

Antiretroviral Therapy and Demand for HIV Testing: Evidence from Zambia *

Nicholas Wilson

Department of Economics, Williams College, Williamstown, MA 01267, United States[†]

October 2010

Abstract

This paper examines the effects of antiretroviral therapy (ART) on demand for HIV testing and of ART-induced testing on demand for risky sexual behavior. I provide a model of sexual behavior decision-making under uncertainty and estimate the structural parameters of the model using nationally representative survey data from Zambia on HIV testing decisions before and after the introduction of ART. The empirical results indicate that although the introduction of ART increased demand for HIV testing, the ART allocation process limited the prevention benefit of ART-induced testing. Simulation results show that eliminating this prevention inefficiency while holding the supply of ART constant would increase the prevention impact of ART-induced testing more than four-fold. More generally, the analysis indicates that existing studies which examine “universal” testing or quasi-experimental testing programs understate the efficacy of standard voluntary counseling and testing programs.

JEL classification: D45; I18; O12

Keywords: HIV/AIDS; Beliefs; Selection; Rationing; Zambia

*I would like to thank Quamrul Ashraf, Daniel Bennett, Aureo de Paula, Andrew Foster, Jason Goldman, Emily Hazan, Vernon Henderson, Alaka Holla, David Love, Mark Lurie, Kaivan Munshi, Peter Pedroni, Ashok Rai, David Weil, and participants at the Conference on Infectious Diseases in Poor Countries at Cornell University for many excellent comments. This research would not be possible without the cooperation of Bates Bucker at MEASURE Evaluation and many individuals in Zambia, including Dr. Albert Mwangi at the Ministry of Health and Nchimunya Nkombo at the Central Statistical Office. All errors are my own. The findings, interpretations, and conclusions expressed in this paper are those of the author and do not necessarily represent the views of the aforementioned individuals or the agencies that employ them.

[†]Tel.: +1 413 597 4868 Fax: +1 413 597 4045 E-mail: nlw3@williams.edu

1 Introduction

Policymakers believe that HIV testing is an important intervention in the HIV/AIDS pandemic. Proponents of HIV testing presume that on average individuals who learn whether or not they are infected will respond to this new information by reducing risky behavior. However, HIV testing rates are low in much of the world and a vast majority of HIV positive individuals are not aware that they are infected. For example, among the 67 percent of the world population infected with HIV/AIDS that reside in Sub-Saharan Africa only 10 percent know their HIV status (UNAIDS 2008, WHO 2006a).

One of the main mechanisms by which policymakers hope to increase demand for HIV testing is increased availability of subsidized antiretroviral therapy (Global HIV Prevention Working Group 2004). The availability of antiretroviral therapy (ART) gives individuals a direct incentive to take a HIV test: if they take a HIV test and are HIV positive then they may begin therapy, reducing morbidity and prolonging life. However, the incentive effects of ART may be heterogeneous and many individuals may not choose to test in response to the introduction of ART. Moreover, those who test because of ART may not demonstrate substantial behavior change subsequent to testing.

This paper examines the effect of ART availability on demand for HIV testing and the effect of ART-induced testing on demand for risky behavior. I present a model of demand for HIV testing and for risky behavior. Initially I assume that the conditions under which individuals have an incentive to test and testing reduces risky behavior hold. In particular, I assume that demand for risky behavior is concave in the prior probability of being HIV positive. Under these conditions, I show that in the pre-ART era it is the riskiest individuals (i.e., individuals with the highest prior probabilities of being HIV positive) who have the greatest incentive to test in the setting examined in the current analysis. The riskiest individuals are also those individuals whose testing decisions should be most responsive to the introduction of ART. I show that under these same conditions the expected reduction in risky behavior associated with taking a HIV test in this setting should be increasing in the prior probability an individual is HIV positive. Thus, ART-induced testing has the potential to substantially reduce the spread of HIV/AIDS.

Using newly assembled data from before and after the introduction of ART in Zambia I examine the empirical evidence on these predictions. Testing behavior among women and men prior to the introduction of ART is consistent with the prediction that individuals with the highest prior probabilities of being HIV positive are those who should be most likely to test. The change in testing behavior among women after the introduction of ART is

consistent with the prediction that individuals with the highest prior probabilities of being HIV positive are those who should demonstrate the greatest response to the availability of ART. However, the change in testing behavior among men after the introduction of ART is not consistent with this prediction. Specifically, testing decisions among older men appear to be particularly responsive to the introduction of ART despite the fact that older men are one of the lowest HIV prevalence groups in Zambia. In contrast, testing decisions among men in the middle of the age distribution are not particularly responsive to the introduction of ART despite the fact that these men are one of the higher HIV prevalence groups. Although data on the age distribution of ART patients in Zambia are not available, I interpret these results as evidence of a non-random rationing mechanism determining the allocation of ART among HIV positive individuals (in favor of older males) and provide evidence rejecting alternative hypotheses. In any case, from the perspective of maximizing the prevention impact of ART-induced testing, the results indicate an inefficiency in the process determining who seeks and receives ART.

I estimate the structural parameters of the model and show that the conditions for individuals to have an incentive to test and for testing to reduce risky behavior are indeed satisfied in this setting. Moreover, the structural parameter estimates allow me to simulate the effect of ART availability on testing demand and the effects of ART-induced testing on demand for risky behavior and on the spread of HIV/AIDS. As part of the simulations, I examine the effects of eliminating the prevention inefficiency in the process determining the allocation of ART and/or expanding the supply of ART. Simulation results show that under the existing policy ART availability increased testing demand by approximately 3 percentage points and ART-induced testing reduced the incidence of HIV by less than 2 percent. Expanding the supply of antiretroviral drugs without eliminating the prevention inefficiency would only have moderate effects on testing rates and risky behavior. In contrast, eliminating the prevention inefficiency while holding fixed the existing supply of antiretroviral drugs would more than quadruple ART-induced testing and the number of new infections avoided due to ART-induced testing.

This analysis yields three broader insights about the economics of HIV/AIDS and health economics more generally. First, this paper shows that with a small amount of theoretical structure we can estimate the parameters of the risky behavior demand function without actually observing risky sexual behavior. Because self-reported sexual behavior is subject to substantial reporting bias (Gersovitz *et al* 1998), the indirect approach to estimating the risky behavior demand function implemented in the current analysis may be superior to a

direct approach.¹

Second, this paper provides a counterpoint to the argument that treatment for an infectious disease diminishes prevention efforts by reducing the private benefit of preventive behavior and, in the case of HIV/AIDS, by increasing vector activity. In contexts where uncertainty about one's own infection status is an important factor in decision-making about preventive behavior, treatment for an infectious disease may actually increase private prevention efforts.²

Third, this analysis shows that the selection induced by the voluntary, unpaid nature of truly voluntary counseling and testing (VCT) means that these programs yield a larger average effect of testing on risky sexual behavior than do "universal" or conditional cash transfer (CCT) approaches to HIV testing. Some have suggested that governments conduct door-to-door testing campaigns (e.g., Donnelly 2005) and/or implement an opt-out approach to testing whereby individuals are offered a HIV test during any visit to a health care provider (e.g., UNAIDS/WHO 2004, Centers for Disease Control and Prevention 2006, Granich *et al* 2009). However, the findings in Boozer and Philipson (2000) suggest that these universal testing campaigns may have little-to-no effect on aggregate risky behavior even if they are successful at getting individuals to take a HIV test. Similarly, the findings in Thornton (2008) suggest that inducing individuals to take a HIV test by offering a financial incentive (i.e., a conditional cash transfer approach to HIV testing) will generate little-to-no change in average risky sexual behavior.

In contrast, I show that if the risky behavior demand function is concave in the prior probability of being HIV positive, then getting everyone to test (or inducing testing among a random sample of the population) will yield less behavior change per person than an approach to testing where individuals simply self-select into testing. In the setting examined in the current analysis, the additional selection into testing induced by the introduction of

¹A related point is that instead of examining the effect of HIV testing on a single or a handful of risky behavior measures, I examine the effect of HIV testing on the cumulative risk of acquiring HIV.

²For the particular disease I examine in this analysis, HIV/AIDS, the net effect of treatment (i.e., ART) on preventive behavior is unclear. I show that ART-induced HIV testing increases preventive behavior. In contrast, Lakdawalla *et al* (2006) show that the introduction of highly-active antiretroviral therapy (HAART) in the United States increased the number of sex partners of the representative HIV positive individual and suggest that this was due to the improved health and longevity of HIV positive individuals receiving HAART. Although this is an important finding, the model I present here focuses on the effect of ART-induced testing on demand for risky behavior. Because Lakdawalla *et al* (2006) do not consider the role of ART-induced testing in behavior change and the current study does not incorporate the Lakdawalla *et al* (2006) mechanism, writing a more general model simultaneously allowing for these two effects may be a useful avenue for future research. However, it may be difficult to reliably estimate the structural parameters of an expanded model.

ART increases the prevention effects of HIV testing. In fact, truly voluntary counseling and testing (VCT) programs are the predominant approach to HIV testing and increasing access to ART remains a priority in HIV/AIDS policy in Sub-Saharan Africa and elsewhere in the developing world.

2 Setting

Zambia has one of the highest rates of HIV prevalence in the world. According to the anonymous HIV testing component of the 2001 Zambia Demographic and Health Survey (ZDHS), the HIV prevalence rate among adults age 15-49 is 14.5 percent. High HIV prevalence in Zambia is not a recent development. Data from antenatal clinics in the sentinel surveillance program are consistent with the hypothesis that HIV prevalence has remained relatively stable from the early 1990s onwards (UNAIDS/WHO 2006).

As in many Sub-Saharan African countries (i.e., high prevalence countries), the highest rates of HIV prevalence are among individuals in the middle of the age distribution. Figure 1 presents the results of a locally weighted smoothed regression of an indicator variable for HIV status on age, estimated separately for females and for males in the 2001 ZDHS. As shown in Figure 1, HIV prevalence is highest among individuals in the middle of the age distribution, prevalence is slightly higher among women than among men, and the age of peak prevalence among females is less than that among males (34 and 40 years, respectively). The average age difference in sexual partners may explain this last fact. Data from the Zambia Sexual Behavior Survey (ZSBS) on the age of sexual partners show that on average males are approximately six years older than their female sexual partners.

[FIGURE 1 APPROXIMATELY HERE]

Magruder (2010) provides an explanation for the inverted U-shape that characterizes the prevalence-age profiles of most high HIV prevalence countries. Individuals in early adulthood engage in “marital shopping” (i.e., the search for a potential marital partner) and experience higher than average turnover in sexual partners. The effect of this behavior on prevalence is magnified by the biology of the HIV virus: infectiousness during the first month of having acquired HIV is much higher than during the majority of the life course of the virus (Wawer *et al* 2005) and rapid turnover of sexual partners during the period of having recently acquired HIV can lead to a sharp increase in incidence and prevalence.

As part of its efforts to combat the HIV/AIDS pandemic, the government of Zambia began offering free voluntary counseling and testing (VCT) in 1998. The VCT approach to

HIV testing, in which individuals have to choose to visit a health clinic to take a HIV test, is the predominant approach to HIV testing around the world. Although initially rates of HIV testing in Zambia were low, they are increasing. According to the ZSBS, the proportion of individuals voluntarily taking a HIV test in the 12 months prior to the survey round increased from between 5.5 and 6 percent in 2000 and 2003 to over 10 percent for women and nearly 8 percent for men in 2005.

One possible explanation for the increase in testing rates is that free antiretroviral therapy (ART) became available in Zambia around this time. The early 21st century witnessed a dramatic expansion in the provision of antiretroviral therapy in Sub-Saharan Africa. Between 2000 and the end of 2005, the number of individuals on ART in the region increased from fewer than 50,000 (Attawell and Mundy 2003) to 810,000 (WHO 2006b). Zambia is a participant in this expansion and in May 2004 the government of Zambia began to provide ART free of charge at primary health care centers (Stringer *et al* 2006).

The introduction of free antiretroviral therapy in Zambia may have increased testing demand by providing those individuals who believed they may have been HIV positive with an added incentive to test.³ However, other factors affecting testing rates could have changed during the period as well. For example, the number of VCT sites increased between 2000-03 and 2005.

If the introduction of ART had a causal effect on testing demand, then we should expect to see an inverse relationship between distance to the nearest ART site and the change in testing rates between 2000-03 and 2005. The ZSBS includes information on the Standard Enumeration Area (SEA) of residence of each respondent.⁴ In addition, the Japanese International Cooperation Agency 2004 Health Facilities Census (HFC) provides global positioning system (GPS) information for each of the ART sites open at the time of the 2005 survey round of the ZSBS. For the 2000, 2003, and 2005 survey rounds of the ZSBS, I cal-

³Although not all HIV positive individuals were clinically eligible for ART, it seems unlikely that clinical eligibility conditioned the testing response to the introduction of free ART. World Health Organization guidelines for resource-limited settings recommend initiating ART after CD4 counts fall below 200 cells/mm³ (WHO 2006c). However, it is unlikely that many individuals in the population were aware of the clinical guidelines for ART eligibility and even more unlikely that individuals had any precise sense of their CD4 count prior to testing positive for HIV and having their CD4 count taken. Moreover, even those individuals who knew the clinical guidelines and had a precise belief about their CD4 count would have an added incentive to test after ART was introduced if they thought they might be HIV positive. Testing positive would initiate a regular series of CD4 tests, allowing the individual to begin ART as soon as clinically possible (or at least take a place in the queue).

⁴The Standard Enumeration Area (SEA) is the smallest administrative unit in the Zambian Census. There are more than 15,000 SEAs in Zambia.

culate the centroid distance from the SEA of residence for each respondent to the nearest ART site that was open at the time of the 2004 HFC.

Figure 2 presents the results of a locally weighted smoothed regression of an indicator variable equal to one if the individual voluntarily took a HIV test in the 12 months prior to the survey month on distance to the (eventual) location of the nearest ART site, estimated separately for the pre-ART and ART periods. In the pre-ART period, individuals closer to sites that eventually provided ART were more likely to test. Nonetheless, the increase in testing rates associated with the ART period is greatest for those individuals closest to ART sites. Of course an important omitted variable in this analysis is the distance to the nearest VCT site; proximity to a VCT site affects the testing decision and is positively correlated with proximity to an ART site. However, health clinics in which ART was introduced in 2004 and 2005 tended to be sites that had previously provided VCT services and the pre-ART regression line reflects this fact.⁵

[FIGURE 2 APPROXIMATELY HERE]

Although it appears that the introduction of subsidized ART caused an increase in aggregate testing demand, there exists substantial heterogeneity across demographic groups in the change in testing demand associated with this intervention. Figure 3 shows the testing-age profile for women prior to the availability of ART and after the introduction of ART. Prior to the availability of ART, women in the middle of the age distribution were most likely to take a HIV test. These women were also those with the highest HIV prevalence. After ART was introduced in Zambia testing increased among women of all ages, but the increase was largest for the women in the middle of the age distribution. These differences by age are consistent with the idea that ART availability increased testing demand and the incentive effect of ART on testing demand is greatest for those individuals with the highest probability of being HIV positive.⁶

[FIGURE 3 APPROXIMATELY HERE]

⁵The expansion in VCT between the pre-ART and ART periods occurred largely in areas that were not close to ART sites and hence may explain the increase in testing rates among individuals further from ART sites in 2005.

⁶Although an expansion in the prevention-of-mother-to-child (PMTCT) program paralleled the introduction of free ART in Zambia, this expansion does not fully explain the results presented in Figure 3. Disaggregating the analysis by whether the respondent reports being pregnant in the 12 months prior to the survey month does not substantially affect the results.

For males, the pattern is noticeably different. Figure 4 shows the testing-age profile for men prior to the availability of ART and after the introduction of ART. As with females, in the “pre” period it is men in the middle of the age distribution that were most likely to take a HIV test. Likewise, these men were also those with the highest HIV prevalence. In the “post” period, however, it is the oldest males that are most likely to take a HIV test. It appears that the idea that individuals with the highest probability of being HIV positive should have testing decisions that are most responsive to ART availability cannot explain the observed patterns of testing demand among males.

[FIGURE 4 APPROXIMATELY HERE]

Although much of the change between 2000-03 and 2005 in the testing-age profiles for females and for males may be attributed to the introduction of ART, it may be that the explanation for the large increase in testing rates among older males (and perhaps that among women in the middle of the age distribution) lies elsewhere. One change that paralleled the introduction of ART in Zambia was a continued expansion in the number of VCT sites. Nonetheless, it is unlikely that the expansion in the number of testing sites caused the observed changes in the testing-age profiles for women and for men. For the expansion in VCT sites to explain these changes, the expansion should have decreased the distance to the nearest VCT site differentially by age. Furthermore, the age differential in the expansion intensity should have differed by gender.

The ideal test of this hypothesis would be to semiparametrically regress the change in the distance to the nearest VCT site on measures of each individual’s age, gender, and interactions thereof. Unfortunately, data on which of the VCT sites open in 2005 existed prior to the introduction of ART in Zambia do not exist. In lieu of such data, we may examine the relationship between distance to the nearest VCT site and age in 2005. If the change in the distance to the nearest VCT site is correlated with age and gender, then we should expect to see evidence of spatial clustering by age/gender and possibly evidence of substantial differences by age in distance to the nearest VCT site in 2005. The HFC includes GPS information for VCT sites and for each respondent in the 2005 ZSBS I construct a measure of the distance to the nearest VCT site. Figure 5 presents the results of a locally weighted smoothed regression of distance to the nearest VCT site on age, estimated separately for females and males. There are small differences by age and gender in the average distance to the nearest VCT site, but these differences do not readily explain the changes in the testing-age profiles (e.g., older males and males in the middle of the age distribution appear

to reside at similar distances from VCT sites). Moreover, the overall flatness of the distance-age profiles suggests that individuals are not spatially clustered by age. Thus, it appears that the expansion in the number of VCT sites between 2003 and 2005 does not explain the large increase in testing rates among older males.

[FIGURE 5 APPROXIMATELY HERE]

A plausible explanation for the large testing response among older males associated with the introduction of ART is that there exists a non-random rationing mechanism for allocating ART. The supply of antiretroviral drugs in Zambia is insufficient to ensure that all individuals who are HIV positive and require ART are guaranteed coverage.⁷ If supply-side constraints mean that priority for access to ART depends on an individual's age and gender or factors correlated with these characteristics, then we would expect to see the observed changes in testing-age profiles deviate from those predicted by a simple model of testing demand where the incentive effect of ART is increasing in the prior probability that the individual is HIV positive.⁸

Social norms in Zambia are consistent with the notion that older males receive priority for access to ART. Although there may exist variation across the more than seventy ethnolinguistic groups in Zambia, in much of traditional Zambian society younger men are subordinate to older men (Bond 1982). The oldest resident male is generally identified as the household head (Scudder 1962). Among the Tonga ethnic group, one of the largest in Zambia, an elder is called *mupati* or "big person" (Colson and Scudder 1981). Among the

⁷Rosen *et al* (2005) examine ART coverage, defined as the ratio of the number of individuals medically eligible for ART therapy to those receiving ART, as of December 2004 in nine Sub-Saharan African countries. Estimated ART coverage ranges from two percent in Nigeria to fifty percent in Botswana. Estimated ART coverage in Zambia, 13 percent, is slightly higher than the median coverage rate in these nine countries. To the extent that not all individuals who were eligible for ART had visited a health clinic and taken a HIV test, this figure overstates the extent of excess demand for ART in Zambia at the time of the 2005 ZSBS. Nonetheless, rough calculations presented in the simulations section of the current analysis suggest that well below 100 percent of individuals who were eligible for ART and took a HIV test in this period were able to receive ART.

⁸Evidence from other contexts indicates that the rationing of scarce medical care is often based on demographic characteristics or factors correlated with these characteristics. McGough *et al* (2005) reviews four major technological developments in health care - insulin for diabetes, penicillin, hemodialysis, and liver transplants - and the associated rationing policies. In all four cases, rationing policies used demographic criteria, or criteria correlated with demographic characteristics (e.g., military service). Moreover, hemodialysis rationing explicitly favored prime-age workers, heads of household, and active participants in community affairs.

Bemba ethnic group, another of the largest ethnic groups in Zambia, males are not considered to be fully adult members of the community until they are married and have children. Older males, or *bakalamba*, control decision-making not just within their kin group, but also more broadly within the community. They are figures of authority and largely control village decision-making (Epstein 1981).

Although scant, direct evidence on the process determining the allocation of ART among HIV positive individuals is also consistent with the idea that older males receive priority in this process. Two of the main groups that are documented as receiving priority for access to ART are richer individuals and individuals with political connections (e.g., civil servants) (Jones 2004, Bwalya 2006). Despite the lack of direct evidence on age-based preferences, it is likely that older men represent a disproportionate share of these two groups.

Thus, although it appears that ART availability in Zambia increased testing demand, the change in demand is consistent with the existence of a non-random ART allocation mechanism. To frame the analysis of the effects of ART availability on demand for HIV testing and of ART-induced testing on demand for risky sexual behavior, the next section presents a simple model of demand for HIV testing and risky behavior. I then estimate the parameters of the model and use these to quantify the magnitude of these effects.

3 Model

The changes in testing behavior associated with the introduction of ART are consistent with the existence of a non-random rationing mechanism determining the allocation of ART among HIV positive individuals. In particular it appears that priority for access to ART may have been given to older males at the expense of males in the middle of the age distribution. To examine this finding more rigorously, I present a model of demand for HIV testing and for risky behavior. I match the model to the data and find that the model fits pre-ART testing behavior quite well but fails to match the change in testing behavior after the introduction of ART. Thus, I expand the model to capture the effect of ART rationing or, more generally, the effects of other factors correlated with demographic characteristics that may have conditioned the testing response to the introduction of ART. Then I use the parameter estimates to simulate the effect of ART availability on testing demand and demand for risky behavior under several ART policies.

3.1 Basic setup

In this model an individual chooses whether to take a HIV test and chooses how much risky behavior in which to engage. There are two components to expected utility: the pleasure the individual derives from engaging in risky behavior (e.g., unprotected sex) and the displeasure associated with acquiring HIV. Taking a HIV test means that the individual learns his HIV status with certainty and allows him to avoid *ex post* errors in forming beliefs about the marginal cost of risky behavior. In addition, when ART is available he may be able to begin treatment if he tests HIV positive.

The individual is endowed with a prior probability of being HIV positive, $p_0 \in [0, 1]$. Given this probability, the individual forms expectations about the net benefit of testing and decides whether to take a HIV test. After deciding whether or not to take a HIV test and learning the results if he chooses to test (and beginning treatment if he is HIV positive and eligible), the individual chooses an amount of risky behavior in which to engage. The amount of risky behavior in which the individual chooses to engage is summarized by the probability that he is HIV positive after engaging in this risky behavior if in fact he was HIV negative prior to this choice. Let $p_1 \in [0, 1]$ denote this probability.

3.2 Demand for risky sexual behavior

Figure 6 provides a graphical representation of the mechanics of the model. For the moment, consider the upper half of this figure. On the x-axis is the exogenous prior probability the individual is HIV positive, p_0 . On the y-axis is the choice of risky behavior, p_1 . The concave function, $p_1(p_0)$, represents the risky behavior demand function if the individual does not take a HIV test. Although the risky behavior demand function depicted in Figure 6 is increasing and concave in p_0 , this function need not be increasing nor concave in p_0 in practice and this model does not assume this to be the case. If it is increasing and concave in p_0 , then the model indicates a role for HIV testing in reducing risky behavior. I will explain why this is the case momentarily.

[FIGURE 6 APPROXIMATELY HERE]

Suppose the individual knows with certainty that he is HIV negative (i.e., $p_0 = 0$). Then he will choose an amount of risky behavior given by $p_1(0)$. This is the y-intercept for the $p_1(p_0)$ function. Now suppose that the individual knows with certainty that he is HIV positive (i.e., $p_0 = 1$). Then he will choose an amount of risky behavior given by $p_1(1)$. Note

that $p_1(0)$ and $p_1(1)$ are bounded below and above by zero and one but they need not equal these values.

In this model, taking a HIV test and receiving the results means that the individual knows his HIV status with certainty. Thus, if the individual takes a HIV test then he will choose one of two possible choices of risky behavior: $p_1(0)$ or $p_1(1)$. The result of the test determines which choice he will make. In expectation, the individual's choice of risky behavior if he were to take a HIV test is simply the linear combination of these two points: $(1 - p_0)p_1(0) + p_0p_1(1)$. This is the linear function in the upper half of Figure 6.

It should now be clear that the expected change in choice of risky behavior associated with taking a HIV test is the difference between the two functions plotted in the upper half of Figure 6. That is, the expected change in risky behavior associated with taking a HIV test is given by

$$(1 - p_0)p_1(0) + p_0p_1(1) - p_1(p_0). \quad (1)$$

If demand for risky behavior (i.e., $p_1(p_0)$) is concave in p_0 , then the sign of this expression is negative and in expectation HIV testing will reduce risky behavior. Conversely, if $p_1(p_0)$ is convex, then the sign of this expression is positive and in expectation HIV testing will increase risky behavior. To fix ideas, I will restrict the following analysis to the case where $p_1(p_0)$ is concave and in expectation HIV testing will reduce risky behavior.

Let p^* denote the value of p_0 that maximizes the expected reduction in risky behavior associated with taking a HIV test. For $p_0 < p^*$, the expected reduction in risky behavior associated with taking a HIV test is increasing in p_0 . For $p_0 \geq p^*$, the expected reduction in risky behavior associated with taking a HIV test is decreasing in p_0 .

3.3 Demand for HIV testing

Now consider the lower half of Figure 6. Again, on the x-axis is the exogenous prior probability the individual is HIV positive, p_0 . On the y-axis is expected utility. There are two components to expected utility in this model: the utility derived from engaging in (risky) sex and the utility loss associated with acquiring HIV. The utility derived from the choice of risky sexual behavior, $p_1(p_0)$, is denoted $V(p_1(p_0))$. The utility loss associated with acquiring HIV is denoted θ . I assume that utility is additively separable in choice of risky sexual behavior and health status.

The convex function in Figure 6, $E[U(\textit{not test})]$, is the expected utility of not taking a HIV test. Additive separability means that I may write the expected utility of not testing

as

$$E[U(\textit{not test})] = V(p_1(p_0)) - [1 - (1 - p_0)(1 - p_1(p_0))]\theta \quad (2)$$

where $[1 - (1 - p_0)(1 - p_1(p_0))]$ is the probability that the individual is HIV positive conditional on the prior probability of being HIV positive, p_0 , and on the probability that he acquires HIV through his choice of risky behavior, $p_1(p_0)$.

Although Figure 6 depicts the expected utility of not taking a HIV test as being decreasing and convex in p_0 this need not be the case in practice and this model does not assume this to be true. If it is decreasing in p_0 , then the increase in the expected utility cost of becoming HIV positive due to an increase in p_0 (and the subsequent choice to increase p_1) outweighs the utility gain (if any) from engaging in a greater amount of risky behavior, $p_1(p_0)$, that results from the increase in p_0 . The convexity of $E[U(\textit{not test})]$ follows from the fact that the marginal cost of additional risky behavior is decreasing in the prior probability of being HIV positive. As p_0 increases, the expected utility of not testing falls because the probability that the individual is HIV positive and incurs the associated utility cost is higher. On the other hand, the loss in expected utility from an increase in p_0 is partly offset by the increase in choice of risky behavior (given the current assumption that $p_1(p_0)$ is increasing in p_0). However, the increase in choice of risky behavior, p_1 , itself increases the probability that the individual acquires HIV. The magnitude of this latter effect (although always non-negative) is decreasing in p_0 and hence $E[U(\textit{not test})]$ is convex.

To understand the expected utility of testing in the absence of ART, first consider the case where the individual takes a HIV test and finds that he is HIV negative. Because he knows that he is HIV negative with certainty, he will choose to engage in an amount of risky behavior given by $p_1(0)$ and the only risk of acquiring HIV is because of this choice. Thus, conditional on testing HIV negative his expected utility is

$$V(p_1(0)) - p_1(0)\theta - c \quad (3)$$

where c is the utility cost of taking a HIV test. Now consider the case where the individual takes a HIV test and finds that he is HIV positive. Because he knows that he is HIV positive with certainty, he will choose to engage in an amount of risky behavior given by $p_1(1)$ and is certain of incurring the utility cost of acquiring HIV. Thus, conditional on testing HIV positive his expected utility is

$$V(p_1(1)) - \theta - c. \quad (4)$$

The expected utility of testing in the absence of ART is the linear combination of these two conditional expectations, where the weights are the prior probabilities of being HIV negative and of being HIV positive, respectively. That is, the expected utility of testing in the absence of ART, denoted $E[U(test)]$ in Figure 6, is

$$E[U(test)] = (1 - p_0)[V(p_1(0)) - p_1(0)\theta - c] + p_0[V(p_1(1)) - \theta - c]. \quad (5)$$

Assuming $V(p_1(0)) - p_1(0)\theta > V(p_1(1)) - \theta$ (i.e., the individual would prefer to engage in $p_1(0)$ and acquire HIV with probability $p_1(0)$ than to engage in $p_1(1)$ and acquire HIV with certainty), implies that $\frac{\partial E[U(test)]}{\partial p_0} < 0$.

With expressions for the expected utility of not testing and the expected utility of testing in the absence of ART, it is possible to examine predictions about the pre-ART relationship between the testing decision and the prior probability of being HIV positive. In the absence of ART, the model predicts that demand for HIV testing should be greatest among individuals in the middle of the p_0 distribution. As long as the cost of testing is sufficiently high, individuals with prior probabilities of being HIV positive that are either very low or very high will not choose to take a HIV test and the only individuals who test will be those in the middle of the p_0 distribution. The prediction is apparent in the lower half of Figure 6: the expected utility of testing is greater than that of not testing only for individuals in the middle of the p_0 distribution. The intuition underlying this result is that the expected change in risky behavior associated with taking a HIV test is relatively small for individuals with very high (or very low) values of p_0 . It is individuals who are in the middle of the p_0 distribution that have the largest expected change in demand for risky behavior associated with taking a HIV test and hence the greatest willingness to pay for the test in the absence of ART.

In this model the introduction of ART (without rationing) does not affect the risky behavior demand function or the expected utility of not testing, but it does affect the expected utility of testing. Specifically, the expected utility of testing increases by the prior probability that the individual is HIV positive times the net utility benefit of beginning ART conditional on being HIV positive. Let α denote the net utility of beginning ART conditional on being HIV positive. Then the expected utility of testing when there is ART, denoted $E[U(test)]$ *with ART* in Figure 6, is

$$E[U(test)] \text{ with ART} = (1 - p_0)[V(p_1(0)) - p_1(0)\theta - c] + p_0[V(p_1(1)) - \theta - c + \alpha]. \quad (6)$$

Suppose the individual takes a HIV test and finds that he is HIV negative. Given the set-up of the model, to receive the utility benefit of beginning ART the individual needs to be HIV positive before engaging in his choice of risky behavior, p_1 . Thus, conditional on testing HIV negative his expected utility when there is ART is identical to that if there were no ART (i.e., $V(p_1(0)) - p_1(0)\theta - c$). Now suppose the individual takes a HIV test and finds that he is HIV positive. Then the individual will receive the treatment benefit of testing with certainty and his expected utility conditional on testing HIV positive when treatment is available is given by

$$V(p_1(1)) - \theta - c + \alpha. \tag{7}$$

The expected utility of testing when there is ART is the linear combination of these two conditional expectations where the weights are the prior probabilities of being HIV negative and of being HIV positive, respectively. Assuming $V(p_1(0)) - p_1(0)\theta > V(p_1(1)) - \theta + \alpha$ (i.e., the individual would prefer to engage in $p_1(0)$ and acquire HIV with probability $p_1(0)$ than to engage in $p_1(1)$ and acquire HIV (and begin ART) with certainty), implies that $\frac{\partial E[U(test)] \text{ with ART}}{\partial p_0} < 0$.

The model predicts that when ART is available (and there is no rationing) demand for testing should be lowest among individuals with the lowest prior probabilities of being HIV positive. For an individual with a low value of p_0 , the expected treatment benefit of testing is relatively low and the expected change in demand for risky behavior associated with taking a HIV test is also relatively low. For an individual with a high value of p_0 , the expected treatment benefit of testing is relatively high but the expected change in demand for risky behavior associated with taking a HIV test is relatively low. Whether individuals with high values of p_0 or individuals with values of p_0 in the middle of the p_0 distribution have the greatest demand for HIV testing when ART is available depends on the magnitude of the treatment benefit of testing relative to the benefit associated with the expected change in demand for risky behavior if the individual were to take a HIV test.

3.4 Change in risky sexual behavior

Linking the upper and lower halves of Figure 6 yields several important insights about the effect of HIV testing on risky sexual behavior. First, consider the state of the world in which ART is not available. The individual chooses to take a HIV test if and only if $E[U(test)] \geq E[U(not\ test)]$. That is, the individual will test if and only $p_0 \in A$. The associated expected change in risky behavior for this individual is $(1 - p_0)p_1(0) + p_0p_1(1) - p_1(p_0)$.

Now suppose that ART is available. As in the absence of ART, if $p_0 \in A$ then the individual will choose to take a HIV test. Moreover, if $p_0 \in A_1$ or $p_0 \in A_2$ then the individual will also test. If individuals are uniformly distributed on the unit interval, the number of individuals in the interval A_1 is greater than those in the interval A_2 and hence the number of individuals induced to test with $p_0 > \max\{p_0 \mid p_0 \in A\}$ is greater than those induced to test with $p_0 < \min\{p_0 \mid p_0 \in A\}$. Because the introduction of ART increases the mean p_0 among those individuals taking a HIV test as compared to the “no ART” case, the per-person reduction in risky behavior associated with HIV testing when ART is available is greater than that in the absence of ART.

3.5 Effect of ART rationing

We may use Figure 6 to examine the effects of ART rationing on demand for HIV testing and for risky behavior. In particular, we may examine the effect a rationing mechanism that favors individuals who happen to have relatively low prior probabilities of being HIV positive (e.g., males age 40-59). As compared to the no-rationing case, such a rationing mechanism would serve to reduce the expected treatment benefit of testing among individuals with relatively high values of p_0 while affecting the expected treatment benefit of testing among low- p_0 individuals less. Because it reduces the expected treatment benefit of testing more among individuals expected to demonstrate larger reductions in post-testing risky behavior, a rationing mechanism of this sort reduces the prevention impact of ART-induced testing. More generally, any other factors that increase (decrease) the cost (benefit) of treatment for high- p_0 individuals relative to low- p_0 individuals reduce the prevention impact of ART-induced testing.

4 Data and estimation

4.1 Data

I use data on the outcomes of individuals’ testing decisions (i.e., whether or not they chose to take a HIV test) and data on HIV prevalence by demographic group to estimate the structural parameters of the model. I do not observe choices about post-testing risky behavior. In addition, I do not observe the results of the voluntary tests reported in the Zambia Sexual Behavior Surveys (ZSBS).

The 2001 Zambia Demographic Health Survey (ZDHS) includes the results of anonymous

HIV testing linked to information on the age, gender, and province of residence of each individual tested. The sample of individuals for whom a HIV test was requested was defined to be all men and all women in one-half of households selected for the men’s survey module in the ZDHS. In total, approximately 5,000 individuals were asked to provide a blood sample. Among women approached to provide a sample, 15.4 percent refused and 5.3 percent were either absent or the result is missing. Among men approached to provide to a sample, 14.8 percent refused and 12.1 percent were either absent or the result is missing. The resulting sample consists of testing results for 3,949 observations.

Because the sample size for any given age-gender-province interaction is small, individuals are grouped into five-year age ranges (e.g., 15-19 years old, 20-24 years old, etc.). The ZDHS includes women age 15-49 and men age 15-59 and there are nine provinces in Zambia, so this procedure yields 144 groups with HIV prevalence ranging from zero percent to 35.3 percent. For each group I treat the proportion that were found to be HIV positive as the initial probability of being HIV positive, p_0 .⁹ I assume that the age-profile of HIV prevalence is static for a given gender-province interaction over the period covered by the repeated cross-sectional testing data in the ZSBS.

The data on whether an individual voluntarily chose to take a HIV test come from the ZSBS. I use the three most recent rounds of the ZSBS: 2000, 2003, and 2005. The sample sizes for each of these survey rounds are 3316, 4471, and 4208 individuals, respectively. To match the testing data to the prevalence data the testing decisions of individuals in the ZSBS are aggregated by demographic group defined as the interaction of five-year age group, gender, and province of residence. For each demographic group I aggregate the two pre-ART survey years (i.e., 2000 and 2003). I then calculate the proportion of individuals who report taking a HIV test in the 12 months prior to the survey date, denoted $test_i$. This procedure yields 288 observations.

4.2 Estimation

For a group of individuals with the homogeneous belief they are HIV positive equal to p_{0i} , the proportion predicted to choose to take a HIV test in the pre-ART period is given by

⁹It is reasonable to believe that on average individuals hold beliefs that are consistent with actual infection probabilities. Nonetheless, it is reassuring that Delavande and Kohler (2007) report that finely measured beliefs about own HIV status in the 2006 Malawi Diffusion and Ideational Change (MDICP) data are very similar to HIV prevalence rates in the sample population.

$$\widehat{test}_i = \frac{1}{1 + e^{-(E[U(test)]_i - E[U(not\ test)]_i)/\delta}} \quad (8)$$

where δ is a smoothing parameter that is to be estimated along with the other parameters of the model. The empirical specification for $E[U(test)]$ includes as additive terms an indicator variable equal to one if the observation is a male demographic group, denoted *male*, and an indicator variable, denoted *post*, equal to one if the observation comes from the 2005 survey round of the ZSBS.

In the ART era, the predicted proportion testing in group i is generated by Equation (8) as well, but with an expanded expression for $E[U(test)]$ with *ART* substituted for $E[U(test)]$. As discussed previously, the changes in the testing-age profiles discussed in Section 3 differ from those predicted by the theoretical model. The model predicts that the individuals who demonstrate the greatest testing response to the introduction of ART should have prior probabilities of being HIV positive that are at least as large as those among individuals who were more likely to test prior to the introduction of ART. In contrast, the observed testing response among males do not fit this prediction. One explanation for this puzzle is that there exists a non-random rationing mechanism for allocating ART. More generally, we may expect that there exists unobserved heterogeneity in the benefits and costs of ART that are correlated with demographic characteristics.

To capture this possibility, I estimate the probability that the individual in group i will seek and receive ART conditional on p_{0i} , denoted p'_{0i} , by including measures of each individual's age, age squared, gender, and interaction terms in the empirical specification for the expected utility of testing when ART is available. Specifically, I include the right-hand side of following expression as an additive term in the empirical specification for $E[U(test)]$ with *ART* (i.e., Equation 6)

$$p'_{0i} = \frac{1}{1 + e^{X_i\beta}} \quad (9)$$

where X_i is a vector whose elements include age, age-squared, an indicator variable for male, and the interactions of age (and age-squared) with the male indicator variable. In addition, I examine the relative importance of these demographic characteristics by estimating weights for the effects of the prior probability of being HIV positive (i.e., p_0) and of the demographic-based probability of seeking and receiving ART conditional on p_0 (i.e., p'_0). These weights are denoted $(1 - \sigma)$ and σ , respectively. Recalling that α denotes the net utility of beginning ART conditional on being HIV positive, the expanded expression for $E[U(test)]$ with *ART* that I use in the empirical specification is

$$\begin{aligned}
E[U(test)] \text{ with ART} &= (1 - p_0)[V(p_1(0)) - p_1(0)\theta - c] + p_0[V(p_1(1)) - \theta - c] \\
&+ (1 - \sigma)\alpha p_0 + \sigma\alpha p_0 p_0'.
\end{aligned} \tag{10}$$

Because the ultimate product of the empirical analysis is a set of policy simulations, I need to parameterize the model. In particular, I need to specify a functional form for the utility derived from choice of risky behavior, $V(p_1)$. I choose a functional form for $V(p_1)$ that is sufficiently flexible to allow a variety of shapes for the risky behavior demand function, $p_1(p_0)$. Specifically, the utility derived from choice of risky behavior is given by

$$V(p_1) = \frac{(p_1 - \frac{b}{e})^2}{ep_1 - 1}. \tag{11}$$

The flexibility of this specification means that I am able to let the data confirm that the risky behavior demand function is in fact increasing and concave in p_0 , as was assumed in the discussion of the model in Section 4.

To estimate the structural parameters of the model, I search for the parameters that minimize a simple weighted sum of squared residuals

$$\sum_{i=1}^{288} \frac{pop_i}{pop} (test_i - \widehat{test}_i)^2 \tag{12}$$

where the weights are given by the proportion of the total adult population (pop) that is in p_0 group i (pop_i).

5 Results

5.1 Parameter estimates

Table 1 provides the parameter estimates from the weighted non-linear least squares minimization problem. Column (1) lists the estimates for the baseline model with no additional controls. The baseline model includes the variables indicated by the theoretical model. Substituting these parameters into Equation (2) and examining the first and second derivatives verifies that the expected utility of not testing is decreasing and convex in the prior probability of being HIV positive. In addition, the utility cost of testing is positive (i.e., $c > 0$) and utility is decreasing in the probability that the individual eventually acquires HIV (i.e., $\theta > 0$).

[TABLE 1 APPROXIMATELY HERE]

The parameter estimates indicate that individual behavior is broadly consistent with a model in which not all HIV positive individuals seek and receive ART, but instead there exist systematic heterogeneity across demographic groups in the probability of seeking and receiving ART conditional on being HIV positive. In the baseline specification, the estimated σ equals 0.955. In other words, the probability that an individual responds to the availability of ART by testing is increasing in the probability that he is HIV positive but is also correlated with the individual’s demographic characteristics conditional on p_0 . Of course the fact that $\sigma < 1$ indicates that individual behavior is consistent with a model where the probability that an individual is HIV positive also affects the testing decision irrespective of these demographic characteristics. Nonetheless, this effect appears to be small and the estimated σ is consistent with the existence of a non-random ART rationing mechanism, or some other source of systematic heterogeneity across demographic groups in the probability of seeking and receiving ART conditional on being HIV positive.

If province-level unobserved heterogeneity in testing demand is correlated with variables in the model, then the estimates presented in Column (1) are biased. For example, we know that there was an expansion in Zambia in the number of VCT sites between the pre-ART and ART periods, effectively lowering the travel cost of testing. Thus, the estimate of the utility benefit of beginning ART shown in Column (1) may be biased upwards. To address this concern, Column (2) presents parameter estimates from Model (2) which includes a “post” indicator variable, equal to one if the observation comes from the 2005 survey round of the ZSBS. Model (2) also includes a “male” fixed effect, fixed effects at the five-year age group level, and province fixed effects.

One reason the gender, age, and province fixed effects are important is that the prior probability of being HIV positive used in the estimation procedure is essentially given by the individual-level reduced form regression of HIV status on the interaction of gender, five-year age group, and province of residence. Social stigma, control over sexual decision-making, biological differences in HIV transmission/progression, and concerns about mother-to-child transmission are all reasons to believe that the net benefit of testing varies by gender. Likewise, it is highly probable that the net benefit of testing/treatment varies by age. For example, as one approaches one’s expected date of death, the personal health cost of acquiring HIV should be decreasing; in the limiting case, the health cost of acquiring HIV on the last day of life should be zero, ignoring concerns about transmitting the virus to others. Finally, the possibility of inter-regional heterogeneity in average distances to the

nearest VCT and ART sites means that the net benefit of testing may vary by province of residence.

As shown in Column (2), the basic results of the model do not change after it is re-estimated with time, gender, age group, and province fixed effects. The expected utility of not testing is still decreasing and convex in the prior probability of being HIV positive, θ and c still have the expected signs, and individual behavior remains consistent with a model in which there exists substantial rationing of treatment based on age. It is reasonable to consider whether these results are robust to the inclusion of additional fixed effects (e.g., interactions of the fixed effects included in Model (2)). However, the computational demands of structural estimation for this non-linear model mean that it is difficult to reliably estimate these additional parameters. In any case, that the parameter estimates changed very little with the controls included in Model (2) suggests that interacting these controls would not have a substantial effect on the parameter estimates.

5.2 Predicted testing-age profiles

Comparing the predicted and observed testing rates is a simple method of assessing the fit of the empirical specification discussed above. Figure 7 presents predicted and observed testing-age profiles for females and for males in the pre-ART period and in the ART era. The predicted testing-age profile in a given period in Figure 7 is the result of a locally weighted smoothed regression of the predicted testing rate for group i on age. The observed testing-age profile in a given period in Figure 7 is the result of a locally weighted smoothed regression of the actual testing rate for group i on age. For a given gender there are two panels in Figure 7, one panel for each of the two empirical specifications discussed above.

[FIGURE 7 APPROXIMATELY HERE]

Each of the empirical specifications appear to generate predicted values that are relatively close to the observed HIV testing rates. The predicted testing rates tend to be within one percentage point or less of the observed testing rate. The mean residuals are largest for the youngest females in the ART era. In general, the empirical specifications generate predicted testing-age profiles in the pre-ART period that take the inverted U-shape that characterizes the observed testing-age profiles and predicted testing-age profiles in the ART era that take the distinct shapes that characterize the observed testing-age profiles for females and males.

5.3 Predicted risky behavior demand function

Figure 8 presents the predicted risky behavior demand function, $\widehat{p}_1(p_0)$, from the empirical specification that includes the full set of controls. Although the empirical specification does not impose that $\widehat{p}_1(p_0)$ be increasing or concave in p_0 , the estimated risky behavior demand function is in fact increasing and concave. The estimated risky behavior demand function implies that if an individual knew with certainty that they were HIV negative then they would choose an amount of risky sexual behavior in which to engage that would imply that they would acquire HIV with probability 0.128. On the other hand, if an individual believed with certainty that they were HIV positive then they would choose an amount of risky sexual behavior in which to engage that would imply that they would acquire HIV with probability 0.591 if they were in fact not HIV positive. The prior probability of being HIV positive that maximizes the expected reduction in risky behavior associated with taking a HIV test is 0.428. However, even at its maximum the expected reduction in risky behavior associated with taking a HIV test in terms of the change in the probability the reference individual acquires HIV - 0.060 - is relatively small.

[FIGURE 8 APPROXIMATELY HERE]

6 Policy scenarios

This section uses the parameter estimates from the preceding section to examine the behavioral and welfare implications of five alternative ART policies: (i) no ART, (ii) the status quo ART policy, (iii) expanding the supply of ART, (iv) eliminating heterogeneity in the probability of seeking and receiving ART conditional on being HIV positive that is correlated with demographic characteristics (i.e., eliminating the prevention inefficiency), and (v) simultaneously eliminating this prevention inefficiency while expanding the supply of ART. I employ three metrics to examine the performance of each of these policies. The first metric is the predicted demand for HIV testing (i.e., the HIV testing rate). The second metric is the mean predicted choice of risky behavior. The third metric is the simulated number of new infections. The first two metrics are relatively straightforward to calculate. The simulated number of new infections, however, requires particular discussion.

I simulate the effect of testing on new infections as follows. Each individual in a given demographic group may be characterized along two dimensions: whether they take a HIV test or not and whether they are HIV positive or not. In a given demographic group, whether an individual chooses to take a HIV test or not (as well as the result of the test) affects the

individual’s choice of risky behavior. Only individuals who are HIV negative are at risk of acquiring a new infection.

The number of new infections among individuals taking a HIV test is given by

$$\sum_{i=1}^N n_i \widehat{test}_i (1 - p_{0i}) \widehat{p}_1(0) \quad (13)$$

where n_i is the number of individuals in group i and N equals 144, the number of demographic groups. For demographic group i , the product of the first three terms in this expression yields the number of individuals in this group that are at-risk of acquiring HIV. Because these individuals are HIV negative and the results of the HIV tests reveal this fact to each of the individuals, they choose to engage in an amount of risky behavior summarized by $\widehat{p}_1(0)$; this is the probability that they acquire HIV.

The number of new infections among individuals not taking a HIV test is given by

$$\sum_{i=1}^N n_i (1 - \widehat{test}_i) (1 - p_{0i}) \widehat{p}_1(p_{0i}). \quad (14)$$

Again, the product of the first three terms in this expression yields the number of individuals in demographic group i that are at-risk of acquiring HIV. Because these individuals choose not to take a HIV test, they engage in an amount of risky behavior summarized by $\widehat{p}_1(p_{0i})$; this is the probability that they acquire HIV. Total new infections are given by the sum of these two expressions.

This method of simulating the number of new infections has one major shortcoming. Namely, it ignores new infections among partners of the reference individuals. Thus, this method measures the expected number of new infections in a given population with some error.

6.1 No ART

The first policy scenario I examine is a scenario where there is no ART. Applying the three metrics to this case provides a baseline for understanding the impact of each of four other ART policies. I calculate the predicted HIV testing rate in the “no ART” scenario by shutting down the treatment benefit to testing (i.e., setting α equal to zero), setting “post” equal to one, and generating predicted values using Equation (8). The first row of Table 2 presents the results of this calculation for each of the two empirical specifications. Model (1) is the baseline empirical specification with no additional controls. Using the parameter estimates from this specification, the predicted HIV testing rate in the state of the world where there

is no ART is 5.7 percent. Model (2) includes “post”, “male”, age group, and province fixed effects, as well as the controls included in Model (1). As Column (2) of Table 2 indicates, the predicted testing rate remains virtually unchanged at 5.8 percent.

[TABLE 2 APPROXIMATELY HERE]

The second metric is the mean choice of risky behavior. The first row of Table 3 presents the mean choice of risky behavior under the “no ART” scenario for each of the two empirical specifications. The parameter estimates from Model (1) imply that the mean choice of risky behavior would be 0.217. This implies that the representative individual would choose to engage in an amount of risky behavior that would lead them to acquire HIV with probability 0.217 if they were in fact HIV negative at the time of this choice. As in the case of the average testing rate, the inclusion of gender, age, and province fixed effects does little to change predicted behavior: Model (2) implies only a slightly higher mean choice of risky behavior, 0.223.

[TABLE 3 APPROXIMATELY HERE]

The third metric is the simulated number of new infections. Table 4 presents the results of this simulation; the first row of this table represents the “no ART” scenario. According to the parameter estimates from Model (1), the number of new infections in the state of the world where there is no ART appears to be relatively large: 907,000 new infections. Of course the theoretical model allows for one chance to take a HIV test and one chance to make a decision about future risky behavior. Thus, it may make sense to think about the simulated number of new infections as representing the number of new infections in the population over the course of the average lifetime of an individual in the population. Given that on average an individual between the age of 15 and 59 in Zambia is expected to live approximately 25 more years (WHO 2007), then the simulated annual incidence is nearly 40,000. As a point of comparison, the annual incidence of new adult HIV cases in Zambia is approximately 50,000. Consistent with the higher mean choice of risky behavior in Model (2) as compared to Model (1), the simulated number of new infections in Model (2), 934,000 infections, is greater than that in Model (1).

[TABLE 4 APPROXIMATELY HERE]

6.2 Status quo ART policy

The second policy scenario I consider is the status quo ART policy. To calculate the three metrics under this policy scenario, I set “post” equal to one and make no additional changes. The second row in Table 2 presents the predicted proportion of individuals choosing to take a HIV test under this scenario. Depending on the empirical specification, between 8.0 and 8.5 percent of individuals are predicted to take a HIV test. This represents an increase in testing demand of between 2.3 and 2.7 percentage points (an increase of between 40 and 47 percent).

Although the effect is not nearly as large, ART provision under the status quo ART policy is also predicted to reduce mean choice of risky behavior. As presented in the second row of Table 3, Models (1) and (2) imply that the mean choice of risky behavior would fall by approximately 0.45 percent (or approximately one-thousandth of a percentage point).

The second row of Table 4 presents the simulated number of new infections under the status quo ART policy. Depending on the empirical specification, the number of new infections ranges from 894,000 to 918,000. These figures represent a decline in the incidence of HIV of between 1.43 and 1.71 percent. In general the effect of ART-induced testing under the status quo ART policy appears to be modest at best. However, comparing these figures with the performance of the alternative policies should yield a more complete understanding of the relative magnitude of this effect.

6.3 Expanding the supply of ART

The third policy scenario I examine is expanding the supply of ART while retaining the existing source of prevention inefficiency in the process determining the allocation of ART. If the quantity of antiretroviral drugs available in Zambia is not sufficient to meet the demand for ART, then treatment providers may ration ART among the pool of eligible individuals without regards to each individual’s identity. I consider the policy of expanding supply to ensure that all individuals who would be allocated ART under the existing non-random rationing mechanism are able to begin treatment. Expanding the supply of ART should yield an increase in the prevention impact of ART-induced testing as compared to the status quo ART policy. Because expanding the supply of ART increases the probability that an individual will receive the payoff associated with taking a HIV test conditional on the prior probability the individual is HIV positive (and on the probability that they would be allocated ART under the non-random rationing mechanism), it should increase the proportion of individuals testing. The cost of this policy is the increase in drug expenditures (and

associated labor costs) required to cover these additional patients.

It appears that the quantity of antiretroviral drugs in Zambia is not sufficient to meet the demand for ART. I calculate the ratio of the number of individuals on ART at the time of the 2005 round of the ZSBS (i.e., 30,000) to the number of individuals who are predicted to test in the 2005 round of the ZSBS and to be eligible for ART. Let q denote this ratio. Then

$$q = \frac{30,000}{\sum_{i=1}^N ((1 - \hat{\sigma})\hat{p}_{0i} + \hat{\sigma}\hat{p}_{0i}')\widehat{test}_i n_i} \quad (15)$$

where N equals 144, the number of demographic groups. Consistent with the existence of excess demand for ART, I find that this ratio is substantially less than one ($q = 0.599$). If there are individuals who did not take a HIV test in the 12 month interval but were still eligible for ART and able to demonstrate their status to the health authorities (e.g., from the results of a test taken prior to the beginning of the 12 month interval), then the q given by Equation (15) is going to understate the extent of rationing of ART using a random lottery. On the other hand, not all HIV positive individuals are recommended to begin ART and so the q given by Equation (15) may overstate the extent of rationing of ART using a random lottery. After discussing the simulation results associated with the assumption that $q=0.599$, I will examine the robustness of the results to changes in this assumption.

To calculate the expected level of testing associated with ART availability under the expanded supply scenario, first I substitute the parameter estimates from Table 1 (except the estimated α) into the predicted testing equation and set “post” equal to one. Then I substitute $\frac{1}{0.599}\hat{\alpha}$ for α in the predicted testing equation (where $\hat{\alpha}$ is the estimated α). Because the effect of rationing without regards to individuals characteristics is not separately identified from the benefit of beginning ART (i.e., α), eliminating the rationing of ART using a random lottery is equivalent to increasing α . The third set of rows of Table 2 present the results of this exercise. Depending on the empirical specification, I predict that expanding the supply of ART while retaining the existing non-random ART rationing mechanism should increase testing rates to between 10.1 and 11.2 percent. This represents an increase in testing rates as compared to the status quo ART policy of between 2.1 and 2.7 percentage points.

The third set of rows in Table 3 presents the mean choice of risky behavior under the “expand supply” scenario. Although the mean choice of risky behavior in this scenario is smaller than that under either the “no ART” or the “status quo ART” scenarios, the reduction in mean risky behavior is still small in an absolute sense. As compared to the “no ART” scenario, mean risky behavior is approximately 0.90 percent smaller under the

“expand supply” scenario.

The simulation results fail to provide support for the idea that the continued roll-out of ART (in the absence of other interventions designed to eliminate the non-random allocation process) will induce substantial reductions in the incidence of new HIV infections. As shown in the third set of rows in Table 4, the number of new infections avoided under this policy as compared to the state of the world without ART is only between 2.76 and 3.43 percent. In relative terms, however, this policy is much more effective than the status quo ART policy: simply expanding the supply of ART to eliminate the random lottery increases the number of new infections avoided by more than 90 percent.

Of course the assumption that $q = 0.599$ may be incorrect. If excess demand is less than that consistent with this assumption (i.e., if the “true” q is greater than 0.599), then the simulated effects of expanding the supply of ART as reported in Tables (2) through (4) are simply overestimates. In the limiting case that there exists no excess demand (conditional on the non-random rationing mechanism), then we cannot consider the effects of expanding the supply of ART without addressing the non-random rationing mechanism (i.e., the simulated effects of this policy are zero).

If excess demand for ART is greater than what I assume, then the simulated effects of expanding the supply of ART as reported in Tables (2) through (4) are underestimates of what we may reasonably believe. To examine the sensitivity of the simulation results to changes in q , the simulations were re-implemented using $q = 0.40$. This choice of q is fairly arbitrary, but the quantitative simulation results change very little and the qualitative simulation results remain the same: simply expanding the supply of ART has only modest prevention effects above and beyond the status quo ART policy.

6.4 Eliminating the prevention inefficiency

The fourth policy scenario is to eliminate the existing source of prevention inefficiency in the ART allocation process while holding constant the quantity of ART supplied. Under this policy, all individuals who are HIV positive are eligible for ART. Age and gender no longer affect the probability of seeking and receiving ART conditional on HIV status. If the non-random rationing story is correct, one way to implement this policy would be to mandate that the enrollment-age profile at ART sites in each province match the province’s prevalence-age profile. Because the existing supply of ART is held constant, eliminating the inefficiency would require no additional expenditure on antiretroviral drugs. The tradeoff or cost of this policy is that individuals who are currently favored in the ART allocation process would face

a lower probability of beginning ART conditional on being HIV positive. Moreover, because the quantity of ART supplied is held constant while the number of individuals eligible for ART increases, the probability that an HIV positive individual who takes a HIV test is allocated ART in the random lottery decreases.

Eliminating the source of inefficiency in the process determining the allocation of ART has the potential to yield a large increase in the number of new infections avoided. The model provided in this paper indicates that the expected reduction in risky behavior associated with taking a HIV test is increasing in the initial probability of being HIV positive over the range relevant in the current setting. However, empirical evidence suggests that current ART policy effectively favors populations (i.e., older males) with prior probabilities of being HIV positive that are much lower than the maximum prior probability observed in the Zambia DHS. Thus, eliminating the source of the inefficiency will increase the number of new infections avoided due to ART-induced testing in part because the composition of the pool of individuals choosing to take a HIV test will shift toward individuals expected to demonstrate relatively large reductions in post-testing risky behavior. In addition, there are many more individuals in groups that are predicted to voluntarily test at high rates in the absence of the source of inefficiency (e.g., males in the middle of the age distribution) than there are in the group (i.e., older males) that currently tests at a high rate. Hence, eliminating the source of inefficiency will increase the number of new infections avoided due to ART-induced testing in part because the number of individuals testing after the inefficiency is addressed would be much greater than the number that test under the status quo ART policy.

I operationalize this policy experiment by generating predicted testing values using the parameter estimates from Table 1 (except the estimated α), setting “post” equal to one, $\sigma = 0$, and recalibrating q . I recalibrate q to reflect the fact that if the policy eliminates the non-random rationing mechanism then the number of individuals who respond to ART by testing will be greater. The new value of q is 0.3. Because the quantity of ART supplied is held constant in this policy, the increase in testing demand means a decrease in the chance of being selected in the random lottery. Therefore, I substitute $0.3\hat{\alpha}$ for α in the predicted testing equation. Again, because the effect of a random lottery determining the allocation of ART among HIV positive individuals is not separately identified from the benefit of beginning ART (i.e., α), the effect of decreasing the probability an individual is selected in the lottery is equivalent to decreasing the benefit of beginning ART.

Regardless of the metric under consideration, the simulated results of this policy provide

clear evidence of the welfare cost of the non-random rationing mechanism revealed in this research. Predicted testing demand is between 16.9 and 18.6 percent, a roughly two-fold increase over testing rates under the status quo ART policy. The percent reduction in mean risky behavior as compared to the “no ART” policy is between 2.69 and 2.76, at least six times greater than the reduction under the status quo ART policy. As compared to the status quo policy, the number of new infections avoided due to ART-induced testing would be more than four times as large if the non-random rationing mechanism were eliminated: between 69,000 and 78,000 new infections avoided.

6.5 Eliminating the prevention inefficiency while expanding the supply of ART

The fifth policy scenario is to expand the supply of ART in combination with eliminating the existing source of inefficiency. In particular, I consider an expansion in the supply of ART that is sufficient to ensure that all individuals who are HIV positive would be able to begin ART. This policy yields the both the mechanical benefit of increasing ART-induced testing by eliminating the random lottery for ART as well as the composition benefit of inducing additional testing among individuals expected to demonstrate greater post-testing reduction in risky behavior. Of course this policy requires the greatest increase in expenditure. The increase in expenditure on antiretroviral drugs (and the associated labor costs) is greater than that under the policy of expanding the supply of ART while retaining the existing source of inefficiency. Moreover, eliminating the source of inefficiency may itself be costly.

To simulate the effect of this new policy, I set “post” equal to one, $\sigma = 0$, replace α with $\frac{1}{0.599}\hat{\alpha}$, and substitute the parameter estimates from Table 1 for the remaining parameters. The simulation results indicate that testing demand under this policy would be between 25.4 and 28.1 percent, a more than four-fold increase as compared to the “no ART” policy. The mean choice of risky behavior would be between 4.04 and 4.61 percent smaller than under the “no ART” policy. In terms of new infections avoided, this policy would have very large effects: the number of new infections would be at least 12 percent (11 percent) smaller than that under the “no ART” scenario (“status quo” scenario).

7 Alternative explanations

The change in testing behavior associated with the introduction of ART in Zambia is consistent with the existence of a non-random rationing mechanism determining the allocation

of ART among HIV positive individuals who test and are clinically eligible for ART. Testing decisions among older men appear to be particularly responsive to the introduction of ART despite the fact that older men are relatively unlikely to be HIV positive. In contrast, testing decisions among men in the middle of the age distribution (e.g., ages 30-35) are much less responsive to the introduction of ART despite the fact that these men are one of the higher HIV prevalence groups. In this section I consider alternative explanations for this puzzle.

1. Concurrent public health campaign specifically targeting older men. If other anti-HIV/AIDS campaigns targeting older men intensified in parallel to the introduction of ART, this might explain the puzzlingly large increase in testing rates among older men. First, this seems unlikely on *a priori* grounds. Older men are one of the lowest HIV prevalence groups in Zambia so they are not at particularly high risk of infecting others. Similarly, their female partners (6 years younger on average) are also relatively unlikely to be HIV positive so older men are not at particularly high risk of acquiring HIV. Finally, all else equal, a prevention-oriented policymaker would target young adults in public efforts at sustained behavior change because they have the greatest expected remaining lifespan and hence the greatest prevention payoff from a given prevention investment.¹⁰

In any case, the ZSBS data include information on exposure to HIV prevention interventions so it is possible to directly test this hypothesis. An analysis of these data reveals that older males do not appear to have been targeted in HIV prevention campaigns implemented simultaneously with the introduction of ART. The fraction of males aged 15-44 who knew of a location where they could take a HIV test increased from 77 percent in the two pre-ART rounds of the ZSBS (i.e., 2000 and 2003) to 87 percent in the 2005 ZSBS. In comparison, this figure for males aged 45-59 increased from 79 percent prior to the introduction of ART to 87 percent in the 2005 ZSBS. Although knowledge of testing locations increased over time, the secular increase does not explain the increase in testing rates among older men.

Likewise, data from the ZSBS on the four large information campaigns included in both a pre-ART round (i.e., 2003) and an ART-era round (i.e., 2005) do not support the hypothesis that anti-HIV/AIDS campaigns targeted older males.¹¹ The proportion of males aged 15-44 (45-59) who had ever seen or listened to the “HEART” program decreased from 29 (32) percent in 2003 to 27 (29) percent in 2005. The proportion of males aged 15-44 (45-59) who had ever seen or listened to the “Your Health Matters” program decreased from 63 (66) percent in 2003 to 57 (53) percent in 2005. The proportion of males aged 15-44 (45-59) who

¹⁰Anecdotal evidence from conversations with policymakers in Zambia is consistent with this conjecture.

¹¹The 2005 ZSBS asked about none of the information campaigns included in the 2000 ZSBS.

had ever seen or listened to the “Kabanana” program decreased from 51 (44) percent in 2003 to 44 (39) percent in 2005. The proportion of males aged 15-44 (45-59) who had ever seen or listened to the “Lifeline” program decreased from 37 (40) percent in 2003 to 30 (25) percent in 2005.

2. VCT and/or ART expansion correlated with HIV prevalence. At the moment, there do not exist spatial data on the change in VCT and ART availability over time. All that exist are the JICA Health Facilities Census data on the spatial distribution at a single point in time. However, discussions with policymakers at the Ministry of Health and with the main ART providers (e.g., the Center for Infectious Disease Research in Zambia (CIDRZ)) indicate that the initial VCT expansion occurred in major urban areas (i.e., Kitwe, Lusaka, Ndola), the initial ART expansion occurred in these same areas, and a follow-up VCT expansion occurred at the same time as the ART expansion but in less populated areas. Although these expansions are likely correlated with HIV prevalence at the province level, the expanded empirical specification (i.e., Model (2)) includes province fixed effects which addresses any potential bias from the possible province-level correlation between VCT (or ART) expansion and HIV prevalence, which is measured at the province level.

Even if ART rationing favoring older men is not the reason older men disproportionately increased their testing rates in parallel to the introduction of ART in Zambia, the conclusion that this increase came at the cost of forgone prevention holds. The model I present is part of a general class of models where the net private benefit of testing conditional on belief about own HIV status varies by demographic group. Within this class of models, the estimated risky behavior demand function does not depend on whether the inference that ART rationing is what is driving this heterogeneous testing behavior is correct. Therefore, the conclusion that there is a foregone prevention impact of having older men test instead of men who are more likely to be HIV positive (e.g., men ages 35-40) would still hold even if ART rationing is not the explanation for the puzzle this paper identifies. Nonetheless, ART rationing appears to be a likely explanation for the changes in the observed testing-age profiles associated with the introduction of ART in Zambia.

8 Conclusion

The debate about appropriate HIV/AIDS policy often features an analytic framework in which interventions are classified as treatment interventions or as prevention interventions (Canning 2006). Antiretroviral therapy (ART) and care for opportunistic infections (e.g.,

tuberculosis) are usually seen as treatment interventions. HIV testing and condoms are usually seen as prevention interventions. However, many proponents of HIV testing argue that antiretrovirals are not simply a treatment intervention and that expanding access to ART in the developing world will induce substantial increases in testing demand and subsequent reductions in risky behavior.

This paper formalizes this intuition using a consumer demand model of sexual behavior decision-making under uncertainty and provides evidence on the magnitude of this effect. In doing so, it generates three broader insights about the economics of HIV/AIDS and health economics more generally. First, the structural model demonstrates that we can recover the parameters of the risky behavior demand function without actually observing risky sexual behavior. Thus, the technique I employ in this paper provides a solution for those concerned about measurement error in self-reported sexual behavior. Second, in contrast to existing analyses of the effect of treatment interventions on preventive behavior, this paper shows that treatment for an infectious disease may actually increase private preventive behavior. The availability of treatment induces individuals to acquire more information about their infection status and this information may lead to increased private prevention efforts. Third, the analysis indicates that standard voluntary counseling and testing (VCT) programs yield greater behavior change per person than if we made everyone test or offered financial incentives for testing. Because existing analyses of the effects of HIV testing on risky behavior largely estimate the average treatment effect under the latter two regimes, these analyses tend to understate the efficacy of the predominant approach to HIV testing.

From a more policy-oriented perspective, the results of the current analysis indicate that the prevention impact of antiretrovirals under the existing ART policy in Zambia is modest at best. The observed changes in testing rates associated with the introduction of ART are not consistent with the prediction of a model of testing demand where the incentive effect of ART is increasing in the prior probability that the individual is HIV positive. Instead, from the perspective of maximizing the prevention impact of ART-induced testing, there appears to exist a source of inefficiency in the process determining who seeks and receives ART. Because of this inefficiency, the number of new infections avoided due to ART-induced testing is relatively small; those individuals disproportionately likely to seek and receive ART are predicted to demonstrate only small reductions in risky behavior subsequent to testing. Simulation results indicate that ART-induced testing under the existing ART policy reduces the number of new infections by less than 2 percent.

Eliminating the inefficiency in the process determining the allocation of ART yields sub-

stantial increases in the prevention impact of ART. Simulation results indicate that this intervention would increase the number of new infections avoided due to ART-induced testing more than four-fold. Moreover, this intervention requires no additional expenditure on antiretroviral drugs. In comparison, expanding the supply of ART without eliminating this inefficiency would only double the number of new infections avoided due to ART-induced testing while requiring a substantial increase in expenditure on antiretroviral drugs. The results of this paper confirm the emphasis among policymakers on the central role of ART-induced HIV testing in reducing the spread of HIV/AIDS and suggest that future research examine the process determining the allocation of ART in regions experiencing a rapid change in the availability of ART.

9 References

- Attawell, K., Mundy, J. 2003. Provision of Antiretroviral therapy in resource-limited settings: A review of experience up to August 2003. Working paper.
- Bond, G. 1982. Education and social stratification in Northern Zambia: The case of the Uyombe. *Anthropology and Education Quarterly* 13(3), 251-267.
- Boozer, M., Philipson, T. 2000. The impact of public testing for Human Immunodeficiency Virus. *The Journal of Human Resources*, 35(3), 419-446.
- Bwalya, E. 2006. The political economy of antiretroviral drugs in Zambia. *Perspectives on Global Development and Technology*, 5(4), 385-409.
- Canning, D. 2006. The economics of HIV/AIDS in low-income countries: The case for prevention. *Journal of Economic Perspectives*, 20(3), 121-142.
- Centers for Disease Control and Prevention. 2006. Revised Recommendations for HIV Testing of Adults, Adolescents, and Pregnant Women in Health-Care Settings. *MMWR* 55(RR-14), September 22.
- Colson, E., Scudder, T. 1981. Old Age in Gwembe District, Zambia. In: Amoss, P., Harrell, S. (Eds.) *Other Ways of Growing Old*. Stanford University Press, Stanford, CA.
- Delavande, A. and H.-P. Kohler. 2007. Subjective expectations in the context of HIV/AIDS in Malawi. University of Pennsylvania Working Paper.
- Donnelly, John. 2005. Dire situation, drastic measures: AIDS testing urged for all in ravaged nation. *Boston Globe*, October 23.
- Epstein, A. L. 1981. *Urbanization and Kinship*. Academic Press, New York, NY.
- Gersovitz, M., Jacoby, H.G., Seri Dedy, F., Goze Tape, A.. 1998. The balance of self-reported heterosexual activity in KAP surveys and the AIDS epidemic in Africa. *Journal of the American Statistical Association*, 93(443), 875-883.
- Global HIV Prevention Working Group. 2004. *HIV Prevention in the Era of Expanded Treatment Access*.
- Granich, R., Gilks, C., Dye, C., De Cock, K., Williams, B. 2009. Universal voluntary HIV testing with immediate antiretroviral therapy as a strategy for elimination of HIV

transmission: A mathematical model. *Lancet*, 373(9657): 48-57.

Jones, P. 2004. Who is getting or will get anti-retroviral treatment in Zambia? Equity, access, and fairness in the governance of “scaling up” HIV/AIDS medication. Working paper.

Lakdawalla, D., Sood, N., Goldman, D. 2006. HIV breakthroughs and risky sexual behavior. *The Quarterly Journal of Economics*, 121(3), 1063-1102.

Magruder, J. 2010. Marital shopping and epidemic AIDS. University of California, Berkeley Working paper.

McGough, L., Reynolds, S., Quinn, T., Zenilman, J. 2005. Which patients first? Setting priorities for antiretroviral therapy where resources are limited. *American Journal of Public Health*, 95(7), 1173-1180.

Philipson, T., Posner, R.. 1995. A theoretical and empirical investigation on the effects of public health subsidies for STD testing. *The Quarterly Journal of Economics*, 110(2), 445-474.

Rosen, S., Sane, I., Collier, A., Simon, J. 2005. Rationing antiretroviral therapy for HIV/AIDS in Africa: Choices and consequences. *PLODS Medicine*, 2(11), 1098-1104.

Scudder, T. 1962. *The Ecology of the Gwembe Tonga*. Manchester University Press, Manchester, England.

Stringer, J., Zulu, I., Levy, J., Stringer, E., Mwangi, A., Chi, B., Mtonga, V., Reid, S., Cantrell, R., Bulterys, M., Saag, M., Marlink, R., Mwinga, A., Ellerbrock, T., Sinkala, M. 2006. Rapid scale-up of antiretroviral therapy at primary care sites in Zambia: Feasibility and early outcomes. *Journal of the American Medical Association*, 296(7), 782-793.

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

UNAIDS. 2008. Report on the Global AIDS Epidemic 2008.

UNAIDS/WHO. 2004. UNAIDS/WHO Policy Statement on HIV Testing.

UNAIDS/WHO. 2006. Epidemiological Fact Sheets on HIV/AIDS and Sexually Transmitted Infections.

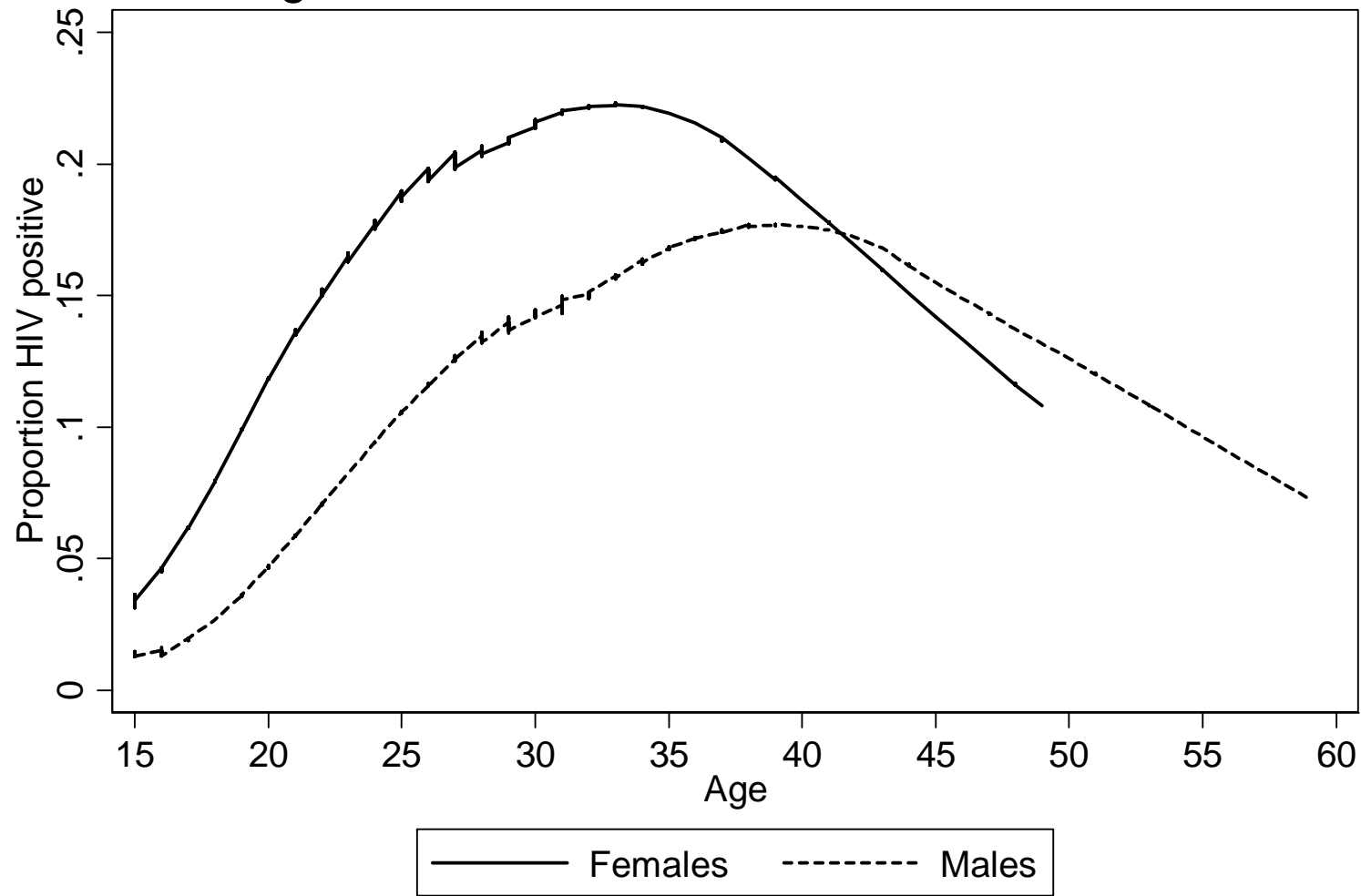
Wawer, M., Gray, R., Sewankambo, N., Serwadda, D., Li, X., Laeyendecker, O., Kiwanuka, N., Kigozi, G., Kiddugavu, M., Lutalo, T., Nalugoda, F., Wabwire-Mangen, F., Meehan, M., Quinn, T. 2005. Rates of HIV-1 transmission per coital act, by stage of HIV-1 infection, in Rakai, Uganda. *The Journal of Infectious Diseases*, 191, 1403-9.

World Health Organization (WHO). 2006a. WHO News Release. August 2006.

World Health Organization (WHO). 2006b. Progress on Global Access to HIV Antiretroviral Therapy: A Report on “3 by 5” and Beyond.

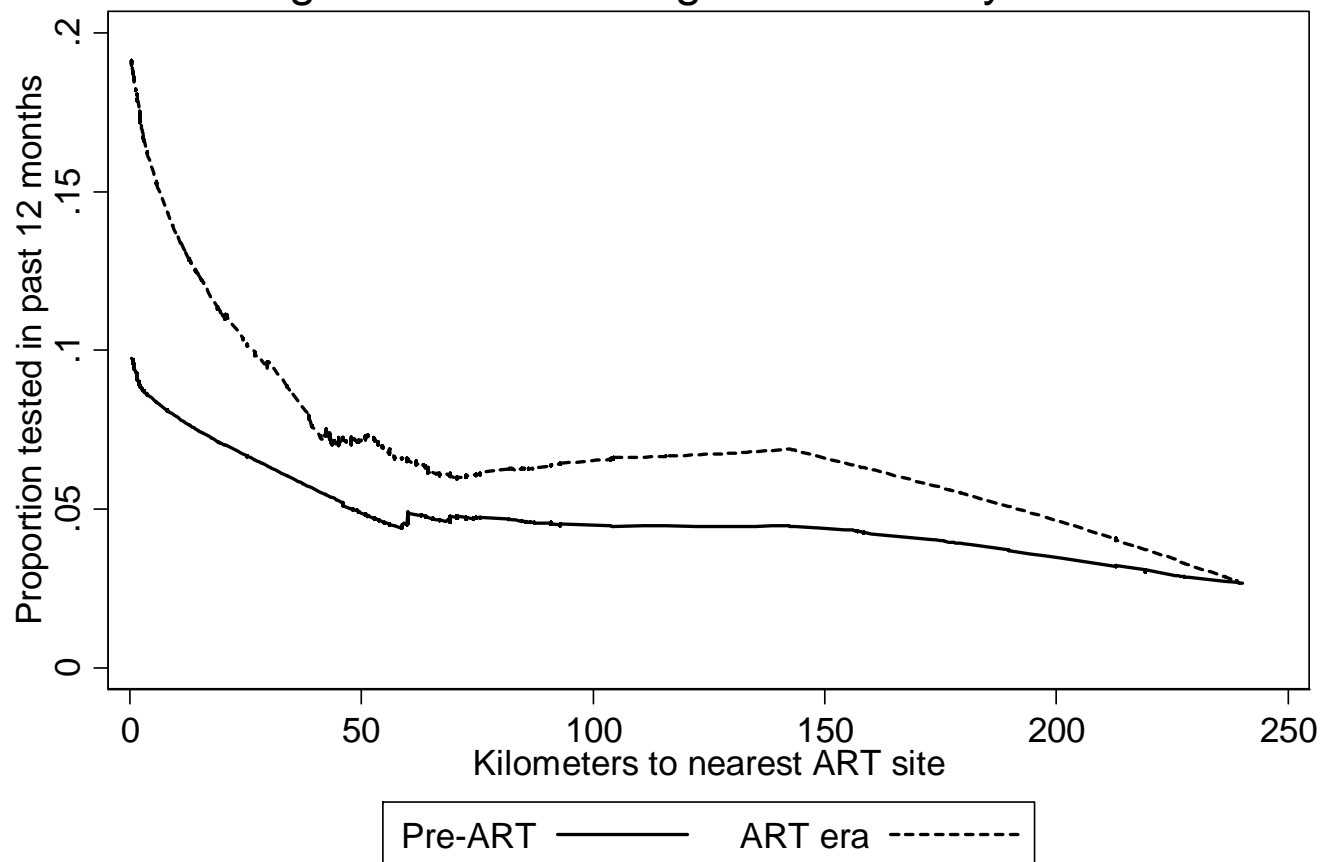
World Health Organization (WHO). 2006c. Antiretroviral Therapy for HIV Infection in Adults and Adolescents in Resource-Limited Settings: Towards Universal Access.

Figure 1: HIV Prevalence, 2001 Zambia DHS



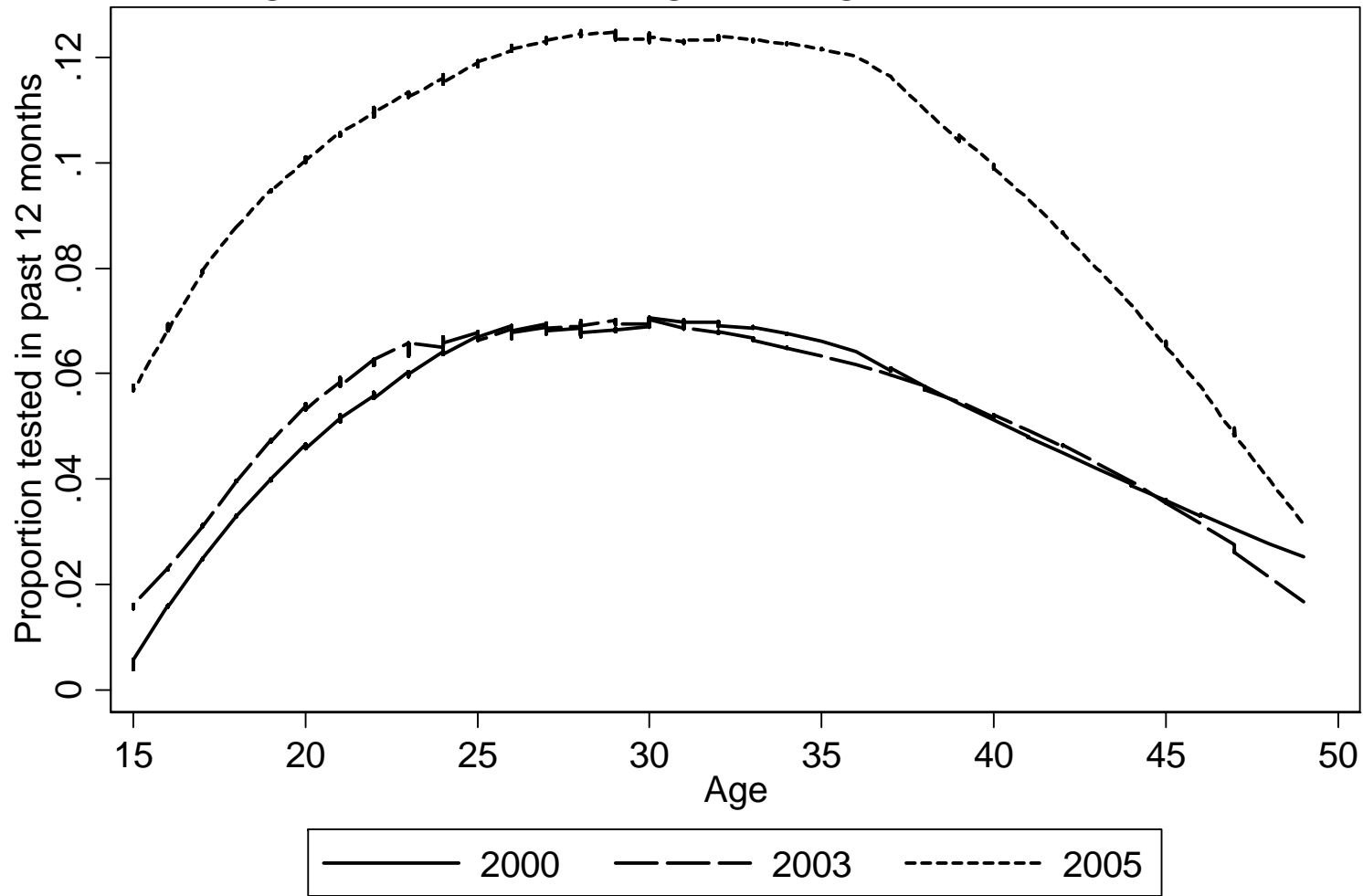
Notes: Locally weighted smoothed regression of an indicator variable for HIV positive on age, estimated separately for females and males. Data come from anonymous HIV testing component of 2001 Zambia Demographic Health Survey (DHS).

Figure 2: HIV Testing and Proximity to ART



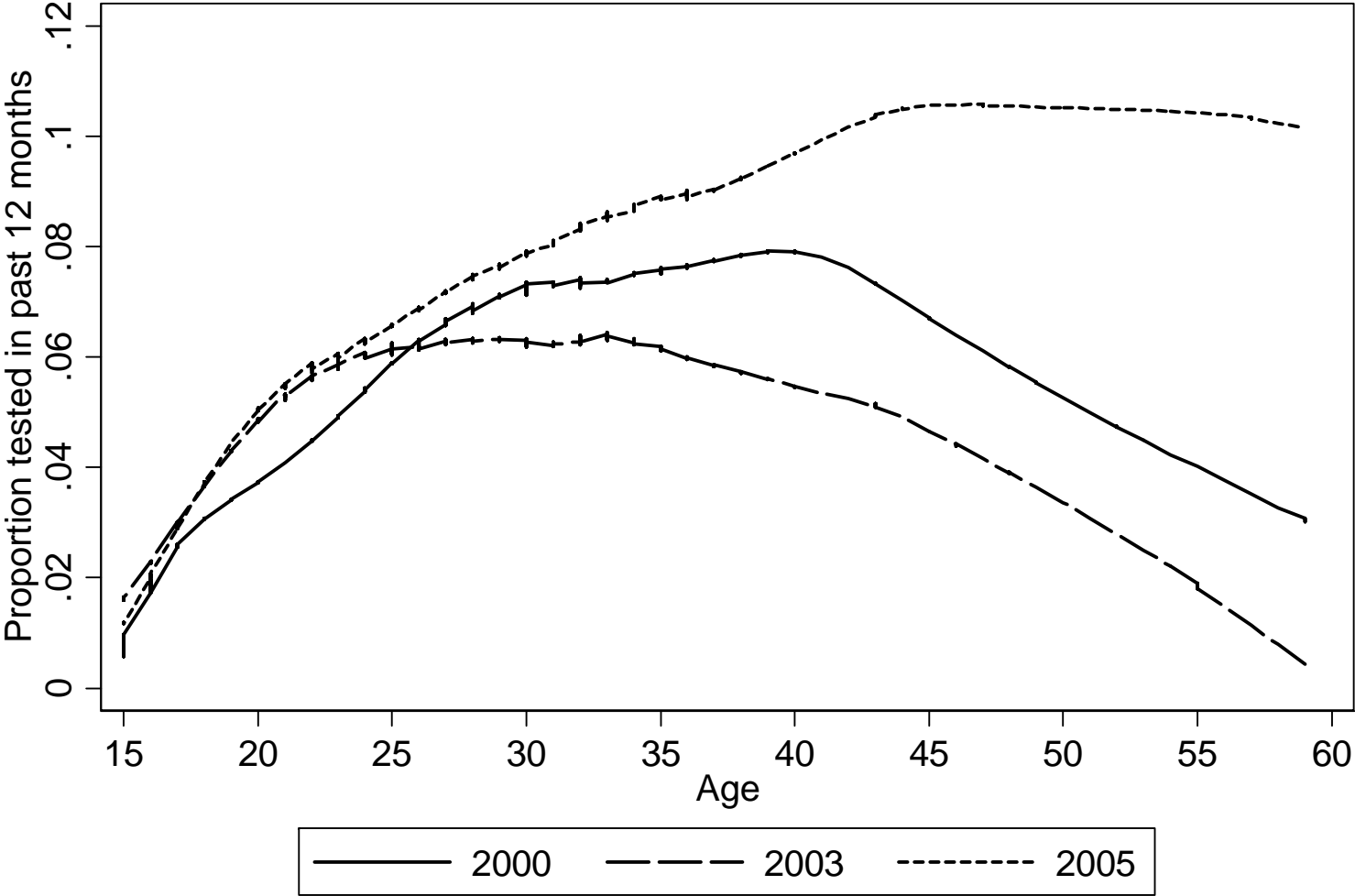
Notes: Locally weighted smoothed regression of an indicator variable for voluntarily took a HIV test in past 12 months on distance to the nearest ART site, estimated separately by pre-ART and ART eras. Testing data for pre-ART era come from 2000 and 2003 Zambia Sexual Behavior Surveys (ZSBS). Testing data for ART era come from 2005 ZSBS. Location of ZSBS respondents calculated as centroid of Standard Enumeration Area (SEA) of residence. Location of nearest ART site comes from the Japanese International Cooperation Agency (JICA) 2004 Health Facilities Census (HFC).

Figure 3: HIV Testing and Age, Females ZSBS



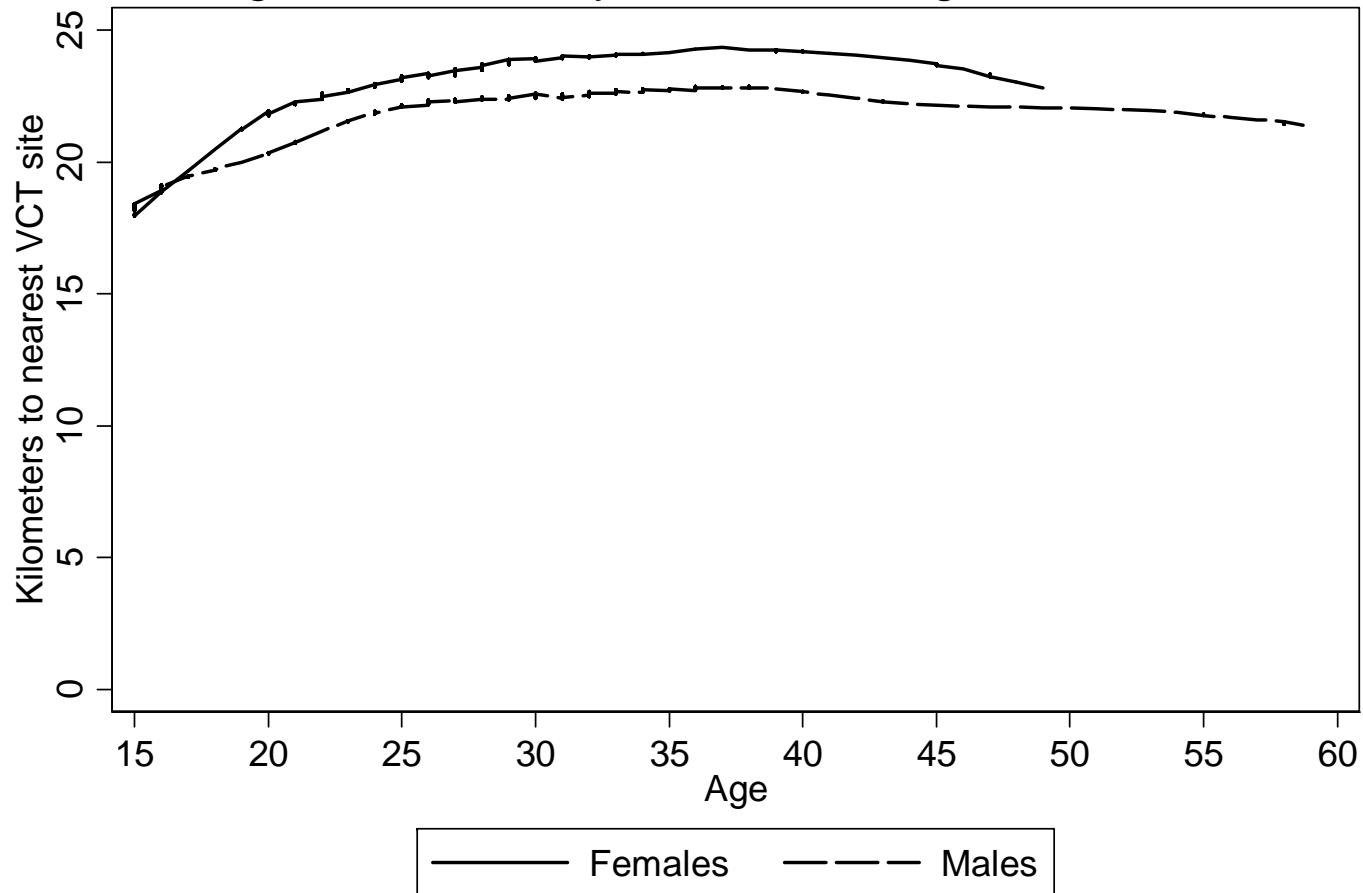
Notes: Locally weighted smoothed regression of an indicator variable for voluntarily took a HIV test in past 12 months on age, estimated separately by survey round of the Zambia Sexual Behavior (ZSBS).

Figure 4: HIV Testing and Age, Males ZSBS



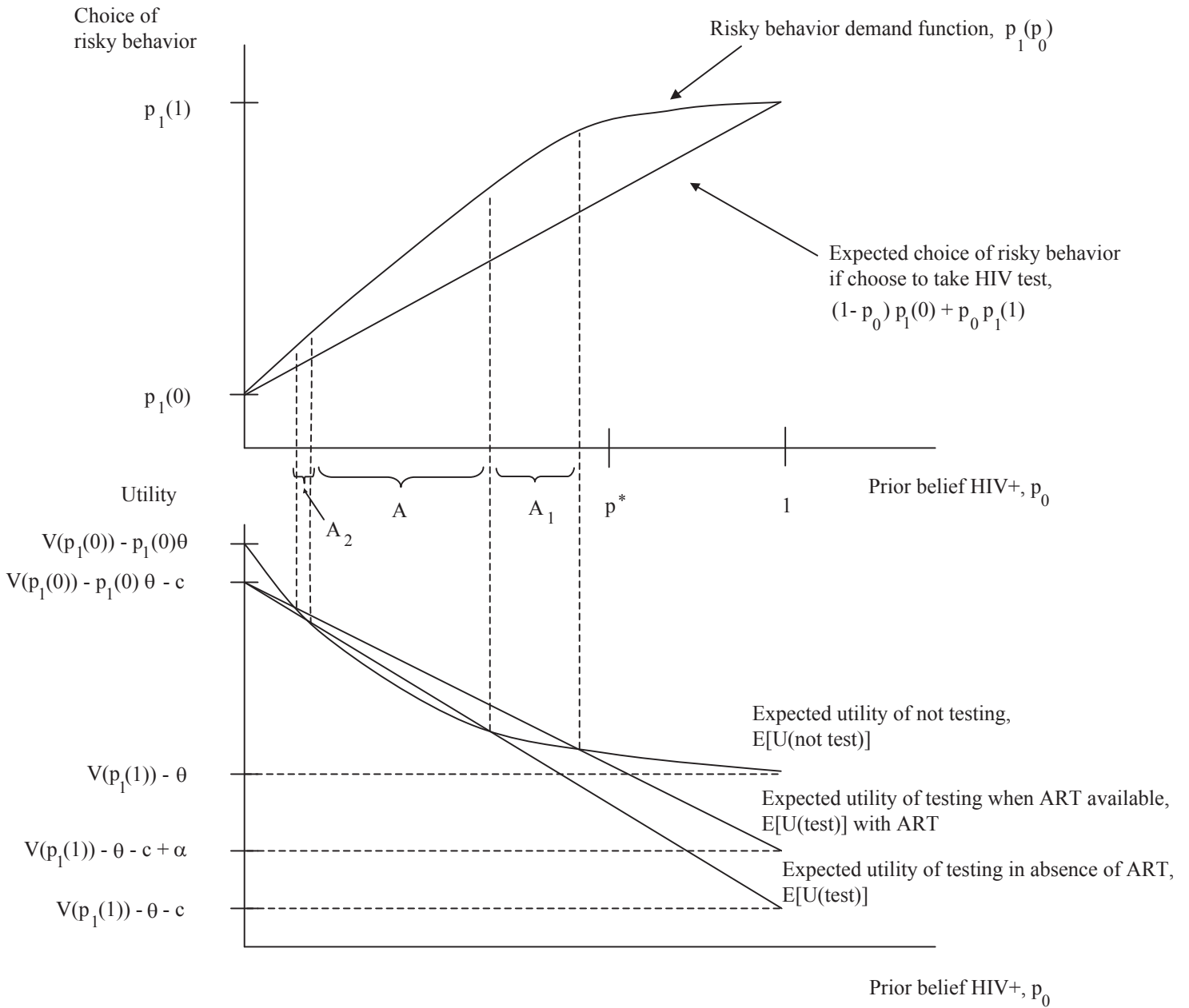
Notes: Locally weighted smoothed regression of an indicator variable for voluntarily took a HIV test in past 12 months on age, estimated separately by survey round of the Zambia Sexual Behavior (ZSBS).

Figure 5: Proximity to VCT and Age, 2005 ZSBS



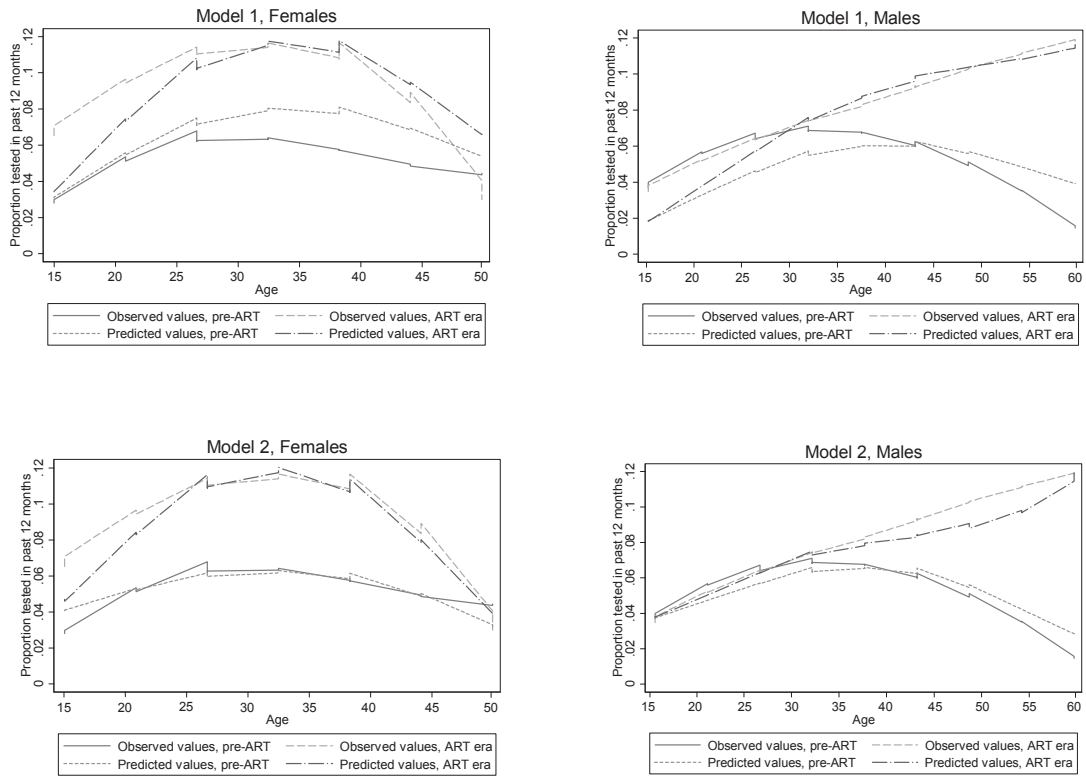
Notes: Locally weighted smoothed regression of distance to the nearest voluntary counseling and testing (VCT) site on age, estimated separately for females and males. Location of Zambia Sexual Behavior Survey (ZSBS) respondents calculated as centroid of Standard Enumeration Area (SEA) of residence. Location of nearest VCT site comes from the Japanese International Cooperation Agency (JICA) 2004 Health Facilities Census (HFC).

Figure 6: Demand for HIV Testing and for Risky Behavior



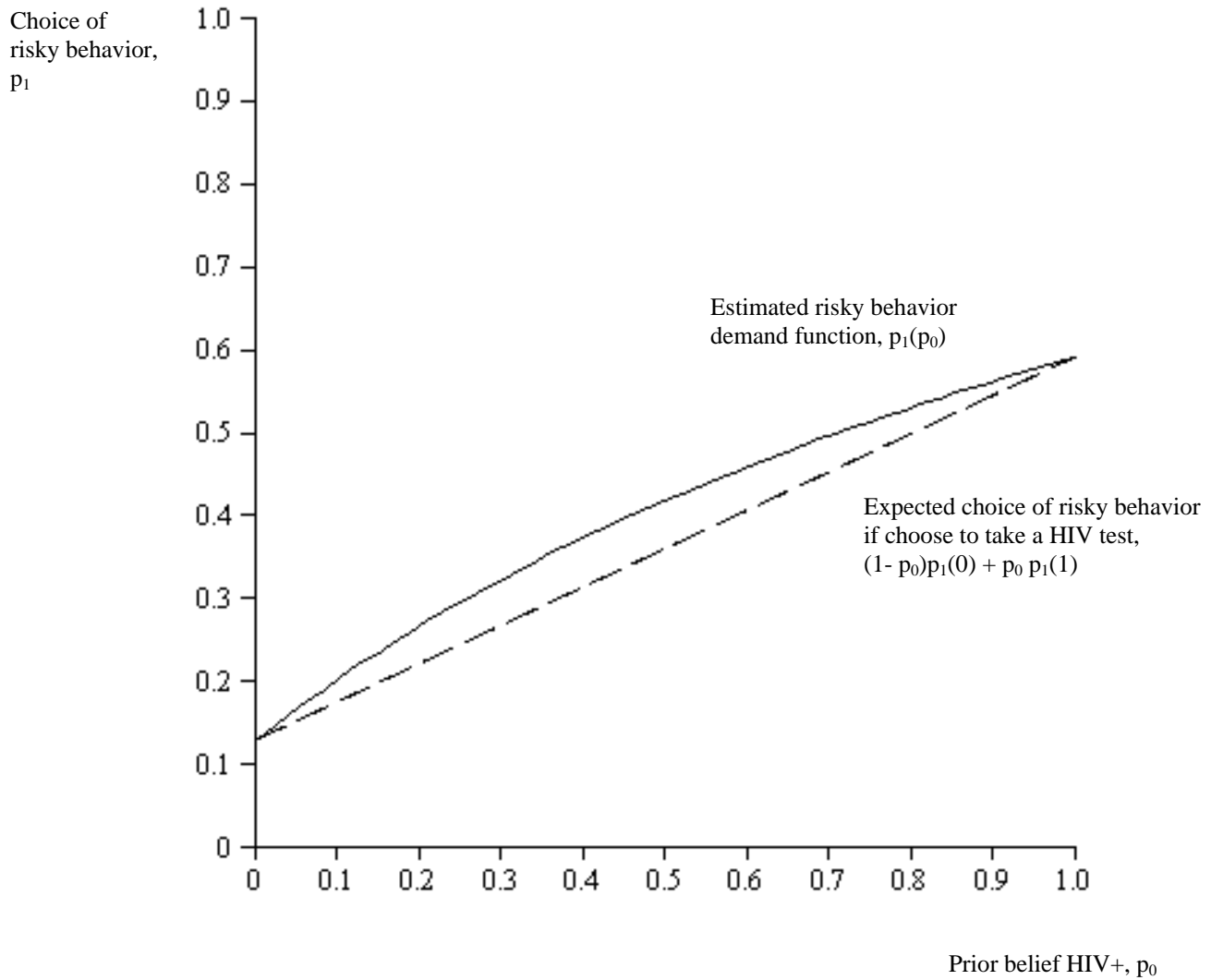
Notes: p_0 denotes prior belief HIV+. $p_1(p_0)$ is the risky behavior demand function. $V(p_1(\cdot))$ denotes utility from risky behavior. θ denotes utility cost of acquiring HIV. c denotes utility cost of taking a HIV test. α denotes net utility of beginning ART.

Figure 7: Predicted and Observed Testing-Age Profiles



Notes: Locally weighted smoothed regression lines of predicted and observed testing rates for five-year age groups on age, estimated separately for pre-ART and ART eras. Model 1 uses the parameter estimates from a regression model with the variables in the structural model and no controls. Model 2 uses the parameter estimates from a regression model with the variables in the structural model, an ART era fixed effect, a male fixed effect, five-year age group fixed effects, and province fixed effects.

Figure 8: Estimated Risky Behavior Demand Function



Notes: p_0 denotes prior belief HIV+. $p_1(p_0)$ is the estimated risky behavior demand function using the structural parameter estimates from the model with the full set of controls.

Table 1: Parameter Estimates

Dependent variable:	tested in past 12 months	
	(1)	(2)
b	1.656	1.652
e	0.588	0.589
c	0.083	0.082
θ	0.850	0.853
α	0.986	0.981
σ	0.955	0.976
δ	0.020	0.020
post dummy	no	yes
male dummy	no	yes
age group controls	no	yes
province controls	no	yes
Observations	288	288

Note: Tested in past 12 months is a binary variable indicating took a HIV test in the 12 months period preceeding the survey. Post dummy is an indicator variable equal to one if the observation comes from the 2005 survey round of the ZSBS. Age group controls are indicator variables for five-year age groups (i.e., 15-19, 20-24, ..., 55-59).

Table 2: Testing Rates Under Alternative Policy Scenarios

Empirical specification:	model 1	model 2
	(1)	(2)
<u>Policy scenario</u>		
No ART	5.7	5.8
Status quo	8.0	8.5
Percent change compared to "No ART"	40	47
Expand supply	10.1	11.2
Percent change compared to "No ART"	77	93
Eliminate inefficiency	18.6	16.9
Percent change compared to "No ART"	226	191
Expand supply and eliminate inefficiency	28.1	25.4
Percent change compared to "No ART"	393	338

Note: Testing rates reported as percent of females age 15-49 and males age 15-59 choosing to take a HIV test. Under the "expand supply" scenario, the supply of ART is expanded to eliminate the random rationing mechanism. Under the "eliminate inefficiency" scenario, the total quantity of ART supplied is the same as in the "status quo" scenario and the non-random rationing mechanism is eliminated. Model 1 uses the parameter estimates from a regression model with the variables in the structural model and no controls. Model 2 uses the parameter estimates from a regression model with the variables in the structural model, an ART era fixed effect, a male fixed effect, five-year age group fixed effects, and province fixed effects.

Table 3: Mean Choice of Risky Behavior Under Alternative Policy Scenarios

Empirical specification:	model 1	model 2
	(1)	(2)
<u>Policy scenario</u>		
No ART	0.217	0.223
Status quo	0.216	0.222
Percent change compared to "No ART"	-0.46	-0.45
Expand supply	0.215	0.221
Percent change compared to "No ART"	-0.92	-0.90
Eliminate inefficiency	0.211	0.217
Percent change compared to "No ART"	-2.76	-2.69
Expand supply and eliminate inefficiency	0.207	0.214
Percent change compared to "No ART"	-4.61	-4.04

Note: Mean choice of risky behavior is mean probability individual acquires HIV through choice of risky behavior, where all individuals treated as if HIV negative at time of choice. Under the "expand supply" scenario, the supply of ART is expanded to eliminate the random rationing mechanism. Under the "eliminate inefficiency" scenario, the total quantity of ART supplied is the same as in the "status quo" scenario and the non-random rationing mechanism is eliminated. Model 1 uses the parameter estimates from a regression model with the variables in the structural model and no controls. Model 2 uses the parameter estimates from a regression model with the variables in the structural model, an ART era fixed effect, a male fixed effect, five-year age group fixed effects, and province fixed effects.

Table 4: New Infections Under Alternative Policy Scenarios

Empirical specification:	model 1	model 2
	(1)	(2)
<u>Policy scenario</u>		
No ART	907	934
Status quo	894	918
Percent change compared to "No ART"	-1.43	-1.71
Expand supply	882	902
Percent change compared to "No ART"	-2.76	-3.43
Eliminate inefficiency	829	865
Percent change compared to "No ART"	-8.60	-7.39
Expand supply and eliminate inefficiency	773	815
Percent change compared to "No ART"	-14.77	-12.74

Note: New infections reported in thousands. Under the "expand supply" scenario, the supply of ART is expanded to eliminate the random rationing mechanism. Under the "eliminate inefficiency" scenario, the total quantity of ART supplied is the same as in the "status quo" scenario and the non-random rationing mechanism is eliminated. Model 1 uses the parameter estimates from a regression model with the variables in the structural model and no controls. Model 2 uses the parameter estimates from a regression model with the variables in the structural model, an ART era fixed effect, a male fixed effect, five-year age group fixed effects, and province fixed effects.