Antiretroviral Drug Access and Behavior Change

Willa Friedman*
Department of Economics, University of California, Berkeley

February 28, 2013

Abstract

Access to antiretroviral (ARV) drugs in Sub-Saharan Africa has rapidly expanded - from fewer than 10,000 people treated in 2000 to more than 8 million in 2011. To measure the impact of this expansion, it is necessary to identify the behavioral response of individuals to drug access. This paper combines geocoded information about the timing of introduction of ARVs in all Kenyan health facilities with two waves of geocoded population surveys to estimate the impact of proximity to an ARV provider on risky sexual behavior. Using a difference in differences strategy that matches survey clusters geographically across waves, I find a relative increase in risky behavior as reflected in pregnancy rates (increase of 82%) and self-reported recent sexual activity (increase of 40%) among young women in areas in which ARVs were introduced between 2004 and 2008. The full impact of ARV access on new infections is estimated through a simulation procedure that combines estimated behavioral responses to ARVs with medical evidence regarding HIV transmission. An increase in ARV drug access is predicted to reduce the rate of new infections despite the induced increase in risk-taking.

JEL Classification: I12, I18, J13, O15

*I am particularly grateful to Edward Miguel, Ernesto Dal Bo, Frederico Finan, and Justin McCrary for helpful advice. The following individuals provided valuable help and support in the collection of and access to the data: Megumi Gordon, Michael Laverty, Patrick Mwangi, Ibrahim Mohamed, and Susan Njogo. I also thank Dan Bennett, Josh Blumenstock, Mariana Carrera, Damien de Walque, Lauren Falcao, Jane Fortson, Tadeja Gracner, Erick Gong, Jonas Hjort, Ashley Langer, Elaine Liu, Jeremy Magruder, Vikram Maheshri, Jamie McCasland, Nicholas Wilson, Andy Zuppann, and various seminar participants for their feedback. Financial support for this project was received from the 2011 Weiss Family Fellowship through the Center for Evaluation and Global Action and the John L. Simpson Memorial Research Fellowship from the Institute of International Studies.

1 Introduction

The HIV epidemic has had an enormous impact on the well-being of millions of people in developing countries. High HIV prevalence rates are associated with falling life expectancy, substantial reductions in human capital accumulation (Ferreira and Pessoa, 2003; Lorentzen et al, 2008; Fortson, 2011), reduced intergenerational human capital transmission (Beegle et al, 2008; Bell et al, 2006; Hunter and Williamson, 2002), and reduced economic growth (Cuddington, Hancock, 1994; Corrigan et al, 2005). The introduction and rapid expansion in access to antiretroviral Drugs (ARVs), which can extend the lives of HIV positive individuals by approximately ten years, is a substantial technological innovation that has changed the course of the epidemic. While ARVs clearly benefit infected individuals and their dependents by delaying the onset of symptoms and revitalize the workforces of many developing countries, ARV provision also shapes future infection rates. Any estimation of the impact of ARVs on future HIV infections fundamentally depends on individual behavioral responses to treatment availability. The direction of this response is theoretically ambiguous because while the cost of infection has gone down, perceptions of the likelihood of infection could increase or decrease depending on beliefs about the impact of ARVs on transmission probabilities. As these beliefs cannot be observed directly, the behavioral response to ARV access must be measured empirically.

This paper uses an original dataset linking individual behavior from two waves of Demographic and Health Surveys (DHS) with a record of the roll-out of ARVs in Kenya to estimate how individual risk-taking responds to ARV access. Using a difference in differences framework with geographically identified survey clusters matched across rounds, I estimate the response to be an 82 percent increase in pregnancies and a 40 percent increase in self-reported risky sexual behavior in the previous 4 weeks. As resulting new infections cannot be empirically identified directly, this paper combines these estimates of the behavioral response to ARV access with medical evidence about a reduction in transmission probabilities for those taking ARVs to simulate the impact of ARV introduction on new infection rates. A sufficiently high level of ARV provision can outweigh even this substantial increase in risk-taking, even with a conservative estimate of the reduction in transmission probability.

A simple theoretical framework demonstrates that the direction of the change in risk-taking in response to ARV access is ambiguous. On the one hand, models of behavioral disinhibition predict

that when individuals are faced with an exogenous decrease in the riskiness associated with an activity over which they have some control, they may take on additional risk (e.g. Peltzman, 1975). In the case of ARV access, individuals who learn that treatment will be available may engage in more risky behavior. This constitutes a specific example of moral hazard associated with access to treatment, which implies that individuals with greater expected access to ARVs would be more likely to risk HIV infection than those who do not anticipate that treatment would be available. On the other hand, ARV access also changes both the true and the perceived probability of becoming infected. While ARV provision means more infected individuals are alive and presumably in the pool of potential sexual partners (Lakdawalla et al, 2006), the medical literature has also demonstrated that these treated individuals have lower transmission probabilities. Perceptions of this can differ widely as some believe that there is no reduction in transmission probability, and others believe that the reduction is complete. This belief determines the direction of the change in the likelihood of becoming infected when ARVs are available.

Estimating the impact of access to ARVs on risk-taking presents a few key challenges that need to be addressed in order to obtain credible estimates. The first challenge to address is the definition of access to ARVs. Self-reported measures of awareness of ARVs introduce endogenous variation in individual characteristics. But proximity to an ARV facility provides an exogenous source of variation in treatment access.[1] Any measure of access based on proximity will inevitably incorporate some misclassification. However, proximity can be thought of as an instrument for access to information about treatment availability, although the first stage cannot be estimated with these data. I exploit detailed geographic information and use the location of respondents relative to health facilities providing ARVs as a proxy for access to treatment. Two measures of proximity are used. First, distance to the nearest facility defines proximity. The primary analysis uses a threshold of 8 kilometers (5 miles) to maximize power, although the results are robust to alternative thresholds. Alternatively, access is defined as being within the same administrative geographic division as a health facility that provides ARVs.[2]

The second challenge is to define a reasonable comparison group to serve as a counterfactual for

[1] In this context, an experiment may not be appropriate as informing people about the presence of ARVs in order to measure whether this encourages additional risk-taking would raise substantial ethical concerns.

[2] A division is the smallest administrative unit in Kenya, with an average size of 2181 km^2. The average size of divisions that are not excluded and contain at least 2 DHS clusters is 2007 km^2

those with access. I use a difference in differences identification strategy with geographic matching to deal with unobserved time-invariant differences across areas. As different villages were surveyed in each wave of the DHS, I use location to match observations across rounds. In the main specification, I match clusters of observations from each wave with those from the nearest clusters from the other wave. With multiple matches, this presents a reasonable counterfactual with which to estimate the treatment effect. This will be explained in more detail in Section 4. A simpler specification is also presented that compares within administrative divisions, using division fixed effects to address time-invariant unobserved differences.

The third challenge is that endogenous placement of ARV facilities raises concerns about omitted variables. Based on policy documents from the Kenyan Ministry of Health, I control explicitly for various factors that were used in targeting facilities for ARV introduction, including HIV rates, urban-rural status, and proximity to other health facilities. Difference in difference estimation addresses time-invariant differences across areas, but it relies on the assumption that in the counterfactual world without ARVs, trends in the control and treatment areas would have been comparable. I use a historical birth register to show that trends in pregnancy rates in treated and control areas were parallel for two decades before ARVs were introduced.

The fourth challenge to be addressed is that, while the outcome of interest in this study is sexual risk-taking, sexual behavior is notoriously misreported (e.g. Jamison and Karlan, 2012; Minnis et al, 2009). To address this, I rely primarily on pregnancy as a proxy for unprotected sexual activity. Pregnancy is a particularly appropriate proxy in this country for a few reasons. First, unlike in many developed countries, in Kenya, as in most of Sub-Saharan Africa, HIV is a generalized epidemic, predominantly spread through heterosexual sex. Second, while abortion exists, it is illegal, and therefore relatively less common. Indeed, the use of pregnancy as a marker of unprotected sex is a commonly used strategy (e.g. Duflo et al, 2011; Dupas, 2011). Still, I will also report impacts on self-reported recent sexual activity.

I estimate a statistically significant relative increase in pregnancy and self-reported sexual activity in areas where ARVs were introduced among women aged 15-18. The point estimate of the treatment effect on pregnancies is 6.56 percentage points, or an increase of approximately 82 percent relative to the fertility rate in control areas. The impact on sexual activity in the previous four weeks is estimated to be a 5.72 percentage point increase or a change of 40 percent. I focus

predominantly on this young demographic as they are the least likely to be in stable relationships, and therefore the most likely to change their willingness to have unprotected sex in response to changes in the threat of HIV infection. This result is consistent across different age and distance thresholds.

A final concern to address is the extent to which alternative mechanisms could explain the observed relationship between ARV access and fertility. In particular, a change in fertility preferences from an increase in life expectancy could generate the observed changes in pregnancy. However, if this were the case, we would expect to also see changes in fertility among married women, yet there is no evidence of a change in behavior among those who are married and no changes in other measures of fertility preferences or access to family planning. Another alternative mechanism relies on an increase in HIV testing that facilitates sero-sorting, namely matching among individuals with the same HIV infection status. I show that the results hold for those who have not been tested and that the size of the population that could sero-sort is sufficiently small that this cannot drive the primary empirical results.

It is not currently possible to empirically estimate the impact of ARV provision on new HIV infections with a purely quasi-experimental approach for reasons related to the biology of HIV transmission and infection. First, the full change in new infections will not be realized immediately, and it is therefore too soon after the introduction of ARVs to measure the full impact. Second, estimating the impact on *new* infections would require distinguishing between new and old infections. As ARVs keep those with HIV alive longer, there will be a mechanical relationship between their introduction and the prevalence of HIV in the population, even if there is no impact on *new* infections. Yet distinguishing between new and old infections is infeasible.

A simulation, incorporating both medical evidence and the behavioral estimates of this paper can provide a reliable prediction of the impact of ARVs on new infection rates. It combines a range of estimates of the reduction in transmission probabilities found in the medical literature with a substantial increase in risk-taking as drugs are made available.[3] I find that even a conservative estimate of the reduction in transmission probabilities can outweigh the effects of a large increase

[3]Previous simulations undertook a similar exercise, but without estimates of either the reduction in transmission rates or of a change in behavior, they were somewhat inconclusive, although the authors suggested that an increase in risky behavior had a significant chance of outweighing the reduction in transmission probabilities (e.g.: Blower et al, 2000; Law et al, 2001).

in risk-taking if a sufficient fraction of those who are positive are treated, predicting reductions in HIV infection from an expansion in ARV access.

This paper provides a test of the theory of risk homeostasis (Peltzman, 1975), which posits that individuals may respond to a decrease in the riskiness of an action by increasing their choice of that action. This risk offset hypothesis is similar to theories of *behavioral disinhibition* due to changes in risk, to theories of *risk compensation*[4] mentioned in the public health literature, and *moral hazard* associated with treatment access. Previous empirical work has found evidence for risk homeostasis in the context of drivers' response to auto safety innovations (Winston et al, 2006).

However, in the context of HIV risk-taking, empirical tests of theories of risk offsetting have found surprisingly little supporting evidence. For example, studies have found no expected responses in risk-taking from information about male circumcision and HIV risk (Godlonton et al, 2011; Wilson et al, 2011).[5] Estimates of the behavioral response to HIV risk generally have found small or no impacts on sexual behavior (Oster, 2009) or fertility (Fortson, 2010; Juhn et al, 2009; Kalemli-Ozcan and Turan, 2010), although Young (2005, 2007) does find a reduction in childbearing associated with HIV prevalence.

A few recent papers have explored the impact of antiretroviral drugs on risk-taking with mixed results. Two studies in the US use variation in behavior among gay men before and after ARVs became available in the US, both finding an increase in risk-taking after their introduction (Mechoulan, 2007; Papageorge, 2012). In Sub-Saharan Africa, where the overall HIV rate is higher, life-expectancy is lower, and access is still limited, ARVs may affect a wider range of outcomes with greater policy implications. In Malawi, Baranov et al (2012) use a method similar to this paper, but rely on measures of risk-taking largely determined before ARVs became available, and they find no impact using the entire sample. This paper focuses on recent behavior among those who could change their actions (young women) resulting in different findings. de Walque et al (2012) study the impacts of beliefs about ARV effectiveness on risk-taking and find a behavioral response, but relying on self-reported beliefs about ARVs introduces concerns about endogeneity.

Although other papers have estimated the impact of ARV access on those who are HIV positive

[4]This term is commonly used but should not be confused with risk compensation in the labor economics literature referring to increased wages paid to employees asked to undertake greater risks.

[5]Male circumcision is associated with a dramatic reduction in the risk of HIV infection (Auvert et al, 2005; Bailey et al, 2007; Gray et al, 2007).

(Bor, 2012; Lakdawalla et al, 2006; Thirumurthy et al, 2008, Thirumurthy et al, 2012), and the impacts on other outcomes including employment (McLaren, 2012), mortality risk-perceptions and productivity (Baranov et al, 2012), human capital investments (Baranov and Kohler, 2012), child health (Lucas and Wilson, 2013), and HIV testing (Wilson, 2010), this paper presents the first causally identified estimates of the impact of ARVs on risk-taking in a context with a generalized HIV epidemic, and this is the behavioral outcome that will determine the course of the epidemic.

This paper proceeds as follows. In section 2, I outline a theoretical framework to formalize the intuition driving the empirical estimation and to demonstrate how the empirical estimates will drive the final simulation. Section 4 describes the data and the context in which it was collected, and the empirical methods are outlined in section 5. Section 5 discusses the main results, and in section 6 I simulate the rate of new infections as a function of the level of ARV distribution, incorporating both mechanical impacts from the medical literature and the behavioral responses estimated in section 5. I conclude in section 7.

2 Theoretical Framework

The theoretical framework presented in this paper builds on the behavior change literature applied to responses to information about HIV. In an early model, Kremer (1996) argues that high HIV prevalence may dissuade those who are low-risk and least-likely to be infected from participating in sexual activity at all while causing those who are less cautious to take more risks because of the low probability of remaining negative. This can generate multiple equilibria at different risk levels. More recently, Gong (2012) shows that HIV testing changes behavior differentially for individuals with different priors about their own status, finding support in data from an early randomized offering of HIV testing in East Africa. Kerwin (2012) constructs a new model that rationalizes a type of fatalism based on previous risk-taking that can generate non-monotonic responses to changes in risk. This model helps to explain a pattern observed in Malawi in which individuals sufficiently overestimate the likelihood that they are currently infected, and stop taking precautions (e.g. Kaler, 2003).

In the framework developed in this paper, individuals from an infinite population of agents of size 1 choose whether or not to have unprotected sex by weighing the individual-specific benefit

from unprotected sex against the expected costs of HIV infection. Access to treatment can change perceptions about both the likelihood of infection and the cost of becoming infected.[6]

The rate of new infections among those previously uninfected, I, is equal to the probability of infection conditional on engaging in unprotected sex, p, multiplied by the proportion of the uninfected population that chooses to do so, A_1. Treatment availability directly affects p by changing the pool of potential partners and their infectivity and indirectly affects both p and A_1 through a behavioral channel.

Individuals can be categorized into three types: 1) Type 1 is HIV negative, 2) Type 2 is HIV positive, without treatment, and 3) Type 3 is HIV positive, with treatment.

I make the following assumptions throughout:

- Each individual has full information about his or her own status. This assumption is included to make the model tractable and to focus on aspects which can be addressed in the empirics. The focus of the analysis is young women, who are likely to accurately perceive that it is unlikely that they are currently HIV positive. This population is old enough that infection from birth is nearly impossible, yet they are young enough that they have not or only recently began having sex. Up until that point, the likelihood of infection was approximately zero, although it could have recently changed. Throughout this section, the impact of weakening the assumption of full information about own status will be directly addressed.

- Each individual knows the distribution of other types among potential sexual partners, but does not observe the status of any particular potential partner.

Each individual chooses **whether to have unprotected sex** based on an individual-specific utility from unprotected sex (incorporating everything including social pressure and desire for children, etc.). Those who are HIV negative also consider the likelihood of becoming infected and the associated utility cost of infection.

Type 1: Formally, an uninfected individual will choose to have unprotected sex if:

$$\theta_i + (1-p) \cdot u^- + p \cdot u^+ \quad > \quad u^- \tag{1}$$

[6]For simplicity, I assume that individuals who have access to ARVs know that they have access and that those who do not do not anticipate future access. This is plausible if proximity brings with it information about the existence of ARVs. I discuss the empirical implications of this assumption in section 3.

where u^- represents the continuation value of staying negative, u^+ represents the continuation value of being positive, and p represents the probability of infection from unprotected sex. θ_i is an individual-specific taste parameter, distributed with cdf, F_θ, which encompasses all non-HIV-related costs or benefits of unprotected sex relative to the alternative. The alternative can be abstinence or protected sex.[7] To be clear, this parameter can also be negative. Rewriting inequality 1 as $\theta_i > p \cdot (u^- - u^+)$, it follows that the proportion of the population that is negative (Type 1) that chooses to have unprotected sex can be written as:

$$A_1 \quad = \quad 1 - F_\theta(p \cdot (u^- - u^+)) \tag{2}$$

Note that ARV availability may change two components of the above equation. First, it reduces the relative cost of becoming infected, $u^- - u^+$, by extending the HIV positive life expectancy. This alone would lead to an increase in risk-taking among individuals of Type 1. However, ARV access can also affect p by changing the population of potential sexual partners. The direction of this effect is ambiguous.

If individuals do not know their HIV status, then the impact of ARV access will be dampened, but the sign will remain the same. If an individual believes that the probability he or she is HIV positive is π, then inequality 1 can be rewritten as

$$\theta_i + (1-\pi) \cdot [(1-p) \cdot u^- + p \cdot u^+] + \pi \cdot u^+ > (1-\pi) \cdot u^- + \pi \cdot u^+ \tag{3}$$

Although this changes the threshold of θ_i over which an individual chooses to have sex, it does not change the direction of the effect of ARV access via the probability of infection from sex or the cost of infection. If, however, drugs change whether people get tested for HIV, then this raises a further complication which is addressed later in the paper.

Types 2 and 3: Those who are already HIV positive do not risk changing their HIV status, and thus the only parameter in their utility optimization is the individual-specific utility from unprotected sex.[8] Altruism, morbidity, fatalism, desire for children, or any other channel through

[7] A number of papers have found evidence of a higher willingness to pay for unprotected sex among those who visit sex workers (e..g.: Gertler et al, 2005; Rao et al, 2003, and Robinson and Yeh, 2011).

[8] Those who are HIV positive do risk re-infection from having sex with another person who is HIV positive. This can moderately increase the speed of the progression of HIV into full-blown AIDS. However, this can credibly be

which treatment changes the utility from unprotected sex for those who are positive can be incorporated into the model by allowing this taste parameter for Types 2 and 3 to be drawn from different distributions.

Thus, an individual of Type 2 (HIV positive, not on treatment) will choose to have unprotected sex if $\gamma_i > 0$, and an individual of Type 3 (HIV positive, on treatment) will choose to have unprotected sex if $\omega_i > 0$, where γ_i and ω_i, are individual-specific taste parameters distributed with cdfs, F_γ and F_ω, respectively. These parameters can be positive or negative, incorporating any utility gains or losses from unprotected sex.

It follows that the proportion of the population that is positive and not on ARVs (Type 2) that chooses to have unprotected sex can be represented as

$$A_2 \quad = \quad 1 - F_\gamma(0) \tag{4}$$

and similarly, the proportion of the population that is positive and on ARVs (Type 3) that chooses to have unprotected sex can be represented as:

$$A_3 \quad = \quad 1 - F_\omega(0) \tag{5}$$

The assumption that those of Types 2 and 3 will not change their behavior in response to treatment access *of others* depends on the claim that while HIV positive individuals bear a utility cost from the possibility of infecting someone who is negative (i.e.: they are altruistic), altruism will have only limited behavior-change consequences. This assumption depends critically on the marginal changes in the probability that one's sexual partner is negative. Where prevalence rates in Kenya are somewhere between 5 and 15 percent, the probability of a heterogeneous match for someone who is HIV positive (i.e.: an HIV negative partner) is much higher than the probability of a heterogeneous match for someone who is HIV negative (i.e.: an HIV positive partner). Changes in the composition of the pool of potential partners induced by the medical life extension of ARVs is therefore proportionally small for those who are positive and proportionally large for those who are negative. Further, supposing that those who are positive are very likely to draw a negative sex

assumed to be negligible with no loss to the applicability of the model.

partner, the altruistic calculation of the cost of infecting someone on the basis of ARV availability is second order for those who are positive (who naturally discount costs for others relative to own benefits) where it is arguably quite substantial for those who are negative.

Another concern with this formulation is that many do not know their status.[9] Again, I expect this to have only a negligible effect overall. First, those of Type 3 necessarily know their status as they are receiving treatment. If some Types 1 and 2 do not know their status, then this will dampen any impact on behavior among those who are negative. As long as some of the Types 1 and 2 knows their status or the two types have different perceptions of the chance that they are infected, then the behavioral response among those of Type 2 will be smaller than those of Type 1. An increase in sexual activity among those who are positive and untreated will feedback, decreasing the utility from unprotected sex among those of Type 1. Without full information, the impact of treatment on behavior in the two types must go in the same direction. However, even if the two types share identical beliefs, this will dampen, but not change the sign of any other impacts.

The probability of becoming infected from unprotected sex, p, depends on the proportion of each type among potential sexual partners and the likelihood of transmission from each type. Denote by N_j the size of the population of each type, because the transmission probabilities can be different with these two groups.

Let q be the reduction in infectivity due to ARVs, and let $\hat{q}$ by individuals' beliefs about q.[10] If individuals believe that ARVs fully eliminate the risk of transmission, then $\hat{q}$ is 0. On the other hand, if individuals are unaware of the reduction in infectivity, then $\hat{q} = 1$.[11] For an individual of Type 1, the likelihood of infection if their partner is of Type 2 is r and the likelihood of infection if the partner is of Type 3 is $r \cdot q$.

The likelihood of infection from unprotected sex can therefore be written as:

$$p = r \cdot \frac{A_2 N_2 + A_3 N_3 q}{A_1 N_1 + A_2 N_2 + A_3 N_3} \tag{6}$$

[9] Of those who tested positive in the 2008/2009 wave of the DHS in Kenya 29 percent had never been tested for HIV previously, and so likely did not know their status.

[10] Based on the medical literature, q could be as small as 0.04 ((Cohen et al, 2011) so the reduction in infectivity from treatment could be quite large. However, individuals respond to their beliefs, $\hat{q}$, which could be anywhere between 0 and 1.

[11] In informal conversations with HIV clinic employees, this was a commonly held belief. Many expressed concern that people who were HIV positive had become healthy and fat and were at risk of infecting others.

and by analogy, the perceived likelihood of infection is:

$$\hat{p} = r \cdot \frac{A_2 N_2 + A_3 N_3 \hat{q}}{A_1 N_1 + A_2 N_2 + A_3 N_3} \tag{7}$$

Changes in access to ARVs affect p by changing the relative sizes of the population of Types 2 and 3 and the proportion of those who are negative who engage (A_1).

Let D represent the share of those who are positive who receive treatment, and let M be the share of the population that was infected as of the beginning of the current period. Besides the possibility of different behavioral parameters, γ_i and ω_i as outlined above, individuals of type 2 and type 3 have different death rates (d_2 and d_3 respectively), as the primary function of ARVs is to keep HIV positive individuals alive. Therefore the size of each population is:

$$\begin{aligned} N_1 &\quad \text{is fixed from the previous period.} \\ N_2 &= M \cdot (1 - D) \cdot (1 - d_2) \qquad (8) \\ N_3 &= M \cdot D \cdot (1 - d_3) \qquad (9) \end{aligned}$$

and we know that $d_2 > d_3$. If treatment is unavailable then $D = 0$ and $N_3 = 0$, and if everybody who is positive receives treatment, then $N_2 = 0$.

An increase in D decreases the cost of becoming infected ($u^- - u^+$), and it changes $\hat{p}$, the perceived likelihood of becoming infected. The sign of this is ambiguous and depends on other parameters.

In particular, if $\hat{q} = 0$, then:

$$\frac{d\hat{p}}{dD} < 0 \tag{10}$$

This is intuitive because every impact of drug provision on p moves it in the same direction. First, with the elimination of infection of those on treatment, the size of the infectious population is necessarily smaller, reducing the likelihood of matching with someone who is infectious. Second, if individuals respond to the reduction in risk from fewer positive matches or from the reduction in the cost of infection, then $A1$ will increase as well, which will further reduce p.

On the other hand, if $\hat{q} = 1$, then the impact of drugs on the likelihood of infection is more complicated. With no reduction in transmission probabilities but a reduction in the mortality probability of those infected, there will be an increase in the size of the infectious population in the pool of potential partners. This will increase p. On the other hand, if the reduction in the cost of infection sufficiently increases A_1 (the fraction of the negatives who choose to have sex), then this could reduce p. Which effect will dominate cannot be determined theoretically because it depends on the response to the perceived cost of infection. If the first effect dominates and p increases, then the effect of drugs on A_1 also becomes ambiguous.[12].

While ARV availability unambiguously decreases the cost to the individual of infection, the sign of the impact of ARV availability on the perceived probability of infection is ambiguous as is the relative magnitude of the cost reduction to the positive or negative change in the perceived probability of infection. Therefore the impact on the likelihood of those who are negative engaging in unprotected sex is ambiguous. The empirical section will estimate this revealed decision.

The theoretical framework was set up in part to show how drugs change new infections directly and through changes in behavior. As previously stated, the infection rate is:

$$I = A_1 \cdot p \tag{11}$$

All parameters that contribute to the above equation can be taken from the existing medical literature, with the exception of the behavioral response to treatment, which determines A_1, and indirectly, p. This response will be measured in the empirical analysis of this paper, and then this estimated response will be used to predict the impact of drugs on new infections.

3 Data and Context

Antiretroviral drugs were developed during the 1980s and became widely available in developed countries in the 1990s. Because of prohibitively high prices, they were almost completely unavailable to residents of Sub-Saharan Africa until the last decade. In the early 2000s, a number of agreements between developing countries and pharmaceutical companies reduced the prices of

[12]If individuals do not know their own HIV status, then A_2 will move in the same direction as A_1, which will reduce the magnitude of, but not change the sign of $\frac{d\hat{p}}{dD}$

ARVs for governments of developing countries. Since then, the price of ARVs paid for by these governments has fallen from more than $10,000 per person per year to under $70 per person per year. With funding from governments and international organizations, ARVs are provided free of charge to eligible patients in Kenya and most other Sub-Saharan African countries.

As reported in Table 1, Kenya has a relatively high rate of HIV infection (6.3% in 2009), and it has seen a large and rapid expansion in access to ARVs in the last decade. In the early stages of the roll-out, the Ministry of Health and other associated government organizations outlined plans to provide geographically dispersed access through capable pre-existing facilities. Although initially only large hospitals were considered to have all the necessary staff and equipment to provide treatment, the requirements for facilities to be designated as capable have been reduced. In 2004, only 7 facilities distributed ARVs in Kenya but this increased substantially to 336 in 2008 (Figures 1a and 1b). Treatment is free for those who are HIV positive and eligible.[13]

Some locations were more likely than others to have ARVs introduced, and the empirical analysis will address these. This includes urban areas and areas with high rates of HIV. Because distribution happened through existing facilities, areas with large hospitals were more likely to distribute ARVs, while areas without nearby health facilities were less likely. The DHS data used in this paper provides the best existing estimates of regional HIV prevalence, and the Kenya Open Data Initiative provides a record of the GPS locations of all health facilities currently in Kenya. This information is included in the analysis to address potential endogeneity from location of ARV sources.

Information about ARV access comes from an original dataset constructed using administrative records obtained from meetings with government and NGO officials in Kenya. The geographic information comes from the Kenya Open Data Initiative,[14] and the timing information comes from reports provided by KEMSA, a procurement agency, and the National AIDS and STI Control Program (NASCOP) of the Ministry of Health. This combined database of health facilities that currently provide ARVs includes information for each facility on the year ARV distribution began and the location of the facility.

I hand matched clinic information across data sources by the name and district of each facility.

[13]Eligibility was initially based on assessments of whether a person was expected to be able to adhere to the medicine, and the progression of the disease. Now the primary metric for eligibility is the progression of the disease. Initially a person was eligible with a CD4 count below 200, but the WHO has increased the threshold to 350.

[14]See opendata.go.ke

The first instance in which a health facility appears in any records is used as the year in which treatment became available.[15] Table 1 shows the number of health facilities and the number of individuals receiving treatment in each year.

The data on individual behaviors come from two waves of geocoded Demographic and Health Surveys (DHS) from 2003 and 2008/2009, [16] which will be referred to throughout the paper as Wave 1 and Wave 2 respectively. Kenya expanded treatment availability largely between 2006 and 2009, so these waves provide information from before and during the middle stages of the expansion. Columns 3 and 4 of Table 1 shows the number of women and the number of clusters in each survey. Each cluster contains an average of 18 households and 21 female respondents. The analysis will focus on women ages 15-18 in order to look at a population that is most likely not to be in stable partnerships. Those who are already married are less likely to change their behavior in measurable ways.[17] I also exclude Nairobi and other areas which were reported to have ARV access in 2004 to mitigate concerns regarding the endogeneity of ARV access. Summary statistics of relevant variables are reported for the sample used in the analysis in Table 2. For clarity, all percentages are reported out of 100.

A few characteristics of the sample should be noted. First, a relatively small fraction of the sample of young women is HIV positive, but treated areas have higher prevalence rates, which will be addressed in the analysis. Nearly the entire sample in both rounds (between 97% and 100%) in both treatment and comparison areas have heard of HIV, and approximately two thirds report that they know someone who currently has or has died of AIDS. Testing increased between rounds in both areas, with a somewhat larger increase in treatment areas, which is consistent with the findings of Wilson (2010). Among both groups, only a very small fraction report STD symptoms or multiple partnerships.

The DHS data contain responses to questions about childbearing and recent sexual activity. There is extensive evidence of misreporting of sexual activity from direct survey questions (e.g.

[15] In conversations with officials working on Monitoring and Evaluation of ARV distribution, I was not told of any health facilities that stopped distributing drugs unless they were replaced by another organization in the same location.

[16] Interviews in the second wave were conducted between November 2008 and March 2009.

[17] While those in stable relationships may change their behavior outside of marriage in response to changes in HIV risk, this is more difficult to measure. I cannot determine paternity from the data, and only a small fraction of respondents report having additional partners. Respondents are asked about STIs, but very few report infections or symptoms.

Jamison and Karlan, 2012; Minnis et al, 2009). In this particular dataset, for example, 609 women reported that the age they first had sex was later than the age at which they first gave birth, and of 2096 individuals in both waves who reported that they had never had sex, 24 tested positive for HIV. All individuals in the sample are over age 15 and therefore very unlikely to have been born with HIV, and this rate is well above the error rate of the set of tests used. Because of these concerns about measurement error, childbearing is a commonly used measure of HIV risk-taking (e.g. Duflo et al, 2011; Dupas, 2011). I follow this convention and use current pregnancy as a preferred proxy for unprotected sex and show additional results using self-reported behavior as the outcome variable.[18] Results are also presented with self-reported unprotected sex in the last four weeks as the outcome.

Throughout the analysis, I proxy for information about access to ARVs with the proximity to a facility providing ARVs. A small fraction of the HIV negative population is aware of ARVs before they are introduced in the area. For this group, proximity only marginally increases access by reducing the cost of obtaining treatment. The bulk of the population has no previous information about ARVs until they are first introduced at a nearby clinic. This change in awareness can happen through several channels, including deliberate information campaigns, posters, and billboards announcing the availability of treatment.[19]

4 Empirical Strategy

With two waves of population surveys combined with a record of the roll-out of treatment, the estimation will rely on a difference in differences identification strategy, using multiple definitions of access based on proximity to an ARV facility and methods of identifying the relevant comparison groups across waves.

[18] Those who report having miscarried recently and would have been pregnant (based on the number of months pregnant at the time of the miscarriage) if not for the miscarriage are coded as pregnant. Results do not change if these are not coded as pregnant.

[19] Other individuals may learn about the presence of ARVs from those who have begun treatment either explicitly via word of mouth, or indirectly by observing health improvements of peers who are rumored to be HIV positive. These two channels of information could lead to the formation of different beliefs about HIV infections. In particular, indirect observation could erroneously signal that a cure is available. In the 2006 Uganda DHS, 34% of women who reported that they had heard of ARVs believed that they were a cure for HIV. As this belief is common, it is possible that behavioral responses to proximity to treatment could be driven by an over-estimate of the benefit of ARVs to those who are HIV positive. In this case, if individuals believe that ARVs are more effective than they are, they might respond more than they would have with accurate information.

In all specifications, all observations are weighted using DHS sampling weights, unless otherwise noted, and each specification includes controls for age, education, and district and division HIV rates.[20] Finally, each specification includes controls for urban-rural status, proximity to large and small health facilities, and each of these interacted with wave 2 to allow different trends.

The basic equation I estimate is:

$$Y_{ijt} = \beta_0 * Treat_j * Wave2_t + \beta_2 * Wave2_t + \gamma_j + \sum_{k=3}^{n} \beta_k * X_{kijt} + \epsilon_{ijt} \tag{12}$$

where Y_{ijt} is the outcome, $Treat_j$ is a binary variable that represents whether the respondent is located in an area in which ARVs were available before Wave 2, and γ_j is an area fixed effect. X_{ijt} is a vector of (n-3) individual-specific controls. Each wave surveys different villages, and therefore the definition of an area j cannot be a village. Each specification will define area differently.

In the preferred specifications, $Treat_j$ is defined as being within 8 kilometers of a facility with ARVs by 2008.[21] Because the same villages were not sampled across waves, the relevant comparison group across waves is not obvious. To address this, observations are matched across waves based on their locations using GPS locations to identify precise comparisons and construct a fixed effect analysis within pairs of neighboring survey clusters.

Each survey cluster in wave 2 is linked with the five closest survey clusters from wave 1.[22] For the analysis, each respondent from wave two is included five times and each observation from wave 1 is included as many times as it is matched. Any pair that is more than 100kms apart is dropped.

Using this expanded and matched sample, I estimate another difference in difference estimate with matched-pair fixed effects. Each specification includes a fixed effect for each matched-cluster pair. Each observation is additionally weighted by the minimum of the inverse of the distance and 1/8.[23] Pairs of cluster with different treatment status are dropped in the primary specification, but estimates including these as well are also presented, and do not generate noticeably different

[20] This is constructed using the DHS sample as this is the standard source of information about HIV rates. Each respondent is excluded from the estimate of the HIV prevalence in her area.

[21] Eight kilometers is chosen to maximize power as it is the closest distance to the median. This generates balance between the treatment and control groups that maximizes the precision of the estimates. This distance (approximately 5 miles) is also a reasonable distance to walk for frequent medical care. For robustness, the analysis is repeated using different distance cut-offs with nearly identical results.

[22] Because the locations of villages is jittered and some villages may be sampled twice, it is possible that some of these matched pairs are truly taken from the same villages at two points in time.

[23] This weighting scheme is used in place of the inverse distance so as not to overweight extremely small distances. Because of the jittered data, these distances are not likely to be precise at this level.

results. Dropping the unmatched pairs is comparable to excluding boundaries between areas in spatial analysis. The standard errors are clustered at the level of the survey cluster to correct for the duplication. I also report standard errors corrected for two-dimensional clustering following Cameron et al (2011). One dimension is a cluster from Wave 1 with all observations from Wave 2 with which it is matched, and the other dimension is the opposite. The standard errors are somewhat larger, but not substantially so. The coefficient of interest remains the interaction between $Wave2_t$ and $ARV\ Access_j$.

In a simpler specification, $Treat_j$ is defined as residing within a division in which at least one health facility provided ARVs by 2008. This specification includes division fixed effects and standard errors clustered at the level of the division.[24,25] While divisions can be large, this measure of proximity may reflect reality in that individuals are likely to visit the center of their division for other business, even if they do not live as close. Therefore it is logical that the relevant proximity that would determine the spread of information about a new HIV treatment could be within the same district. The geographic distribution of treated divisions is shown in Figure 2. One weakness of this specification is that observations from divisions with clusters in only one round do not contribute the estimates, so information is lost, which is why the matched specification is preferred.

Robustness is verified using multiple age cut-offs, and results are also reported separately for those married and unmarried. The theoretical framework suggests a change in behavior among those who are HIV negative. The analysis that follows includes a very small fraction of respondents who tested positive for HIV. The results are robust to excluding this group. All estimates include controls for age, education, district and division HIV prevalence, urban-rural status, and proximity to other health facilities, along with survey wave and location fixed effects as described.

The primary assumption to justify the difference in difference specification is that the trends in the treatment and control areas would have been the same in the absence of treatment. Figure 4 plots pregnancy rates in treated and control areas (defined using the 8km distance threshold), before 2003, based on the birth registry in the DHS data. While the levels are not the same, the

[24]During the time between waves, administrative boundaries have shifted. For consistency, I use current borders and place observations within them using their GPS locations.

[25]Due to jittering, 11 clusters were placed outside of the borders of Kenya. These observations were manually linked with the closest administrative division within the country so that they could be included in this analysis.

trends are clearly similar, and we cannot reject that the two curves are parallel.

5 Results

The main results are reported in Panel A of Table 3. Columns 1 and 2 present the results using the specifications with matched clusters of observations. In Column 1, this estimation includes all matches, and Column 2 excludes the pairs with different treatment status from the analysis. The treatment effect is the coefficient on the interaction term, reported in the first row. This shows a treatment effect of 6.7 percentage points. Column 3 presents the specification in which treatment is defined as having a facility with ARVs in the same division, showing a treatment effect of 9.5 percentage points. In all three specifications, the coefficient of interest is positive and statistically significantly different from zero.

Panels B and C of Table 3 repeat the same estimation, using whether the respondent reports that she has had sex in the last 4 weeks as the outcome. In Panel C, the outcome is reporting having had sex in the last 4 weeks and reporting having not used a condom with the most recent sexual partner. In the first and second columns, the treatment effect is measured to be approximately 5 percentage points. While sexual activity would need to change by a larger magnitude to generate the observed changes in pregnancy rates, the lower estimated treatment effects could reflect attenuation from noise resulting from misreporting. In the third column, the coefficient of interest is insignificant, but the point estimate and standard error are both large, so a substantial increase cannot be rejected.

5.1 Fertility preferences

Changes in pregnancy rates could also reflect differences in fertility preferences, questioning the applicability of the proposed theory of risk-taking to explain the observed results. Panel A of Table 4 estimates the impact on other measures of fertility preferences or access to family planning, using the matched specification with only pairs with similar treatment status (Estimation Strategy 3). Column 1 shows the impact on having been visited by a family planning worker as a test of whether the introduction of ARVs also brought changes in the provision of broader reproductive health services. This coefficient is negative, small, and insignificant. The second column estimates the impact on the stated ideal number of children. Treatment areas - which were observed to have

had relatively higher conditional fertility rates - reported lower numbers of ideal children. Columns 3 and 4 find no impacts on the use of birth control, conditional and not conditional on having had sex respectively.

Another way to test whether the impacts on pregnancy reflect changes in general fertility preferences is to look at which segment of the population changes their behavior. If ARVs changed fertility preferences, then we would expect to see a change in fertility among those who are married at least as strongly as among those who are not married. Panel B of Table 4 repeats the main analysis using different subgroups. Column 1 includes married women, and column 2 includes women who have been married for at least 1 year. Column 3 includes women who report that they are cohabiting. In each of these three specifications, the estimated treatment effect is either negative or extremely small and insignificant. However, Column 4 includes those who are unmarried and over 25 (in order to have a completely distinct population from those in the previous estimates), including those who never married or are divorced or widowed. In this specification, the treatment effect estimated is 4.6 percentage points, similar to that estimated for young women. Selecting on these subgroups is problematic because the criteria for selection are potentially endogenous and could themselves be responses to ARV access. Still, these estimates suggest that the measured differences are less likely to be a reflection of changes in fertility preferences, and they may reflect differences in risk preferences with regards to unprotected sex among populations with the possibility for marginal behavior changes.

Columns 1 and 2 of Appendix Table A1 present estimates of the impact on unwanted pregnancies, as these are more likely to reflect changes in risk-taking rather than fertility preferences. I code pregnancies as unwanted if the respondent reports that she did not want to become pregnant or did not want to become pregnant at that time. For those who have recently miscarried but would have been pregnant otherwise, respondents are not asked whether they wanted the pregnancy. I code all pregnancies resulting in miscarriage as unwanted, and the results are nearly identical if these are all coded as wanted. Rates of reported unwanted pregnancies are substantially lower than for all pregnancies, and so the estimated impacts are correspondingly smaller, but still positive and substantial.

5.2 HIV testing

As discussed earlier, Wilson (2010) demonstrates that demand for HIV testing is likely to increase with ARV access. This increase in testing could facilitate partner sorting based on HIV status or *sero-sorting*. This presents an alternative channel by which ARV access increases testing which facilitates sero-sorting, which increases pregnancies among those who know their partners status and thus are not putting themselves at risk of HIV infection. While this could be part of the story, there is evidence that it is not the entire story. First, in this sample, even in the second wave, only 27 percent of those in areas with ARVs had been tested, while 21 percent of those in control areas had been tested. Of those who were tested in treatment areas in wave 2, only one third (or 9 percent of the entire group) had been tested more than one year before the survey. Columns 3 and 4 of Appendix Table A1 repeat the main analysis excluding those who had been tested at least one year before the survey, and the results remain the same. While sero-sorting may marginally contribute to the increase in pregnancy among young women, it cannot explain the observed relative increase in risky behavior in areas that received ARVs.

HIV testing could also change beliefs about own HIV status, as many individuals overestimate the probability that they are infected. If this is the case, as ARV access encourages testing more people will believe that they are HIV negative. Following Gong (2012), this would predict a reduction in risk-taking as those who believe they are likely to be positive and find they are negative were demonstrated to respond by taking fewer risks, and so this cannot be driving the results. On the other hand, if this also changes beliefs about the prevalence of HIV in the general population, this becomes more complicated. Without detailed information about beliefs, this is beyond the scope of this paper.[26]

The threshold of 8 kilometers was chosen because it is near the median in order to maximize power, but - like any other distance cutoff - it is somewhat arbitrary. Panel A of Table A2 allows the distance threshold to vary from 8-12 kilometers. Each column repeats the analysis of the first column of Panel A of Table 3 with pregnancy as the outcome using a different distance cut-off. The results are remarkably consistent across these specifications. The sample sizes varies somewhat because of the restriction that matched pairs have the same treatment status. Some

[26] For more information about changes in beliefs as a result of ARV access, see Baranov et al (2012).

misclassification is inevitable as any distance cut-off will necessarily put some individuals who know about treatment outside of the circle while including others who do not know about it within it. However, this demonstrates that the particular choice of the threshold does not determine the estimated results.

The age cut-off can also be varied to show that there are consistent results using alternative age thresholds. While the main cut-off restricts the analysis to teenagers, a demographic that is of particular interest in research on changes in fertility behavior, others are possible. For example, the majority of those aged 21 and under do not have children, while those above are more likely than not to have had a child. The majority of those 22 and under do not report that they are cohabiting and the majority of those 23 and under do not report that they are married. Panel B of Table A2 repeats the analysis from Column 3 of Table 3 varying the age cutoff from 19 to 24, and Panel C of Table A2 repeats the analysis from Column 3 of Panel A of Table 3 using the administrative area to determine treatment status. In both tables, the results are reasonably consistent, although the estimated treatment effect declines as the threshold increases. The increase in age increases the proportion of the sample that is already married, cohabiting, or otherwise in a stable partnership, and thus unlikely to respond to changes in risk of unprotected sex, and this is likely to generate the decline in the estimated effect.

6 Simulation

The introduction of antiretroviral drugs could influence the spread of HIV both through changing behavior and through biological channels - reducing infectiousness of those on treatment and keeping more people who are HIV positive alive. This is formalized in Section 2, demonstrating how the sign of the impact of ARVs on new infections is ambiguous and depends on behavior.

The empirical analysis above showed a relative increase in risk-taking among those with access to antiretroviral treatment. This can directly increase the rate of new infections by increasing those who put themselves at risk. However, it also can indirectly decrease the rate of new infections as the increase in A_1 means that a larger fraction of the pool of potential sexual partners is HIV negative, decreasing the risk of infection for those who engage, p. This is formally demonstrated by Kremer (1996).

In addition, the reduction in transmission risk from treatment, q, can outweigh a substantial change in behavior among those who are negative so that the rate of new infections will decline with treatment. It bears mentioning that beyond the impact on new infections, ARV access has large and important welfare impacts for those who are infected and receive treatment.

In practice, the effect of D (the level of ARV provision) on behavior is likely to be non-linear with substantially larger effects on behavior when the marginal person put on treatment is sicker. The benefit to an individual who is HIV positive of being on treatment is high when he or she has a low CD4 count, which means being close to AIDS onset and opportunistic infections. However, especially given the toxicity and unpleasant side-effects, earlier treatment is not likely to provide a significant additional benefit to the individual. Thus while access to treatment provided to individuals with a CD4 count below 200 (which was previously the WHO recommended threshold) can generate the observed difference in behavior, the behavioral response is not likely to grow as the CD4 count threshold increases. However, the change in this threshold will change the probability of infection as more infected individuals are put on treatment and present a lower transmission probability.[27]

Based on the reasoning above, a low level of ARV access could change behavior but not lead to a significant reduction in infectiousness, while a very high level in which treatment is available upon diagnosis of HIV infection would reduce incidence of HIV. This is outlined in Over et al (2006) and Granich et al (2009) who propose beginning treatment immediately after a positive HIV test.

This will be demonstrated via simulation. Recall

$$I = A_1 * p$$

where I is the rate of new infections, A_1 is the fraction of the negative population that has unprotected sex, and p is the likelihood of transmission conditional on unprotected sex.

[27] WHO changed the recommended CD4 count threshold to determine ARV eligibility from 200 to 350, however most countries in Sub-Saharan Africa have not reached full coverage even with the lower threshold due to a lack of supplies. Rwanda is one exception, reporting nearly 100 percent coverage of those eligible, and experimenting with using 500 as a threshold for those in sero-discordant couples to reduce the likelihood of transmission to the uninfected partner.

This probability can be written as:

$$p = r * \frac{A_2N_2 + A_3N_3q}{A_1N_1 + A_2N_2 + A_3N_3}$$

where N_j is the size of group j, A_j is the proportion of each group that has unprotected sex, r is the transmission risk from sex with a Type 2 individual, and $r * q$ is the transmission risk from sex with a Type 3 individual. The simulation will use available estimates of each of these parameters to estimate the impact of drugs on new infections. For clarification, treatment changes A_1, N_2, and N_3. The assumptions used in the simulation are summarized in Table 5.

As described above, treatment changes behavior most at the low end, but would not be expected to change dramatically as access is available to anyone with a sufficiently low CD4-count, while the impact on transmission rates continues as treatment is provided to those based on higher CD4 thresholds. Based on Williams et al (2006), if the CD4 count threshold is set at 200, then 17% of those who are HIV positive will receive treatment. This number climbs to 44% if the threshold is 350 and 67% if the threshold is 500. For simplicity, I assume that below 17%, treatment is given to a fraction of those who need it and behavior changes for this fraction of the negative population. Above this threshold, behavior change is constant, at the level estimated in the empirical analysis. This assumed relationship between the fraction positive on treatment and the fraction negative who have sex is demonstrated in Figure 5.

I simulate new infection rates at all levels of drug provision up to 67%. This is done using 10,000 individuals. First, HIV status is assigned, then some are assigned to treatment based on the level of distribution. Death rates determine survival, and some choose to have unprotected sex. Of those who choose to, they are matched randomly. Some become infected. This is repeated for each percentage on treatment from 0-67% 500 times with and without behavior change, and with q equal to 1, 0.5 and 0.04.

Figures 6 and 7 present the estimated infection rates. Figure 6a assumes that there is no behavioral response and no reduction in transmission, and clearly, there is nearly no difference in new infection rates, except for a moderate increase explained by keeping more people who are infected alive. Figure 6b also presents estimates with no reduction in transmission, but with a change in behavior. This presents a much larger increase in infection rates. Figure 7a presents infection rates

for different levels of treatment distribution if the reduction in transmission probability from ARVs is substantial ($q = 0.04$). Here, there is a slight jump in infection rates when behavior changes (at the CD4 count threshold of 200), but there is a substantial decline in infection rates that outweighs this. Figure 7b uses $q = 0.5$ to show the impact of ARV provision if the reduction in transmission is more modest. In this case, the increase in infection due to behavior change is outweighed only if a sufficient fraction of the population is put on treatment. This suggests that provision of treatment can decrease infection rates, but that overcoming a behavioral response depends on reaching a sufficient threshold.

7 Conclusion

Previous models of the impacts of ARVs insufficiently acknowledged the importance of behavior change in shaping HIV incidence. With the absence of evidence about the magnitude or sign of this behavioral response, even the direction of the response could only be guessed. However, taking this response seriously is necessary for credibly evaluating drug provision to inform developing country governments and international donors as they weigh competing demands on tight budgets. This paper fills two prominent holes in the existing literature on HIV treatment provision in Sub-Saharan Africa: First, it provides the first causally identified estimates of the change in risky behavior due to treatment access in the context of a generalized epidemic. Second, it shows how these estimates work with existing medical evidence about the mechanical effects of ARVs to determine the predicted impacts of treatment provision on new HIV infections.

Using an original dataset that combines administrative records of the roll-out of treatment facilities in Kenya with two national population surveys, I estimate a substantial increase in risk-taking in response to treatment access. Among young women, this demonstrates an increase in pregnancies of 82% and an increase in self-reported sexual behavior of 40%. Identifying this response is crucial to estimating the impact of ARVs on the course of the HIV epidemic. Incorporating the behavioral response into a simulated model of the impact of different levels of ARV provision demonstrates that treatment provision can reduce new infection rates, even with the substantial increase in risk-taking estimated in the empirical section of the paper.

Like any study with data from a single country, the question of generalizability remains. Future

work will apply the same method of analysis in Uganda and Rwanda, combining administrative records of ARV distribution, which I have already collected, with recently released DHS data from these countries. While these results diverge from previous studies in Sub-Saharan Africa that do not find significant changes in risk-taking in response to information about HIV risk (e.g.: Godlonton et al, 2011; Oster, 2009; Wilson et al, 2011), they match evidence of behavioral responses to ARV provision among gay men in the US (Mechoulan, 2007; Papageorge, 2012). While previous changes in the risk environment were generated by variation in the likelihood of infection, ARVs change the costs of infection. As the likelihood of infection from a single encounter is low, perhaps the changes in probabilities are not easily understood or perceived, whereas a change in life expectancy and the cost of infection is more salient.

While this paper provides some evidence of the extent to which risky sexual behavior responds to changes in the cost of HIV infection, more work remains to be done to assess the generalizability of these results and variation in responses among different populations. Hopefully future assessments of proposed policy changes regarding HIV treatment provision will acknowledge the potential strength and importance of behavioral responses.

References

Auvert, Bertran, Dirk Taljaard, Emmanuel Lagarde, Joelle Sobngwi-Tambekou, Remi Sitta, and Adrian Puren (2005), "Randomized, controlled intervention trial of male circumcision for reduction of HIV infection risk: The ANRS 1265 trial," *PloS Medicine*, 2(11).

Bailey, Robert, Stephen Moses, Corette Parker, Kawango Agot, Ian Maclean, John Krieger, Carolyn Williams, Richard Campbell, Jeckoniah Ndinya-Achola (2007), "Male Circumcision for HIV prevention in young men in Kisumu Kenya: a randomised controlled trial" *The Lancet* 369(9562).

Baranov, Victoria, Daniel Bennett, and Hans-Peter Kohler (2012), "The Indirect Impact of Antiretroviral Therapy," Unpublished working paper.

Baranov, Victoria, and Hans-Peter Kohler (2012), "AIDS Treatment on Savings and Human Capital Investment in Malawi," Unpublished working paper.

Beegle, Kathleen, Joachim De Weerdt, and Stefan Dercon (2008), "Adult Mortality and Consumption Growth in the Age of HIV/AIDS," *Economic Development and Cultural Change*, 56.

Bell, Glive, Shantayanan Devarajan, and Hans Gersbach (2004), "Thinking About the Long-Run Economic Costs of AIDS," in *The Macroeconomics of HIV/AIDS*, Markus Haacker ed. Washington DC: International Monetary Fund.

Blower, Sally, Hayley Gershengorn, and Robert Grant (2000), "A Tale of Two Futures: HIV and Antiretroviral Therapy in San Francisco," *Science*, 287.

Bor, Jacob (2012), "HIV Treatment and Labor Supply in Rural South Africa," Unpublished working paper.

Cameron, Colin, Jonah Gelbach, and Douglas Miller (2011) "Robust Inference with Multi-way Clustering," *Journal of Business and Economic Statistics,* 29(2).

Cohen, Myron, Ying Chen, and Marybeth McCauley et al (2011), "Prevention of HIV-1 Infection With Early Antiretroviral Therapy," *New England Journal of Medicine*, 365.

Corrigan, Paul, Gerhard Glomm, and Fabio Mendez (2005), "AIDS Crisis and Growth," *Journal of Development Economics*, 77(1).

Cuddington, John, and John Hancock (1994), "Assessing the Impact of AIDS on the Growth Path of the Malawian Economy." *Journal of Development Economics*, 43(2).

de Walque, Damien, Harounan Kazianga, and Mead Over (2012), "Antiretroviral Therapy Awareness and Risky Sexual behaviors: Evidence from Mozambique," *Economic Development and Cultural Change*, forthcoming.

Dupas, Pascaline (2011), "Do Teenagers Respond to HIV Risk Information? Evidence from a Field Experiment in Kenya," *American Economic Journal: Applied Economics*, 3(1).

Duflo, Esther, Pascaline Dupas, and Michael Kremer (2011), "Education, HIV and Early Fertility: Experimental Evidence From Kenya," Unpublished working paper.

Ferreira, Pedro, and Samuel Pessoa (2003), "The Long-Run Economic Impact of AIDS," Unpublished working paper.

Fortson, Jane (2009), "HIV/AIDS and fertility." American Economic Journal: Applied Economics, 1(3): 170-194.

Fortson, Jane (2011), "Mortality Risk and Human Capital Investment: The Impact of HIV/AIDS in Sub-Saharan Africa," *Review of Economics and Statitstics*, 93(1).

Gertler, Paul, Manisha Shah, and Stefano M Bertozzi (2005), "Risky Business: The Market for Unprotected Commercial Sex," *Journal of Political Economy*, 113(3).

Godlonton, Susan, Alister Munthali, and Rebecca Thornton (2011), "Behavioral Response Information? Circumcision, Information, and HIV Prevention," unpublished working paper.

Gong, Erick (2012), "HIV Testing and Risky Sexual Behavior", Unpublished working paper.

Granich, Reuben, Charles Gilks, Christopher Dye, Kevin De Cock, and Brian Williams (2009) "Universal voluntary HIV testing with immediate antiretroviral therapy as a strategy for elimination of HIV transmission: a mathematical model" *The Lancet* 373(1).

Gray, Ronald, Godfrey Kigozi, David Serwadda, Frederick Makumbi, Stephen Watya, Fred Nalugo, et al (2007), "Male circumcision for HIV prevention in men in Rakai, Uganda: a randomised trial," *The Lancet*, 369(9562).

Hunter, Susan and John Williamson (2002), "Children on the Brink: Strategies to Support Children Isolated by HIV/AIDS," USAID.

Jamison, Julian and Dean Karlan, "Measuring Preferences and Predicting Outcomes," Unpublished working paper.

Juhn, Chinhui, Sebnem Kalemli-Ozcan, and Belgi Turan (2010), "HIV and Fertility in Africa: First Evidence from Population Based Surveys" Unpublished working paper.

Kalemli-Ozcan, Sebnem and Belgi Turan (2010), "HIV and Fertility Revisited" NBER Working Paper 16115.

Kaler, Amy (2003), "'My Girlfriends Could Fill a Yanu-Yanu Bus': Rural Malawian Men's Claims About Their Own Serostatus," *Demographic Research* 1(11).

Kerwin, Jason (2012), "'Rational Fatalism': Non-monotonic choices in response to risk," unpublished working paper.

Kopczuk, Wojciech, and Joel Slemrod (2005), "Denial of Death and Economic Behavior," *Advances in Theoretical Economics*, 5(1).

Kremer, Michael (1996), "Integrating Behavioral Choice into Epidemiological Models of AIDS" *The Quarterly Journal of Economics*

Lakdawalla, Darius, Neeraj Sood, and Dana Goldman (2006) "HIV Breakthroughs and Risky Sexual Behavior" *The Quarterly Journal of Economics.*

Law, Matthew, Garrett Prestage, Andrew Grulich, Paul Van de Ven, and Susan Kippax (2001), *AIDS* 15(10).

Lorentzen, Peter, John McMillan, and Romain Wacziarg (2008), "Death and development," *Journal of Economic Growth*, 13.

Lucas, Adrienne, and Nicholas Wilson (2013), "Adult Antiretroviral Therapy and Child Health: Evidence from Scale-up in Zambia," *American Economic Review: Papers and Proceedings.*

Magruder, Jeremy(2011), "Marital Shopping and Epidemic AIDS," *Demography* 48(4).

McLaren, Zoe (2011), "The Effect of Access to AIDS Treatment on Employment Outcomes in South Africa," unpublished working paper.

Mechoulan, Stephane (2007), "Risky Sexual behavior, Testing, and HIV Treatments," *Forum for Health Economics & Policy*, 10(2).

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

Oster, Emily (2009) "HIV and Sexual Behavior Change: Why not Africa?" *NBER working paper.*

Over, Mead; Elliot Marseille Kurapati Sudhakar, Julian Gold, Indrani Gupta, Abhaya Indrayan, Subhash Hira, Nico Nagelkerke, Arni Srinivasa Rao, Peter Heywood (2006), "Antiretroviral Therapy and HIV Prevention in India: Modeling Costs and Consequences of Policy Options," *Sexually Transmitted Diseases*, 33(10).

Papageorge, Nick (2012), "HIV and Incentives for Monogamy among Gay Men" unpublished working paper.

Peltzman, Sam (1975) "The Effects of Automobile Safety Regulation," *Journal of Political Economy,* 83(4).

Philipson, T. and RA Posner (1995), "A Theoretical and Empirical Investigation of the Effects of Public Health Subsides for STD Testing," *The Quarterly Journal of Economics* 110.

Rao, Vijayendra, Indrani Gupta, Michael Lokshin, and Smarajit Jana (2003), "Sex Workers and the Cost of Safe Sex: The Compensating Differential for Condom Use Among Calcutta Prostitutes," *Journal of Development Economics*, 71(2).

Robinson, Jonathan and Ethan Yeh (2011), "Transactional Sex as a Response to Risk in Western Kenya," *American Economic Journal: Applied Economics*, 3.

M Roura, M Urassa, J Busza, D Mbata, A Wringe, B Zaba: "Scaling up stigma? The effects of antiretroviral roll-out on stigma and HIV testing. Early evidence from rural Tanzania," *:* Sexually Transmitted Infections, 2009.

Thirumurthy, Harsha, Josha Graff Zivin, and Markus Goldstein (2008) "The Economic Impact of AIDS Treatment: Labor Supply in Western Kenya," *Journal of Human Resources*, 18:511-52.

Thirumurthy, Harsha, Markus Goldstein, Joshua Graff Zivin, James Habyarimana, and Cristian Pop-Eleches (2012), "Behavioral Responses of Patients in AIDS Treatment Programs: Sexual Behavior in Kenya," Forum for Health Economics and Policy, 2012, 15 (2).

Thornton, Rebecca (2008) "The demand for, and impact of, learning HIV status" *The American Economic Review.*

Wilson, Nicholas (2010), "Antiretroviral Therapy and Demand for HIV Testing: Evidence from Zambia," Unpublished working paper.

Wilson, Nicholas (2012), "Prevention of Mother-to-Child Transmission of HIV and Reproductive Behavior in Zambia," NBER Working Paper No. 18226.

Wilson, Nicholas, Wentao Ziong, Christine Mattson (2011), "Is Sex Like Driving? Risk Compensation Associated with Randomized Male Circumcision in Kisumu, Kenya," Unpublished working paper.

Winston, Cliff, Vikram Maheshri, and Fred Mannering (2006), "An exploration of the offset hypothesis using disaggregate data: The Case of airbags and antilock breaks," *Journal of Risk and Uncertainty.*

Young, Alwyn (2005), "The Gift of the Dying: The Tragedy of AIDS and the Welfare of Future Generations," *Quarterly Journal of Economics.*

Young, Alwyn (2007), "In Sorrow to Bring Forth Children: Fertility Amidst the Plague of HIV," *Journal of Economic Growth.*

8 Figures

Figure 1: ARV distribution sites in Kenya

(a) 2004

(b) 2009

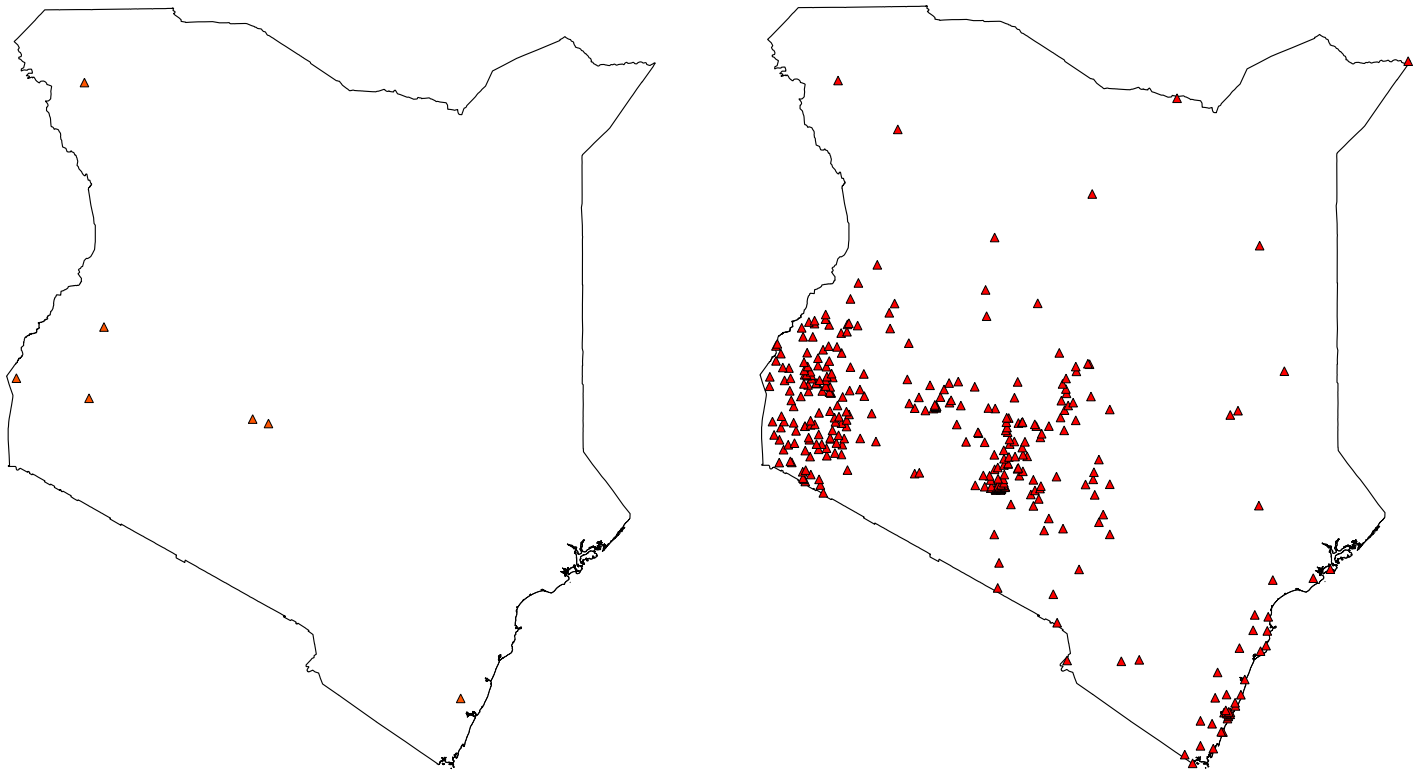


Figure 2: Treated divisions

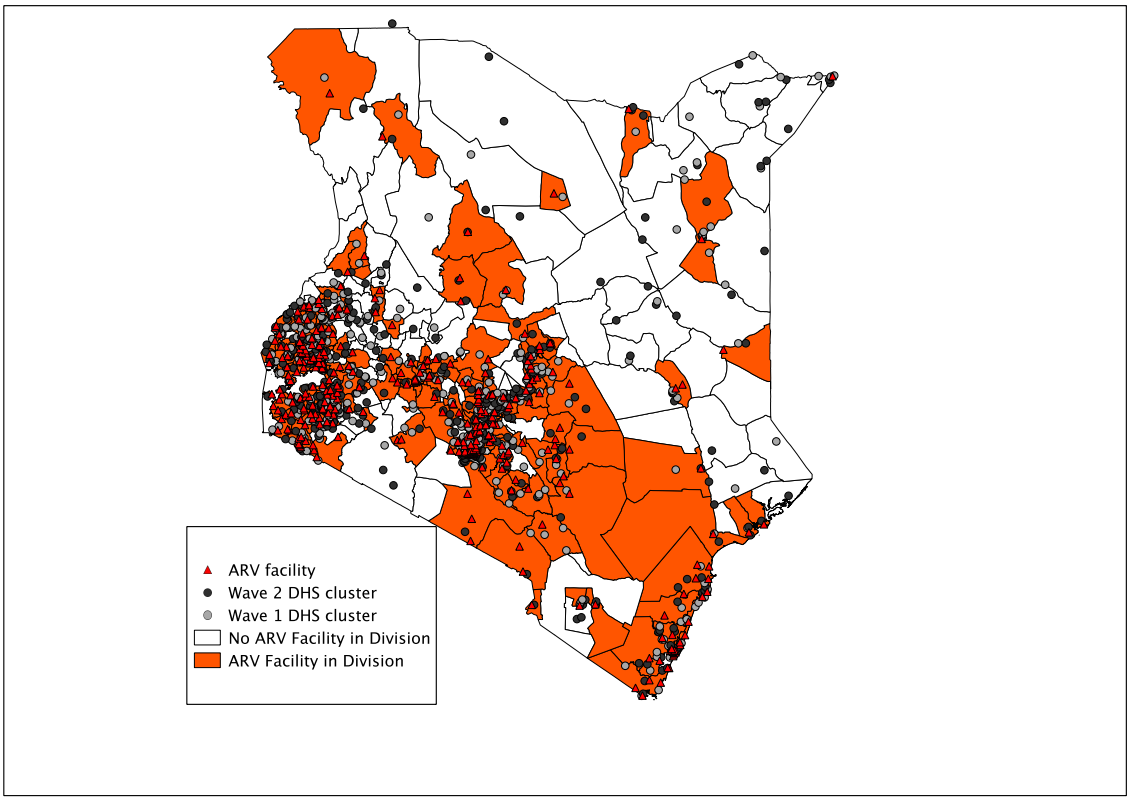


Figure 3: DHS clusters in Kenya, 2003, 2008/2009

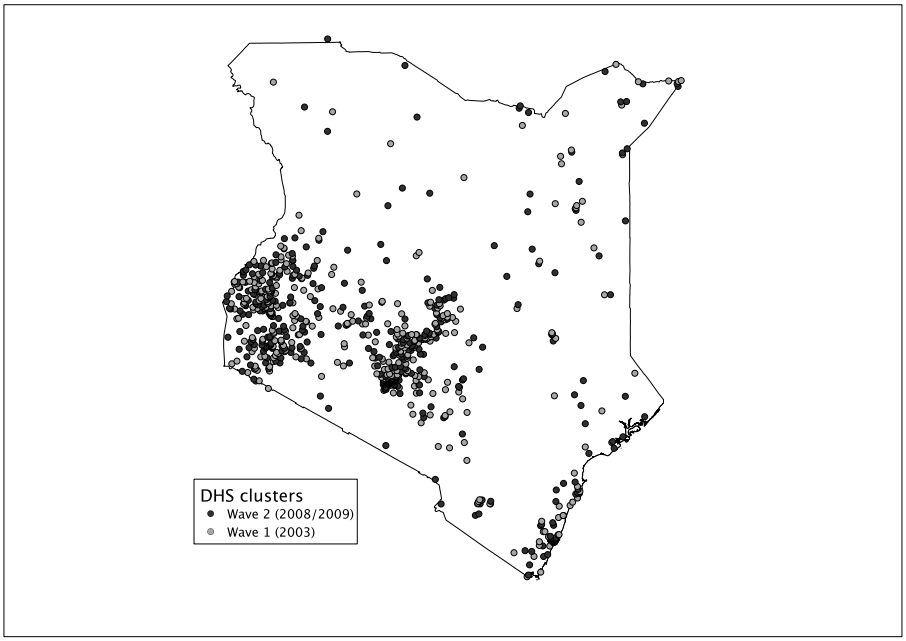


Figure 4: Parallel trends: Pregnancy rates before 2003

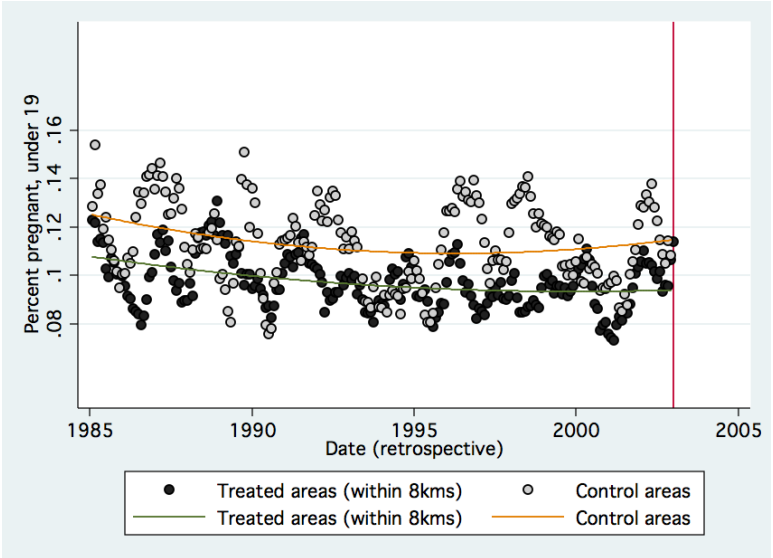


Figure 5: Simulation assumption of behavior change

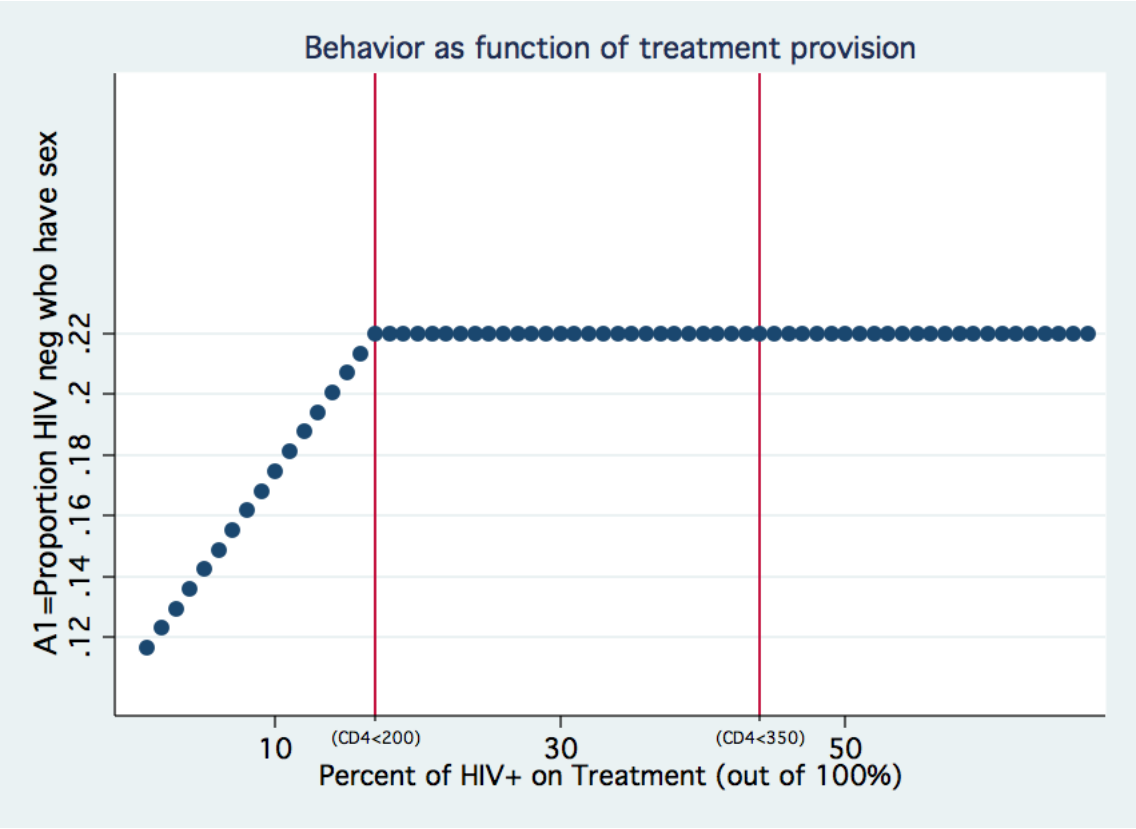
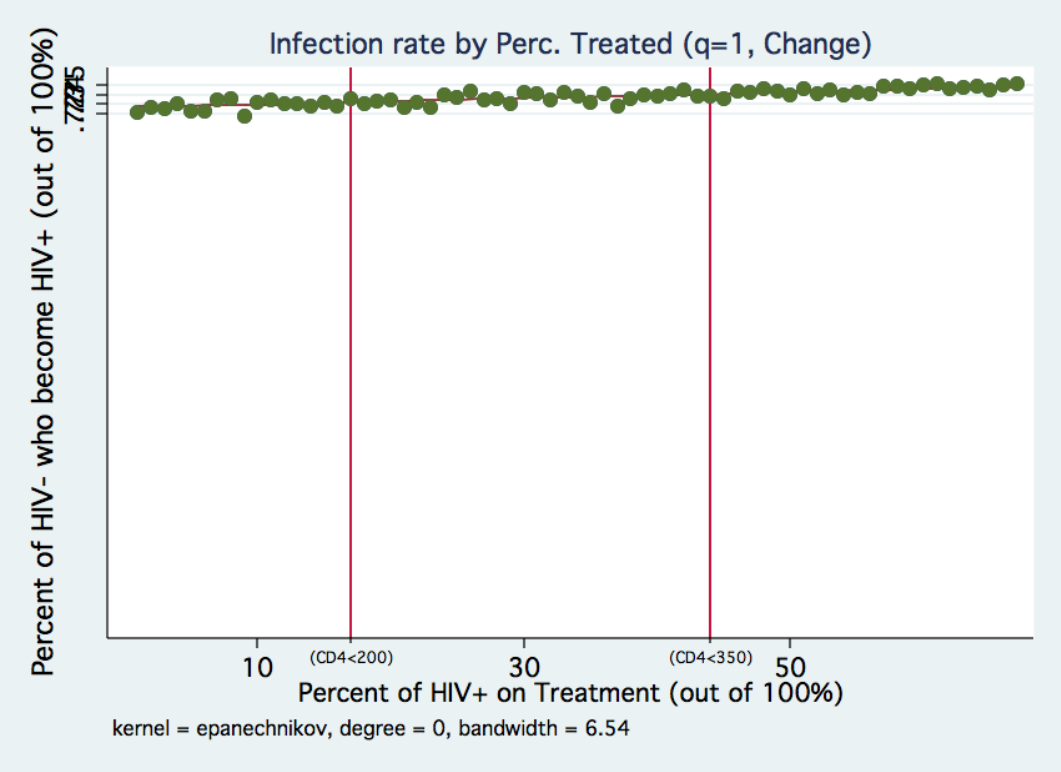


Figure 6: No reduction in transmission probability (q=1)

(a) No behavior change



(b) Behavior change

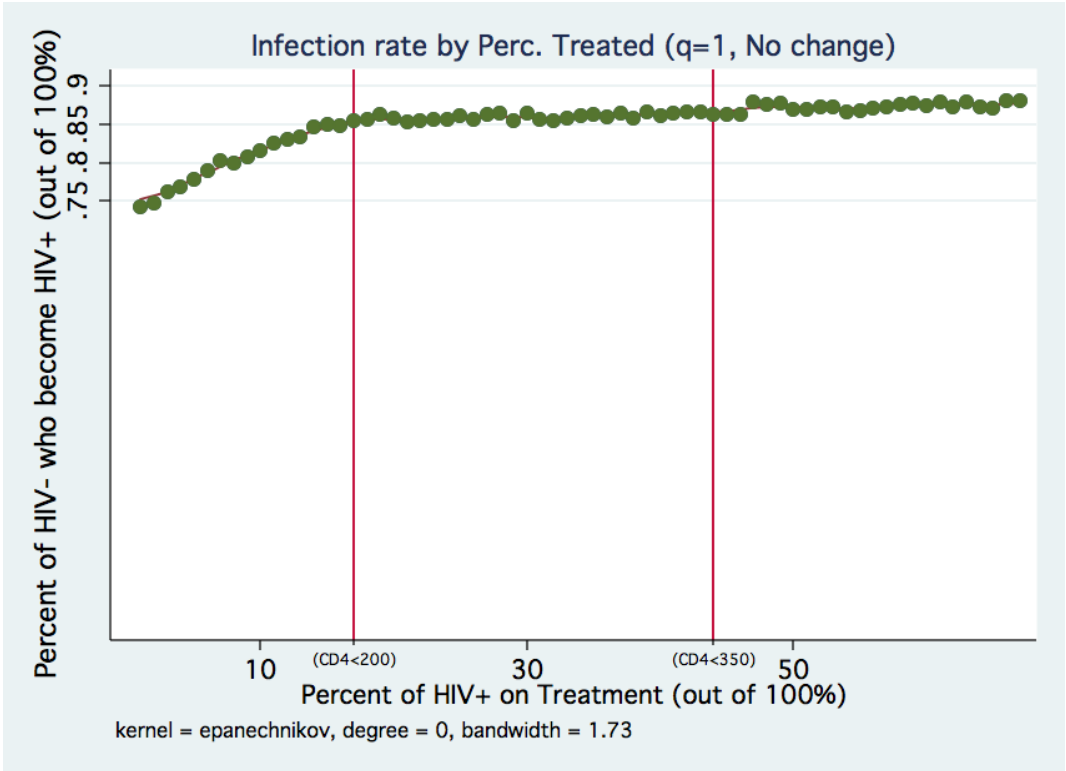
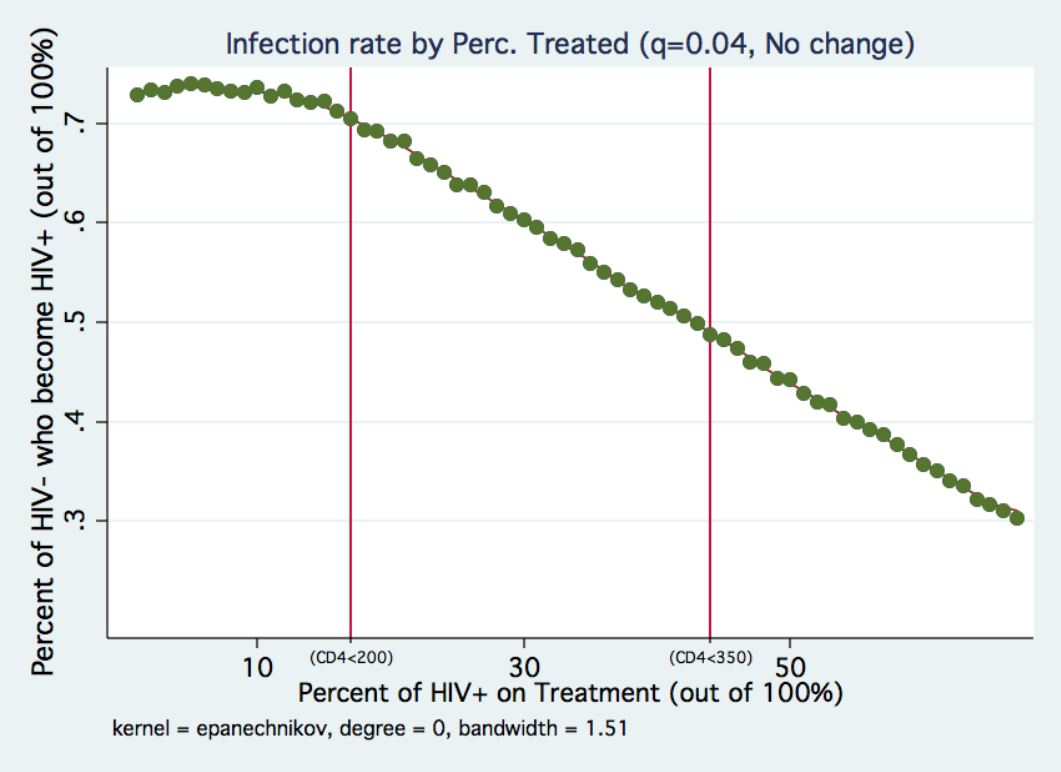
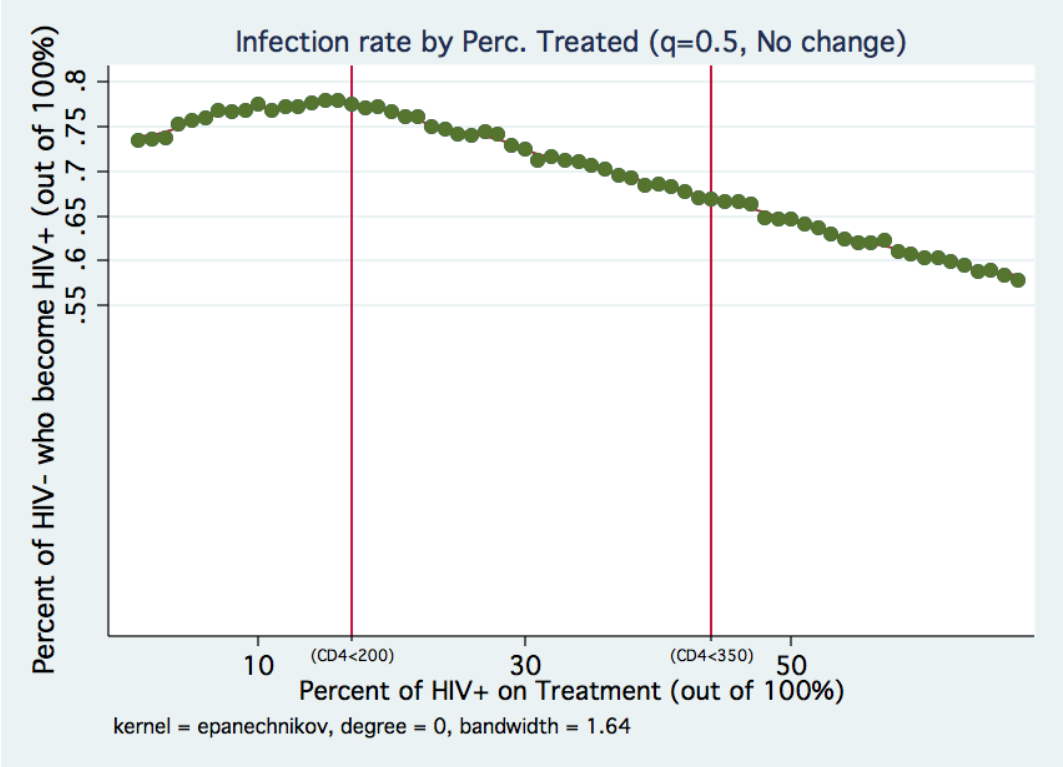


Figure 7: Behavior change and reductions in transmission probability

(a) Reduction in transmission probability of 96% (q=0.04)



(b) Reduction in transmission probability of 50% (q=0.5)



9 Tables

Table 1: Summary of ARV roll-out, HIV prevalence, and survey timing

Year	Number Facilities with ARVs	HIV Prevalence (WHO)	DHS survey Female Respondents	DHS survey Clusters
2003	1	7.5	8,195	400
2004	7	7.1		
2005	153	6.8		
2006	188	6.6		
2007	263	6.4		
2008	336	6.3	8,444	398
2009	392	6.3		
2010	610			

Note: Facilities counted as distinct only if in different locations.

Table 2: Summary Statistics 1

	2003		2008/2009	
	No ARVs in 8kms by 2008	ARVs in 8kms by 2008	No ARVs in 8kms by 2008	ARVs in 8kms by 2008
HIV positive	.014	.037	.01	.059
	(.119)	(.189)	(.099)	(.236)
Years of education	5.978	6.867	6.979	7.764
	(2.853)	(2.515)	(2.751)	(2.285)
Married	.151	.062	.062	.082
	(.358)	(.241)	(.241)	(.275)
Heard of AIDS	.966	.995	.979	.994
	(.183)	(.071)	(.145)	(.079)
Knows someone who has or died of AIDS	.656	.669	.614	.72
	(.476)	(.471)	(.487)	(.449)
Ever been tested for AIDS	.033	.054	.21	.282
	(.179)	(.225)	(.408)	(.451)
Ever had sex	.358	.351	.296	.331
	(.48)	(.477)	(.457)	(.471)
Had sex in the last 4 weeks	.144	.115	.066	.118
	(.351)	(.319)	(.249)	(.323)
Currently Pregnant/Miscarried	.083	.04	.035	.047
	(.276)	(.195)	(.185)	(.212)
Current unwanted pregnancy/miscarriage	.038	.024	.012	.035
	(.19)	(.153)	(.108)	(.184)
Ideal number of children	3.808	3.318	3.636	3.174
	(2.449)	(1.741)	(1.927)	(1.585)
Used any birth control method	.034	.055	.032	.06
	(.181)	(.228)	(.177)	(.238)
Used any birth control if had sex	.082	.141	.11	.18
	(.275)	(.349)	(.314)	(.385)
Has at least two sexual partners	.009	.021	.008	.011
	(.095)	(.144)	(.087)	(.106)
Had any STD in last 12 months	.002	.006	.003	.003
	(.047)	(.079)	(.055)	(.056)
Had STD symptoms in last 12 mos.	.009	.014	.005	.017
	(.096)	(.118)	(.074)	(.129)

Note: Standard deviations in parentheses. Includes women ages 15-18. Excludes areas with ARVs before 2004.

Table 3: Impacts of ARV access on pregnancy, self-reported sexual activity

	VARIABLES	(1) Matched Same Treatment Status	(2) Matched	(3) ART in Division
Panel A: Currently Pregnant	ARV Access*Wave2	.064 ***	.067 ***	.095 ***
		(.02)	(.019)	(.033)
		[.033]	[.033]	
	Observations	11,391	7,538	2,494
	Clusters	621	583	207
Panel B: Sex in the last 4 weeks	ARV Access*Wave2	.057 ***	.057 ***	.027
		(.024)	(.024)	(.041)
		[.042]	[.042]	
	Observations	11,391	7,538	2,494
	Clusters	621	583	207
Panel C: Unprotected sex in the last 4 weeks	ARV Access*Wave2	.048 **	.048 **	.03
		(.022)	(.022)	(.041)
		[.037]	[.037]	
	Observations	11,391	7,538	2,494
	Clusters	621	583	207

Note: All estimates include controls for age and education, district and division HIV prevalences, urban-rural status, the presence of large and small health facilities within 10kms, and each of these location characteristics interacted with Wave 2. Columns 1 and 2 include pair fixed effects with standard errors clustered at the level of the survey cluster. Two-way clustering adjusted standard errors, following Cameron et al (2006) are reported in square brackets. All estimates are weighted using DHS sampling weights. Estimates in columns 1 and 2 are additionally weighted by the DHS sampling weights multiplied by the minimum of 1/8 and the inverse of the distance between the pair. Column 3 includes division fixed effects and standard errors, clustered at the division level.

Table 4: Impacts of ARV access on fertility preferences, pregnancy in alternative subsets

		(1)	(2)	(3)	(4)
		Visited by FP worker	Ideal number of children	Used any birth control, if had sex	Used any birth control
Panel A:	ARV Access*Wave2	.007	-.207*	.017	.009
Fertility Preferences		(.011)	(.123)	(.057)	(.019)
		[.018]	[.219]	[.161]	[.033]
	Observations	7521	7538	2534	7538
	Clusters	583	583	402	583
		(1)	(2)	(3)	(4)
		Pregnant	Pregnant	Pregnant	Pregnant
Panel B:	ARV Access*Wave2	-.014	-.016	-.003	.059***
Alternate Subsets		(.013)	(.013)	(.012)	(.022)
		[.021]	[.021]	[.021]	[.04]
	Observations	23343	22376	25523	6224
	Clusters	620	620	620	561
	Subset:	Married	Married one year or more	Cohabiting	Unmarried Over25

Note: All estimates include controls for age and education, district and division HIV prevalences, urban-rural status, the presence of large and small health facilities within 10kms, each of these location characteristics interacted with Wave 2, and pair fixed effects. Standard errors in parentheses are clustered at the level of the survey cluster. Two-way clustering adjusted standard errors, following Cameron et al (2006) are reported in square brackets. All estimates are weighted using DHS sampling weights multiplied by the minimum of 1/8 and the inverse of the distance between the pair.

Table 5: Simulation Assumptions

Parameter	*Value*	*Notes*
r (transmission probability)	0.23	(representing one year)
q (reduction in transmission probability with ARVs)	0.04, 0.5, 1	0.04 represents estimates from Cohen et al (2011) 0.5 represents the lowest end of medical estimates 1 represents no reduction
d_1 (death rate among HIV negative)	0.027	Average mortality for 15-19 year-olds in Kenya between 2000 and 2005: World *Population Prospects: The 2010 Revision* UN Department of Economic and Social Affairs, Population Division (2011)
d_2 (death rate among HIV positive, untreated)	0.12	
d_3(death rate among HIV positive, treated)	0.06	
A_2 (proportion of positive untreated who have unprotected sex)	0.37	Fraction of HIV positive DHS respondents who reported having had sex in previous four weeks in untreated areas
A_3 (proportion of positive and treated who have unprotected sex)	0.33	Fraction of HIV positive DHS respondents who reported having had sex in previous four weeks in treated areas
A_1 without ARVs	0.11	Assuming: pregnancy lasts 9 months, individuals have sex twice per week, the pregnancy rate when drugs are not available is 0.6, the likelihood of becoming pregnant from unprotected sex once is 0.01: A_1 (without ARVs)= $\frac{0.06}{1-(1-0.01)^{78}} = 0.11$
A_1 with ARVs	0.11	With a pregnancy rate when drugs are available of 0.12: A_1 (with ARVs)= $\frac{0.06}{1-(1-0.01)^{78}} = 0.11$

10 Appendix

Table A1: Robustness Checks

VARIABLES	(1) ART in Division Unwanted Preg	(2) Matched Same Status Unwanted Preg	(3) ART in Division Pregnant (Untested)	(4) Matched Same Status Pregnant (Untested)
ARV Access*Wave2	0.0485**	0.0134	0.0785**	0.0589***
	(0.0245)	(0.0114)	(0.0338)	(0.0184)
Observations	2,494	7,538	2,367	7,100
R-squared	0.105	0.188	0.141	0.225
Clusters	207	583	207	579

Note: All estimates include controls for age and education, district and division HIV prevalences, urban-rural status, the presence of large and small health facilities within 10kms, and each of these location characteristics interacted with wave 2. Columns 1 and 3 define treatment as an ARV provision facility in the same division, and they include division fixed effects and standard errors, clustered at the division level. Columns 2 and 4 define treatment by distance and include include pair fixed effects with standard errors clustered at the level of the survey cluster. The dependent variable in columns 1 and 2 is current unwanted pregnancy. Columns 3 and 4 restrict the sample to those who have not been tested for HIV in the previous 12 months. All estimates are weighted using DHS sampling weights. Estimates in columns 2 and 4 are additionally weighted by the DHS sampling weights multiplied by the minimum of 1/8 and the inverse of the distance between the pair.

Table A2: Robustness: Impacts of ARV access on pregnancy, Alternative specifications

Panel A: Treatment defined as within fixed distance, Varying cutoff distance

	(1) Cutoff: 6km	(2) Cutoff: 7km	(3) Cutoff: 8km	(4) Cutoff: 9km	(5) Cutoff: 10km	(6) Cutoff: 11km
ARV Access*Wave2	.06***	.054***	.064***	.055***	.061***	.066***
	(.019)	(.02)	(.02)	(.021)	(.022)	(.025)
	[.032]	[.033]	[.033]	[.034]	[.036]	[.041]
Observations	7,422	7,442	7,538	7,947	8,050	8,242
Clusters	588	585	583	589	585	576

Panel B: Treatment defined as within 8kms, Matched specification, Varying ages

	(1) Under 18	(2) Under 19	(3) Under 20	(4) Under 21	(5) Under 22	(6) Under 23
ARV Access*Wave2	.035**	.064***	.055***	.049***	.036**	.021
	(.016)	(.02)	(.018)	(.018)	(.016)	(.015)
	[.027]	[.033]	[.029]	[.03]	[.026]	[.026]
Observations	5,653	7,538	9,313	11,266	12,744	14,464
Clusters	561	583	602	612	615	617

Panel C: Treatment defined as within same division, Varying ages

	(1) Under 18	(2) Under 19	(3) Under 20	(4) Under 21	(5) Under 22	(6) Under 23
ARV Access*Wave2	.061**	.095***	.105***	.083***	.073***	.064**
	(.03)	(.033)	(.03)	(.029)	(.028)	(.031)
Observations	1,867	2,494	3,077	3,728	4,205	4,766
Clusters	207	207	207	208	208	209

Note: In all specifications, the outcome is an indicator for whether the respondent is currently pregnant All estimates include controls for age and education, district and division HIV prevalences urban-rural status, the presence of large and small health facilities within 10kms, and each of these location characteristics interacted with Wave 2. Panels A and B include pair fixed effects with standard errors clustered at the level of the survey cluster. Two-way clustering adjusted standard errors, following Cameron et al (2006) are reported in square brackets. Estimates in panels A and B are weighted by DHS sampling weights multiplied by the minimum of 1/8 and the inverse of the distance between the pair. Estimates in panel C are weighted by DHS sampling weights and include division fixed effects and standard errors, clustered at the division level.