive at baseline and those who became infected with HIV between baseline and follow-up.[21] This implies that the counterfactual HSV-2 prevalence for the HTC treatment is contaminated by the presence of HIV-seroconverters: we would like to observe the follow-up HSV-2 prevalence among only those who were HIV-positive at baseline in the control group, but instead we observe it among them **and** HIV seroconverters. The bias caused by this latter group is a function of two quantities: its size in our sample and the HSV-2 prevalence among them. We can estimate the former by calculating the difference between the HIV prevalence in the control group at follow-up and the HIV prevalence in the treatment group at baseline. Both of these groups are untested, and the samples are comparable due to random assignment into treatment. Using only this quantity, we can then calculate bounds for how large the bias caused by HIV-seroconverters to the impact estimates of HSV-2 can possibly be, under two extreme scenarios: (i) all of the HIV-seroconverters are HSV-2 positive in the follow-up; or (ii) none of them are.

IV. Results

Table 2 presents our main findings. It first reports estimates for outcomes objectively measured at follow-up (HIV, HSV-2, and achievement test results, unadjusted and adjusted, in columns 1-6), and then for outcomes measured using self-reports (beliefs on HIV infection, subjective life

---

our empirical approach allows us to estimate the intention-to-treat (ITT) effect and provide bounds on the value of the true ITT by simply offering testing in some clusters and not in others.

[21] The description we provide here is for the case of the HIV-positive group, but the argument is symmetric for the estimator for the HIV-negative group.

expectancy, sexual behavior, and school enrollment – all adjusted estimates – in columns 7-12). Panel A contains HTC impacts on the entire sample and shows that HTC had no statistically significant effect on the prevalence of HIV, HSV-2, or test scores at follow-up. These null effects are measured fairly precisely: we would have been able to detect a 2.7 percentage point effect on HSV-2 prevalence (follow-up prevalence in the control group is 7.4%) and a 0.23 standard deviation effect on test scores with 95% confidence. When we examine self-reported data, we find no significant changes in subjective life expectancy, likelihood of HIV infection, self-reported school enrollment, or unsafe sex, although there is a small but statistically significant increase in the likelihood of being sexually active during the past 12 months. Hence, overall, our findings suggest that the HTC intervention did not have any significant effects on STI risk or investments in human capital in the study population.

Panel A shows that the follow-up HIV prevalence in our sample of young women is fairly small at approximately 5%, indicating that more than 95% of the HTC group received a HIV-negative test result at baseline. Panel B shows that the HTC impacts among this group are generally very similar to those among the entire sample: HTC caused no changes in HSV-2 risk or in achievement test scores among this group, although they assign a lower average probability of being infected with HIV at follow-up and report a slightly higher number of sexual partners over the past 12 months.

Table 2: HTC Impacts on Objective and Self-Reported Outcomes

	Panel A: HTC Impacts on the Entire Sample											
	Objective						Self-Reported					
	=1 if HIV Positive	=1 if HIV Positive	=1 if HSV-2 Positive	=1 if HSV-2 Positive	Achievement Test Score	Achievement Test Score	Subjective Likelihood of HIV Infection	Probability Live to 50	=1 if Sex Active Past 12 Months	Number Partners Past 12 Months	=1 if Engage in Risky Sex	=1 if in School
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
HTC	-0.006	-0.003	0.003	0.010	0.137	0.073	0.001	-0.159	0.042	0.059*	0.024	-0.007
	(0.011)	(0.010)	(0.014)	(0.014)	(0.118)	(0.099)	(0.008)	(0.222)	(0.027)	(0.034)	(0.028)	(0.022)
Number of observations	1,681	1,681	1,680	1,680	1,660	1,660	1,673	1,673	1,671	1,671	1,669	1,673
Mean in the control group	0.050	0.050	0.074	0.074	0.000	0.000	0.041	5.370	0.353	0.360	0.332	0.525
	Panel B: HTC Impacts on the HIV-Negative Sample											
HTC			-0.004	0.000	0.146	0.088	-0.010*	-0.097	0.040	0.057*	0.030	-0.003
			(0.012)	(0.013)	(0.122)	(0.104)	(0.006)	(0.227)	(0.028)	(0.033)	(0.029)	(0.022)
Number of observations			1,607	1,607	1,591	1,591	1,602	1,602	1,600	1,600	1,600	1,602
Mean in the control group			0.065	0.065	0.004	0.004	0.033	5.331	0.338	0.346	0.319	0.534
	Panel C: HTC Impacts on the HIV-Positive Sample											
HTC			0.231*	0.246***	-0.125	-0.233	0.291*	-0.423	0.109	0.133	-0.030	-0.060
			(0.137)	(0.069)	(0.230)	(0.172)	(0.162)	(0.949)	(0.098)	(0.092)	(0.079)	(0.135)
Number of observations			73	73	69	69	71	71	71	71	69	71
Mean in the control group			0.255	0.255	-0.069	-0.069	0.192	6.140	0.629	0.629	0.579	0.356
Includes Controls	No	Yes	No	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Regressions are OLS models with robust standard errors clustered at the EA level. All regressions are weighted to make them representative of the target population in the study EAs. The achievement test score is the average percent across math, English and cognitive tests (standardized). The subjective likelihood of HIV infection takes on a value from 0 to 1 where zero is no self-reported chance of HIV infection. Risky sex takes on a value of 1 if the respondent does not always use condoms with all partners, and is zero for those who always use condoms or have not engaged in sexual activity in the past 12 months. Regressions that include controls include the following: age, an indicator for never having had sex, and school enrollment status at baseline, all of which are demeaned and fully interacted with the treatment indicator. Parameter estimates statistically different than zero at 99% (***), 95% (**), and 90% (*) confidence.

However, when we examine the effects of HTC on those who tested HIV-positive at baseline, the story is starkly different (Panel C). Among the 73 HIV-positive individuals in our sample, the follow-up prevalence of HSV-2 is 23.1 percentage points higher in the HTC group than the control group mean of 25.5% -- statistically significant at the 10% level.[22] Self-reports suggest changes in subjective beliefs as well: those in the HTC arm who receive HIV-positive tests assign a higher average probability to being HIV-positive at follow-up.[23] The self-reported outcomes on sexual behavior suggest some increase in sexual activity and the number of partners during the past 12 months but none of these variables is statistically significant. We do not find any significant effects on test scores.[24]

As described earlier, our naïve estimator of HTC impacts by baseline HIV status is biased by the presence of HIV-seroconverters among the counterfactual group at follow-up. To assess whether the HTC effect on HSV-2 prevalence among the HIV-positive subgroup is robust to this issue, we conduct a bounding exercise, which was summarized in the previous section and formally outlined in Appendix C.[25] While the lower and upper bounds for the true ITT estimate can be calculated analytically, in order to provide standard errors on the bounds we calculate them empirically. We do this by repeatedly and randomly excluding the expected number of HIV-seroconverters from our analysis and recalculating the HTC effect on HSV-2. The baseline HIV prevalence in the HTC group is 3.96% while the follow-up HIV prevalence in the control group is 5.03%. Neither of these figures is affected by HTC as they both reflect prevalence rates in each study arm prior to HTC. The difference between these two figures, which is equal to the

---

[22] The adjusted estimate is very similar at 24.6 percentage points and significant at the 1% level.

[23] While the average HTC effect on beliefs of HIV infection is strong and statistically significant, 43% of those who received HIV-positive test results reported *no* (zero) chance of being infected with HIV at the follow-up survey. This finding is unlikely to be due to the testing procedures. As detailed in Appendix A, those who tested positive with one rapid test received a second independent test using a different test kit. On the very unlikely event that the two test results were discordant, a third test was administered to break the tie. It is consistent with two other studies from Malawi, both of which find that approximately 40% of individuals who received HIV-positive test results through a VCT intervention reported *no* likelihood of being infected in a follow-up survey (Thornton 2012; Delavande and Kohler 2012).

[24] The reader should note the limited statistical power of the impact analysis among this subgroup due to its small size in our sample. We have a total of 73 HIV-positive individuals in our analysis, which is similar in size to Thornton (2012) who uses a sample of 79 HIV-positive individuals, and Delavande and Kohler (2012) who use a sample of 97 HIV-positive individuals. Gong (2013), using data on an older sample of people who were seeking HIV-related services, analyzes a sample of 465 HIV-positive individuals.

[25] We do not present the results of the bounding exercise for the HIV-negative group because the small number of HIV-seroconverters forms a negligible share of the sample of 1,607 HIV-negative individuals and, hence, does not bias our estimates in any meaningful way.

estimated HIV incidence rate, implies that, of the 40 individuals in the control group who tested HIV-positive at follow-up, 32 are expected to have been HIV-positive at baseline and eight to have seroconverted between the two HTC rounds. To estimate a lower bound HTC effect, we randomly exclude eight HSV-2-negative individuals from the control group, rerun the adjusted regression described in equation (1), record the ITT estimate β and its standard error, and repeat this 500 times.[26] Repeating the same exercise, this time randomly excluding eight HSV-2-positive individuals from the control group 500 times, yields the upper bound ITT estimate.

Table 3 presents the results and shows that the lower bound estimates of the HTC effect is 0.224 while the upper bound estimate is 0.328 – both statistically significant at the 1% level (columns 2 & 3). Hence, under the assumption of baseline balance in HIV prevalence between the HTC and control groups, the true HTC effect on HSV-2 prevalence among the HIV-positive sample in this study must lie somewhere between 22.4 and 32.8 percentage points and the potential bias caused by the seroconverters cannot be large enough to overwhelm the perverse effect of HIV testing. The naïve ITT estimate of 0.246 (column 1) reported in Panel C of Table 2 is robust to extreme assumptions about the HSV-2 status of HIV-seroconverters.

The above bounding exercise simply requires us to assume the randomization led to baseline balance of HIV prevalence. We can examine how imbalanced HIV prevalence would have to be between the HTC and the control group to explain away our results. To estimate this, we simply extend the calculation of the lower bound estimate described above. We remove a randomly selected number of HSV-2 negative individuals from the control group and rerun the adjusted regression (equation 1); we keep removing individuals until we no longer are able to reject the null of no effect of HTC at the 10% level. We find that it would take 22 HIV-seroconverters who are all HSV-2 negative in order to fail to estimate significant effects of HIV-positive tests on follow-up HSV-2 rates. 22 HIV-seroconverters implies that HIV-prevalence at baseline for the control group would have been 2.28% compared to 3.96% in the HTC treatment arm.

---

[26] This simulation produces an average HIV prevalence of 4.03% in the control group at baseline – very close to the 3.96% figure in the HTC group at baseline.

Table 3: Bounding HTC Impacts on HSV-2 for the HIV-Positive Sample

	Estimated ITT	Lower Bound ITT	Upper Bound ITT
	(1)	(2)	(3)
HTC	0.246***	0.224***	0.328***
	(0.069)	(0.072)	(0.068)
Number of Observations	73	65	65

Notes: Estimated ITT in column 1 is identical to the estimate in column 2, Panel C of Table 2. To estimate the lower bound ITT, we randomly exclude 8 HSV-2-negative individuals from the control group, which is the expected number of HIV seroconverters in the control group, rerun the regression described in equation (1), record the ITT estimate β and its standard error, and repeat this 500 times. The mean values of the coefficient estimate and its standard error are reported in column 2. Repeating this simulation exercise, this time randomly excluding 8 HSV-2-positive individuals from the control group, we calculate the upper bound ITT, which is reported in column 3. Please see details in Results section and Appendix C. Regressions are OLS models with robust standard errors clustered at the EA level. All regressions are weighted to make them representative of the target population in the study EAs. All regressions include the following controls: age, an indicator for never having had sex, and school enrollment status at baseline, all of which are demeaned and fully interacted with the treatment indicator. Parameter estimates statistically different than zero at 99% (***), 95% (**), and 90% (*) confidence.

We conclude our empirical analysis by testing the hypothesis that the effects of HIV-positive (negative) tests should be concentrated among those who considered themselves at no (some) risk of HIV at baseline because HTC should lead to behavior change only if it is providing new information.[27] Table 4 provides support for this hypothesis. Panel A presents HTC effects for those surprised by their HIV-negative status: the first four columns show the effects on HSV-2 and test scores (unadjusted and adjusted) while the remaining columns show effects on the self-reported outcome measures (adjusted) presented in Table 2.[28] We find an economically meaningful and statistically significant increase in achievement test scores in this subgroup: the unadjusted and adjusted estimates of HTC impact on test scores are 0.487 and 0.400 standard deviations, respectively – both significant at the 5% level. Self-reported outcomes point to the revelation of new, important information on health and increased investment in schooling among as potential channels for this finding: the young women in this subgroup are 12.3 percentage

---

[27] Appendix Tables B4 and B5 show baseline balance among these sub-groups.

[28] We define this subgroup as those who reported some likelihood of being infected with HIV prior to participating in HTC and found out otherwise. Similarly we define those surprised by their HIV-positive status as those who reported no chance of being infected with HIV prior participating in HTC but tested positive.. . We find that the self-reported likelihood of being infected with HIV is positively and significantly associated with actual HIV-status at baseline amongst those in the HTC sample. .

points less likely to think they are HIV-positive (compared to a control mean of 15.9%; p-value=0.000) and 13.3 percentage points more likely to report being enrolled in school at follow-up (compared to a control mean of 29.5%; p-value=0.158). The impact estimates on HSV-2 are in the expected direction but not statistically significant.

Panel B presents HTC effects for those surprised by their HIV-positive status. The unadjusted and adjusted HTC effects on HSV-2 prevalence in this subgroup are 0.299 and 0.215, respectively – both significant at the 5% level. Individuals in this group are more than 50 percentage points more likely to think they are infected with HIV at follow-up (compared with less than 2% in the control group; p-value=0.000). The impact estimates for test scores are, again, in the expected direction but not statistically significant. We find no significant impacts on any of the other self-reported outcomes. The reader should note, however, that the size of this subgroup is small and the impact estimates are accordingly imprecise.

Table 4: HTC Impacts on Objective and Self-Reported Outcomes by Prior HIV Beliefs

	Panel A: High Prior Beliefs of HIV Infection Before HTC and Tested Negative									
	Objective				Self-Reported					
	=1 if HSV-2 Positive	=1 if HSV-2 Positive	Achievement Test Score	Achievement Test Score	Subjective Likelihood of HIV Infection	Probability Live to 50	=1 if Sex Active Past 12 Months	Number Partners Past 12 Months	=1 if Engage in Risky Sex	=1 if in School
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
HTC	-0.029	-0.039	0.487**	0.400**	-0.123***	0.219	0.011	-0.029	-0.005	0.133
	(0.040)	(0.030)	(0.217)	(0.185)	(0.031)	(0.592)	(0.080)	(0.084)	(0.079)	(0.093)
Number of observations	143	143	142	142	143	143	143	143	143	143
Mean in control	0.091	0.091	-0.299	-0.299	0.159	4.816	0.449	0.505	0.421	0.295
	Panel B: Low Prior Beliefs of HIV Infection Before HTC and Tested Positive									
HTC	0.299**	0.215**	-0.010	-0.166	0.524***	0.195	-0.035	0.009	-0.100	0.090
	(0.139)	(0.098)	(0.304)	(0.288)	(0.146)	(1.242)	(0.149)	(0.155)	(0.156)	(0.209)
Number of observations	53	53	50	50	51	51	51	51	49	51
Mean in control	0.218	0.218	-0.254	-0.254	0.016	5.885	0.664	0.664	0.569	0.231
Includes Controls	No	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Regressions are OLS models with robust standard errors clustered at the EA level. All regressions are weighted to make them representative of the target population in the study EAs. The achievement test score is the average percent across math, English and cognitive tests (standardized). The subjective likelihood of HIV infection takes on a value from 0 to 1 where zero is no self-reported chance of HIV infection. Risky sex takes on a value of 1 if the respondent does not always use condoms with all partners, and is zero for those who always use condoms or have not engaged in sexual activity in the past 12 months. Regressions that include controls include the following: age, an indicator for never having had sex, and school enrollment status at baseline, all of which are demeaned and fully interacted with the treatment indicator. Parameter estimates statistically different than zero at 99% (***), 95% (**), and 90% (*) confidence

V. Concluding Discussion

We conducted a cluster randomized home based HIV testing and counseling intervention among young females in Malawi to disentangle its effects on risky sexual behavior and human capital investments. While HIV testing plays an unambiguously important role as a gateway to ART and, hence, would have indirect effects on HIV transmission through suppressed infectiousness after the start of therapy, its direct behavioral effects appear to be complex and not uniformly protective. Our main finding is that while HTC has no overall effect on STI risk or achievement test scores, it does have, consistent with economic theory, heterogeneous effects in subgroups. Among those who discover that they are HIV-positive, the risk of HSV-2 infection at follow-up increases substantially. On the other hand, individuals surprised by HIV-negative test results experienced a significant improvement in achievement test scores.

Until very recently, the literature on the effects of HIV testing on subsequent sexual behavior almost uniformly suggested that it caused HIV-positive individuals to become safer, not more risky. However, this study and Gong (2013), which share random assignment of HIV testing and the use of STI biomarkers as primary outcomes, come to the opposite conclusion. The sample in Gong (2013) consists of male and female adults seeking HIV-related services in Kenya and Tanzania in the mid-1990s, while our study examines a random sample of young females in Malawi in 2009. The common findings in these two studies, which utilize independent data sets from different countries, decades, and demographic groups, cast some doubt on the assumption that standard HIV testing interventions have direct prevention benefits through behavior change. We suggest that caution is warranted in interpreting findings from testing studies that do not simultaneously tackle both the endogeneity of testing and the risk of bias from self-reported sexual behavior data.

The study's main finding, i.e. that HIV-positive tests lead to increases in HSV-2 rates suggesting increases in unprotected sex, warrants some caution given that it is based on a small sample size (N=73). One concern is that a small sample size increases the chance of baseline imbalance for our primary outcome measures. Given the study design, whereby baseline STI status is unknown in the control group, it is impossible to demonstrate balance for HIV and HSV-2. However, we note that the main findings are robust to the inclusion of baseline covariates that are prognostic of HSV-2 infection at follow-up. Another concern is that our

statistical inference relies on asymptotic theory, which may not hold at smaller samples such as ours. To address this, we implement a non-parametric permutation test as defined by Anderson (2008) that does not rely on asymptotic theory (details in Appendix D). Using the same specification as in Table 2, Panel C, Column 4, "HTC Impacts on the HIV-Positive Sample", we find a p-value<.01, allowing us to reject the null of no effect at the 1% level - which is exactly what we concluded earlier.

There are two additional limitations to this study. First, we do not have the ideal study design required to assess testing impacts by HIV serostatus, which requires collecting blood samples from all study participants at baseline but not informing the control group of their HIV serostatus. We show that under random assignment of clusters to testing or no testing, researchers can obtain an ITT estimate for the causal effect of testing separately for HIV-positive (and negative) individuals and construct bounds around the true ITT effect. Secondly, the study does not measure subsequent HIV transmission to sexual partners. If there is serosorting (i.e. HIV-positive females pair with HIV-positive males upon learning their status) then the increase in HSV-2 infections we observe among HIV-positive females could occur without any effects on the total number of new HIV infections. We find this to be unlikely in this study as our data show no indications of such serosorting among HSV-2 seroconverters in the HTC arm.[29]

Keeping these limitations in mind, our study makes some contributions to the literature. Using survey questions designed to test several hypotheses regarding the effects of HIV testing, we find support for the hypothesis that the effects of testing are concentrated among those 'surprised' by HIV test results, both in terms of sexual behavior and educational achievement, the latter of which is consistent with increased effort and investment in human capital in anticipation of a longer, healthier life. This finding suggests that HIV tests can affect important investment decisions among young individuals by revealing an HIV-negative status to those who previously

---

[29] An examination of the answers of HIV-positive HSV-2 seroconverters to follow-up survey questions about their sexual partners reveals that (i) all of them attach either 'no likelihood or 'low likelihood that their partners are HIV-positive; (ii) none of them think that their partner had other sexual partners while they were together; (iii) only half of them report that their partner was ever tested for HIV; and (iv) they all had age appropriate partners. Furthermore, the distributions of these variables are similar both among all of those in the VCT group who tested HIV-positive at baseline and among those in the control group who tested HIV-positive at follow-up. Another possibility is that individuals who found out they are HIV-positive switched to oral sex to protect their partners. We find this explanation also unlikely since the self-reported likelihood of receiving oral sex by men is very small in Malawi – only 2% of rural men and 11% of urban men report ever having received oral sex. Furthermore, men substantially overestimate the probability of HIV infection through oral sex, indicating that it would not necessarily be seen as safer sex (Kerwin 2012; Kerwin, Thornton and Foley 2012).

thought they might be infected with HIV. This finding also suggests that exposure to infection risk and uncertainty about HIV status may cause young people to lower their investments in education.

These insights regarding the information signal provided by HIV tests help explain the main findings in Table 2. Among this young population, most individuals have no expectation of being HIV-positive, so an HIV-negative test result does not surprise many. This is consistent with finding no overall effect of HIV tests among the HIV-negative group while finding an improvement in educational achievement once we limit our analysis to those surprised by an HIV-negative test. Similarly, an HIV-positive test result is an unwelcome surprise to many in our sample, which explains the large effects of tests on HSV-2 among the entire HIV-positive group.

The troubling perverse effect of HIV-positive test results on HSV-2 serves as a caution to practitioners and merits careful consideration in policy terms. Following the logic presented in Gong (2013), this result is suggestive of high short-term benefits from risky activity and low altruism towards marginal partners – a scenario consistent with transactional sex. If such relationships are partially to blame for this finding, policies that reduce transactional sex may also reduce these unwanted effects of HIV testing. Our results also suggest that standard testing interventions, which simply convey HIV-status and provide brief pre- and post-test counseling sessions, may be insufficient to cause declines in risky sexual activity among those who find out they are infected. Other types of post-test support may be needed to ensure that such individuals are both linked to treatment and minimize the risk of transmission to others. Future RCTs that experiment with alternative means of informing, counseling, and providing post-test support to those who receive HIV-positive test results can be useful in creating effective positive prevention strategies among young people.

References

Allen, Susan, Antoine Serufilira, Joseph Bogaerts, Philippe Van de Perre, Francois Nsengumuremyi, Christina Lindan, Michel Carael, William Wolf, Thomas Coates, Stephen Hulley, 1992a. "Confidential HIV testing and condom promotion in Africa." *JAMA* 268: 3338-3343.

Allen, Susan, Jeffrey Tice, Philippe Van de Perre, Antoine Serufilira, Esther Hudes, Francois Nsengumuremyi, Joseph Bogaerts, Christina Lindan, Stephen Hulley, 1992b. "Effect of serotesting with counseling on condom use and seroconversion among HIV discordant couples in Africa." *BMJ* 304: 1605-1609.

Allen, Susan, Jareen Meinzen-Derr, Michele Kautzman, Isaac Zulu, Stanley Trask, Ulgen Fideli, Rosemary Musonda, Francis Kasolo, Feng Gao, Alan Haworth, 2003. "Sexual behavior of HIV discordant couples and HIV counseling and testing." *AIDS* 17: 733-740.

Anderson, Michael, 2008. "Multiple Inference and Gender Differences in the Effects of Early Intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects." *Journal of the American Statistical Association 103: No. 484.*

Baird, Sarah, Richard Garfein, Craig McIntosh, and Berk Özler, 2012. "Effect of a cash transfer programme for schooling on prevalence of HIV and herpes simplex type 2 in Malawi. A cluster randomized trial." *Lancet* 379: 1320-1329.

Baird Sarah, Craig McIntosh, and Berk Özler. 2011. "Cash or condition? Evidence from a cash transfer experiment." *Quarterly Journal of Economics* 126: 1709-1753.

Baird, Sarah, and Berk Özler. 2012. "Examining the reliability of self-reported data on school participation." *Journal of Development Economics* 98: 89-93.

Beegle, Kathleen, Michelle Poulin, and Gil Shapira, 2012. Does HIV testing change behaviors of young adults in Malawi? Paper presented at the Sixth Annual Research Conference on Population, Reproductive Health and Economic Development, Accra, Ghana.

Boozer, Michael and Tomas Philipson. 2000. "The impact of public testing for human immunodeficiency virus." *Journal of Human Resources* 35(3): 419-446.

Cassell, Micahel M. and Alison Surdo. 2007. "Testing the limits of case finding for HIV prevention." *Lancet Infectious Diseases* 7: 491-495.

Cohen, Myron S., Ying Q. Chen, Marybeth McCauley, Theresa Gamble, Mina C. Hosseinipour,

Nagalingeswaran Kumarasamy, James G. Hakim, Johnstone Kumwenda, Beatriz Grinsztejn, Jose H.S. Pilotto, Sheela V. Godbole, Sanjay Mehendale, Suwat Chariyalertsak, Breno R. Santos, Kenneth H. Mayer, Irving F. Hoffman, Susan H. Eshleman, Estelle Piwowar-Manning, Lei Wang, Joseph Makhema, Lisa A. Mills, Guy de Bruyn, Ian Sanne, Joseph Eron, Joel Gallant, Diane Havlir, Susan Swindells, Heather Ribaudo, Vanessa Elharrar, David Burns, Taha E. Taha, Karin Nielsen-Saines, David Celentano, Max Essex, and Thomas R. Fleming for the HPTN 052 Study Team, 2011. "Prevention of HIV-1 infection with early antiretroviral therapy." *The New England Journal of Medicine* **365**, 493-505 (2011).

Corbett, Elizabeth L., Beauty Makamure, Yin Bun Cheung, Ethel Dauya, Ronnie Matambo, Tsitsi Bandason, Shungu S. Munyati, Peter R. Mason, Anthony E. Butterworth, and Richard J. Hayes, 2007. "HIV incidence during a cluster-randomized trial of two strategies providing voluntary counseling and testing at the workplace, Zimbabwe." *AIDS* 21: 483-489.

Delavande, Adeline and Hans-Peter Kohler. 2012. "The impact of HIV testing on subjective expectations and risky behavior in Malawi." *Demography* 49: 1011-1036.

de Paula, Aureo, Gil Shapira, and Petra Todd. 2011. "How beliefs about HIV status affect risky behaviors: evidence from Malawi, sixth version." PIER Working Paper No. 11-005.

Dupas, Pascaline. 2010. "Do teenagers respond to HIV risk information? Evidence from a field experiment in Kenya." *American Economic Journal: Applied Economics* 3(1): 1-34.

Fonner, Virginia A., Julie Denison, Caitlyn E. Kennedy, Kevin O'Reilly, and Michael Sweat. 2012. "Voluntary counseling and testing (VCT) for changing HIV-related risk behavior in developing countries." *Cochrane Database of Systematic Reviews*, Issue 9.

Gersovitz, Mark. 2010. "HIV testing: principles and practice." *World Bank Research Observer* 26: 1-41.

Goldstein, Markus, Joshua G. Zivin, James Habyarimana, Cristian Pop-Eleches, and Harsha Thirumurthy. "The Effect of Absenteeism and Clinic Protocol on Health Outcomes: The Case of Mother-to-Child Transmission of HIV in Kenya." *American Economic Journal: Applied Economics*, 5(2): 58-85.

Gong, Erick. 2013. "HIV testing and risky sexual behavior." sites.google.com/site/erickgong/. [Accepted, The Economic Journal].

Haile, Beliyou A. 2011. "Promoting HIV testing and safe sexual behavior: evidence from a field experiment in Ethiopia." Job market paper, Department of Economics, Columbia University.

Halder, Amal K., Carole Tronchet, Shamima Akhter, Abbas Bhuiya, Richard Johnston, Stephen P. Luby. 2010. "Observed hand cleanliness and other measures of handwashing behavior in rural Bangladesh." *BMC Public Health* 10: 545.

Jensen, Rob. 2010. The (Perceived) Returns to Education and the Demand for Schooling. *Quarterly Journal of Economics* 125(2): 515-548.

Kamb, Mary L., Martin Fishbein, John M. Douglas, Jr, Fen Rhodes, Judy Rogers, Gail Bolan, Jonathan Zenilman, Tamara Hoxworth, C. Kevin Malotte, Michael Iatesta, Charlotte Kent, Andrew Lentz, Sandra Graziano, Robert H. Byers, Thomas A. Peterman, for the Project RESPECT Study Group, 1998. "Efficacy of risk reduction counseling to prevent human immunodeficiency virus and sexually transmitted diseases: a randomized controlled trial." *JAMA* 280(13): 1161-1167.

Kerwin, Jason T. 2012. "'Rational fatalism': non-monotonic choices in response to risks." Paper presented at the Working Group in African Political Economy meeting, University of California, Berkeley, CA.

Kerwin, Jason, Rebecca Thornton, and Sallie Foley. 2012. "Prevalence of and factors associated with oral sex among rural and urban Malawian men." Paper presented at the Population Association of America Annual Meeting

Kremer, Michael. 1996. "Integrating behavioral choice into epidemiological models of the AIDS epidemic." *Quarterly Journal of Economics* 111(2): 549-573.

Lee, David. 2009. "Training, wages, and sample selection: Estimating sharp bounds on treatment effects." *Review of Economic Studies* 76(3): 1071-1102.

Lin, Winston. 2013. "Agnostic Notes on Regression Adjustments to Experimental Data: Reexamining Freedman's Critique." *The Annals of Applied Statistics* 7(1): 295-318.

Minnis, Alexandra M., Markus J. Steiner, Maria F. Gallo, Lee Warner, Marcia M. Hobbs, Ariane van der Straten, Tsungai Chipato, Maurizio Macaluso, and Nancy S. Padian, 2009. "Biomarker validation of reports of recent sexual activity: results of a randomized controlled study in Zimbabwe." *American Journal of Epidemiology* 170(7): 918-924.

National Statistical Office (Malawi) and ORC Macro. 2004. Malawi Demographic and Health Survey 2004. Maryland: NSO and ORC Macro.

Nguyen, Trang. 2008. "Information, role models and perceived returns to education: Experimental evidence from Madagascar." Unpublished.

Oster, Emily, Ira Shoulson, and E. Ray Dorsey, 2013. "Optimal expectations and limited medical testing: evidence from Huntington disease." *American Economic Review*, 103(2): 804-830.

Padian, Nancy S., Sandra I. McCoy, Shanti Manian, David Wilson, Bernhard Schwartländer, Stefano M. Bertozzi, 2011. "Evaluation of Large-Scale Combination HIV Prevention Programs: Essential Issues." *Journal of Acquired Immune Deficiency Syndrome* 58(2): e23-e28.

Potts, Malcolm, Daniel T. Halperin, Douglas Kirby, Ann Swidler, Elliot Marseille, Jeffrey D. Klausner, Norman Hearst, Richard G. Wamai, James G. Kahn, Julia Walsh, 2008. "Reassessing HIV prevention." *Science* 320: 749-750.

Poulin, Michelle. 2007. "Sex, money, and premarital partnerships in Southern Malawi." *Social Science & Medicine*, 65: 2383-2393.

Ram, Pavani K., Amal K. Halder, Stewart P. Granger, Therese Jones, Peter Hall, David Hitchcock, Richard Wright, Benjamin Nygren, M. Sirajul Islam, John W. Molyneaux, and Stephen P. Luby, 2010. "Is structured observation a valid technique to measure handwashing behavior? Use of acceleration sensors embedded in soap to assess reactivity to structured observation." *American Journal of Tropical Medicine and Hygiene* 83(5): 1070-1076.

Swidler, Ann and Susan Watkins. 2007. "Ties of dependence: AIDS and transactional sex in rural Malawi." *Studies in family planning* 38 (3): 147-162.

The Voluntary HIV-1 Counseling and Testing Efficacy Study Group. 2000. "Efficacy of voluntary HIV-1 counseling and testing in individuals and couples in Kenya, Tanzania, and Trinidad: a randomized trial." *Lancet* 356: 103-112.

Thornton, Rebecca L. 2008. "The demand for, and impact of, learning HIV status." *American Economic Review* 98(5): 1829-1863.

Thornton, Rebecca L. 2012. "HIV testing, subjective beliefs and economic behavior." *Journal of Development Economics* 99: 300-313.

UNAIDS, 2008 Report on the global HIV/AIDS epidemic (UNAIDS/08.25E/JC1510E; http://whqlibdoc.who.int/unaids/2008/9789291737116_eng.pdf)

U.S. Preventive Services Task Force. 2012. "Screening for HIV: U.S. Preventive Services Task Force draft recommendation statement." AHRQ Publication No. 12-05173-EF-3.

Weinhardt, L.S., M. P. Carey, B. T. Johnson, and N. L. Bickham. 1999. "Effects of HIV counseling and testing on sexual risk behavior: A meta-analytic review of published research, 1985-1997." *American Journal of Public Health* 89(9): 1397-1405.

World Development Indicators, 2010, Database. Accessed November 2010 (http://data.worldbank.org/data-catalog/world-development-indicators/wdi-2010).

Yeatman, Sara E., 2009. "The Impact of HIV Status and Perceived Status on Fertility Desires in Rural Malawi," *AIDS Behavior* 13: S12-S19.

Appendix A: Home Based HIV Testing and Counseling (HTC) Details

Between July and September, 2009, 52 of the 88 EAs in the comparison group of the larger experiment were randomly sampled for inclusion in the data collection for biological outcomes. Study participants in these EAs were visited at their homes and invited to receive counseling and rapid testing for HIV, HSV-2, and syphilis by a trained counselor.[30] Then, between the months of March and August 2010, i.e. after an average of approximately 10 months, participants in all 88 EAs were visited at their homes and invited to receive counseling and rapid testing for HIV and HSV-2. All participants provided written informed consent. Additional consent was obtained from parents or legal guardians of all unmarried girls under the age of 18, with assent obtained from the girls.

HTC was conducted by Malawian nurses and counselors certified in conducting rapid HIV tests through the Ministry of Health HIV Unit HCT Counselor Certification Program. The counselors were not the same individuals who conducted the survey questionnaires. Testing and counseling were performed in a private location at or near the participant's house. Whole blood samples were obtained using a finger-stick. For HIV, the testing algorithm described in the Malawi Ministry of Health guidelines was followed. The collected blood sample was first tested using a Determine HIV/1-2^{TM} assay (Inverness Medical, UK). If the test result was positive, then the sample was tested using a Uni-GoId® HIV assay (Trinity Biotech, Ireland). If these two test results were discordant, then the sample was tested with SD BIOLINE HIV 1/2 3.0 assay (Standard Diagnostics, Inc., Korea) for a tie-breaker. Participants who tested positive on the Determine plus either of the other two assays were interpreted to be HIV infected; those testing negative on the Determine or negative on the other two assays were interpreted to be HIV uninfected. Recognizing that HIV-negative participants could be in the window period, they were counseled to be retested after 3 months if they had any risk factors for infection.

For HSV-2, we used the **Biokit**HSV-2 Rapid Test assay (Biokit USA, USA) following the manufacturer's instructions. The Biokit assay is a point-of-care test that is easy to use in the field and produces results in minutes; however the results are dependent on the subjective judgment of the tester. Therefore, all tests were reviewed by at least two readers. The tests were

---

[30] Prevalence rates for syphilis were below 1% during the initial testing, thus syphilis was not included during the second round of HTC.

read immediately by the tester who also photographed the test kit with a digital camera, and a second study staff member interpreted the results from the photos blinded to the tester's interpretation. When the readers disagreed, the testing supervisor reviewed the photos to break the ties.

Appendix B: Tables

Appendix Table B1: Attrition			
	=1 if Have Baseline Survey Data	=1 if Have Survey Data and HIV Data	=1 if Have Survey Data, HIV data, and all Controls
=1 if Respondent in Round 2	0.031	0.005	0.007
Sampled HTC EA	(0.020)	(0.027)	(0.027)
Control mean	0.906	0.849	0.852
Number of observations	1948	1948	1948

Notes: Regressions are OLS models with robust standard errors clustered at the EA level. All regressions are weighted to make them representative of the target population in the study EAs. Parameter estimates statistically different than zero at 99% (***), 95% (**), and 90% (*) confidence.

Appendix Table B2: Baseline Balance Among the HIV-Positive Sample

	Control	HTC	p-value (control-HTC)
Age	17.748	18.600	0.351
	(2.716)	(2.869)	
=1 if Live Inside 16km	0.629	0.445	0.366
	(0.489)	(0.505)	
=1 if Live Outside 16km	0.104	0.060	0.631
	(0.309)	(0.241)	
=1 if Live in Urban Area	0.267	0.495	0.279
	(0.448)	(0.508)	
Asset Index	0.027	0.319	0.639
	(2.451)	(2.352)	
=1 if Female Headed Household	0.792	0.671	0.305
	(0.411)	(0.477)	
=1 if in School	0.531	0.641	0.372
	(0.505)	(0.487)	
Highest Grade	8.911	8.834	0.898
	(2.357)	(1.979)	
=1 if Ever Married	0.125	0.100	0.719
	(0.334)	(0.304)	
=1 if Ever Pregnant	0.427	0.344	0.502
	(0.501)	(0.482)	
=1 if Never Had Sex	0.217	0.368	0.250
	(0.417)	(0.490)	
=1 if Sexually Active in Past 12 months	0.344	0.455	0.370
	(0.481)	(0.506)	
Number of Partners in Past 12 Months	0.365	0.475	0.388
	(0.529)	(0.546)	
=1 if Engage in Risky Sex	0.303	0.330	0.811
	(0.465)	(0.477)	
=1 if Partner Over 25	0.021	0.165	0.013
	(0.144)	(0.377)	
=1 if Any Chance Infected with HIV Now	0.181	0.325	0.272
	(0.390)	(0.476)	
Number of observations	40	33	

Notes: Column (1) shows the baseline means for the control group while column (2) shows the means for the tested group. Column (3) presents the p-value on the difference between the two means. All variables are weighted to make them representative of the target population in the study EAs with robust standard errors clustered at the EA level to calculate the p-value.

Appendix Table B3: Baseline Balance Among the HIV-Negative Sample

	Control	HTC	p-value (control-HTC)
Age	16.592	16.630	0.811
	2.172	2.133	
=1 if Live Inside 16km	0.603	0.581	0.879
	0.490	0.494	
=1 if Live Outside 16km	0.091	0.097	0.910
	0.288	0.296	
=1 if Live in Urban Area	0.306	0.322	0.918
	0.461	0.467	
Asset Index	0.187	0.425	0.549
	2.429	2.606	
=1 if Female Headed Household	0.413	0.355	0.096
	0.493	0.479	
=1 if in School	0.761	0.810	0.129
	0.427	0.393	
Highest Grade	8.149	8.392	0.149
	1.966	1.979	
=1 if Ever Married	0.089	0.073	0.417
	0.284	0.261	
=1 if Ever Pregnant	0.167	0.162	0.852
	0.373	0.369	
=1 if Never Had Sex	0.595	0.651	0.133
	0.491	0.477	
=1 if Sexually Active in Past 12 months	0.228	0.208	0.526
	0.420	0.406	
Number of Partners in Past 12 Months	0.233	0.221	0.694
	0.437	0.452	
=1 if Engage in Risky Sex	0.180	0.165	0.583
	0.384	0.371	
=1 if Partner Over 25	0.036	0.029	0.462
	0.186	0.168	
=1 if Any Chance Infected with HIV Now	0.074	0.078	0.772
	0.262	0.269	
Number of observations	680	927	

Notes: Column (1) shows the baseline means for the control group while column (2) shows the means for the tested group. Column (3) presents the p-value on the difference between the two means. All variables are weighted to make them representative of the target population in the study EAs with robust standard errors clustered at the EA level to calculate the p-value.

Appendix Table B4: Baseline Balance Among HIV-Positive
Individuals with Low Prior Beliefs of HIV Infection Before HTC

	Control	HTC	p-value (control-HTC)
Age	18.418	18.056	0.699
	(2.682)	(3.003)	
=1 if Live Inside 16km	0.711	0.509	0.301
	(0.461)	(0.511)	
=1 if Live Outside 16km	0.124	0.090	0.758
	(0.336)	(0.293)	
=1 if Live in Urban Area	0.165	0.401	0.208
	(0.377)	(0.501)	
Asset Index	-0.625	-0.137	0.478
	(2.169)	(2.381)	
=1 if Female Headed Household	0.751	0.635	0.399
	(0.440)	(0.492)	
=1 if in School	0.533	0.608	0.588
	(0.508)	(0.499)	
Highest Grade	8.605	8.664	0.932
	(2.527)	(2.078)	
=1 if Ever Married	0.156	0.090	0.446
	(0.369)	(0.293)	
=1 if Ever Pregnant	0.404	0.271	0.288
	(0.499)	(0.454)	
=1 if Never Had Sex	0.156	0.293	0.315
	(0.369)	(0.465)	
=1 if Sexually Active in Past 12 months	0.249	0.500	0.073
	(0.440)	(0.511)	
Number of Partners in Past 12 Months	0.280	0.530	0.099
	(0.523)	(0.568)	
=1 if Engage in Risky Sex	0.218	0.341	0.347
	(0.420)	(0.484)	
=1 if Partner Over 25	0.031	0.190	0.136
	(0.177)	(0.401)	
Number of observations	29	24	

Notes: Column (1) shows the baseline means for the control group while column (2) shows the means for the tested group. Column (3) presents the p-value on the difference between the two means. All variables are weighted to make them representative of the target population in the study EAs with robust standard errors clustered at the EA level to calculate the p-value.

Appendix Table B5: Baseline Balance Among HIV-Negative
Individuals with High Prior Beliefs of HIV Infection Before HTC

	Control	HTC	p-value (control-HTC)
Age	17.578	17.645	0.869
	(2.489)	(2.395)	
=1 if Live Inside 16km	0.705	0.613	0.531
	(0.460)	(0.490)	
=1 if Live Outside 16km	0.147	0.093	0.513
	(0.357)	(0.292)	
=1 if Live in Urban Area	0.148	0.295	0.281
	(0.358)	(0.459)	
Asset Index	-0.164	-0.086	0.881
	(2.518)	(2.343)	
=1 if Female Headed Household	0.468	0.390	0.406
	(0.503)	(0.491)	
=1 if in School	0.690	0.691	0.987
	(0.466)	(0.465)	
Highest Grade	8.077	8.541	0.277
	(2.337)	(2.452)	
=1 if Ever Married	0.169	0.093	0.227
	(0.378)	(0.292)	
=1 if Ever Pregnant	0.254	0.206	0.551
	(0.439)	(0.407)	
=1 if Never Had Sex	0.432	0.479	0.620
	(0.499)	(0.503)	
=1 if Sexually Active in Past 12 months	0.350	0.362	0.875
	(0.481)	(0.484)	
Number of Partners in Past 12 Months	0.350	0.383	0.680
	(0.481)	(0.530)	
=1 if Engage in Risky Sex	0.254	0.271	0.806
	(0.439)	(0.447)	
=1 if Partner Over 25	0.070	0.081	0.819
	(0.258)	(0.274)	
Number of observations	64	79	

Notes: Column (1) shows the baseline means for the control group while column (2) shows the means for the tested group. Column (3) presents the p-value on the difference between the two means. All variables are weighted to make them representative of the target population in the study EAs with robust standard errors clustered at the EA level to calculate the p-value.

Appendix C: Bounding the estimates of HTC impacts by HIV serostatus

Let Y_i^T indicate the HSV-2 biomarker outcome at follow-up for an HIV-tested individual and Y_i^C for an HIV-untested individual. Then, define three exhaustive and mutually exclusive groups based on HIV status in a previously untested population. These are $H_i = 0$, those who are HIV-negative in both rounds, $H_i = 1$, those who are HIV-positive in both rounds, and $H_i = \tilde{S}$, those who HIV seroconvert in between the two rounds if no testing occurred (throughout this appendix we use a tilde to indicate quantities that are unobserved). Each of these three groups has a population prevalence, where $p^0 + p^1 + p^S = 1$, and by the properties of randomization and the fact that strata are defined in the absence of testing, these quantities are the same in the testing treatment and control. We observe p^1 as the prevalence rate in the treatment in the baseline, p^0 as one minus the prevalence rate in the control at follow-up, and p^S as the difference between these two quantities.

The quantity that we wish to estimate is the causal effect of an HIV-positive test on the HSV-2 prevalence rate, which would be given by:[31]

$$ITT^{1*} \equiv E(Y_i^T \mid H_i = 1) - E(\tilde{Y}_i^C \mid H_i = 1)$$

While we can calculate the fraction of the sample p^S that seroconverted in the control group, we cannot tell which specific (HIV-positive at follow-up) individuals belong to this group and hence we cannot separately measure the quantities $E(\tilde{Y}_i^C \mid H_i = 1)$ or $E(\tilde{Y}_i^C \mid H_i = S)$, only the HSV-2 prevalence in these combined groups $E(Y_i^C \mid H_i \in \{S,1\})$.

Therefore, what we are able to measure directly is the 'Estimated' $\widehat{ITT}^{+E}$:

---

[31] The estimation of bounds for the causal effect of an HIV-negative test is similar and not detailed here.

$$\widehat{ITT}^{+E} \equiv E(Y_i^T \mid H_i = 1) - E(Y_i^C \mid H_i \in \{1, S\}) = ITT^{+*} + bias$$

$$= \frac{p^1 \left[E(Y_i^T \mid H_i = 1) - E(\tilde{Y}_i^C \mid H_i = 1) \right] + p^S \left[E(Y_i^T \mid H_i = 1) - E(\tilde{Y}_i^C \mid H_i = S) \right]}{p^1 + p^S}$$

$$= \frac{p^1 ITT^{+*} + p^S \left[E(Y_i^T \mid H_i = 1) - E(\tilde{Y}_i^C \mid H_i = S) \right]}{p^1 + p^S}$$

This shows that the resulting estimand Fis a weighted average of the correct ITT and the incorrect estimate that we would get by subtracting the outcome among seroconverters from the outcome among tested always-positives. This expression will equal the correct ITT either if $p^S = 0$ (there are no seroconverters) or if $E(\tilde{Y}_i^C \mid H_i = 1) = E(\tilde{Y}_i^C \mid H_i = S)$ (the seroconverters look just like the control always-positives). Hence, this expression gives us a way to understand how this imperfect counterfactual distorts the quantity that we estimate relative to the true ITT.

We can do this by taking advantage of the binary nature of Y_i. Specifically, we can plug in the extreme possible values of prevalence of HSV-2 in the follow-up for the seroconverters, and rearrange the above expression to calculate the extreme bounds on the possible values of the true ITT given the other parameters:

$$ITT^{+*} \in \left[\frac{\widehat{ITT}^{+E} \left(p^1 + p^S \right) - p^S E(Y_i^T \mid H_i = 1)}{p^1}, \quad \frac{\widehat{ITT}^{+E} \left(p^1 + p^S \right) - p^S \left(E(Y_i^T \mid H_i = 1) - 1 \right)}{p^1} \right]$$

Clearly, as $p^S \to 0$ these bounds converge to the estimated value, and $ITT^{+E} \to ITT^{+*}$. This technique is the analogue to Lee bounds (Lee 2009), where rather than facing attrition we face the addition of a new group whose outcomes may be different from those within the counterfactual we wish to form. These bounds can be formed with no assumptions beyond simple random assignment.

Appendix D: Non-Parametric Permutation Test

We follow Anderson (2008), and implement a non-parametric permutation test. Starting with our HIV-positive sample (N=73), we randomly assign HTC treatment status to 40 of the HIV-positive individuals, and the remaining 33 individuals are assigned to the control group. Using this sample, we estimate equation 1 and collect the t-statistic for $\hat{\beta}$ (the effect of an HIV-positive test on HSV-2). We repeat this 10,000 times and generate a distribution of t-statistics from which we based our statistical inference on. Using this distribution and our original t-statistic from our estimation (t-statistic = 3.58) we calculate our p-value, which is less than .01. This allows us to reject the null hypothesis of no effect of HTC on HSV-2 at the 1% level without relying on large sample asymptotic properties.