Journal of Development Economics 135 (2018) 392–411

Contents lists available at ScienceDirect

Journal of Development Economics

journal homepage: www.elsevier.com/locate/devec

Antiretroviral drug access and behavior change☆

Willa Helterline Friedman

University of Houston, USA

ARTICLE INFO

Jel:
I15
I18
J13
O15

ABSTRACT

Assessing the impact of antiretroviral-drug access on future HIV infections in Sub-Saharan Africa requires identification of the behavioral response. This paper combines geocoded information about the timing of introduction of ARVs in Kenyan health facilities with population surveys to estimate the impact of proximity to ARV providers on adolescent risky sexual behavior. A variety of difference-in-difference strategies yield a range of estimates of behavioral effects on pregnancy rates and self-reported sexual activity among 15-18 year-olds in areas where ARVs were introduced, from small to quite large. A simulation combining estimated behavioral responses with medical evidence regarding HIV transmission suggests increasing ARV access will reduce new HIV infections even with a very large increases in risk-taking.

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 (Cavalcanti Ferreira and Pessoa, 2003; Lorentzen et al., 2008; Fortson, 2011), reduced intergenerational human capital transmission (Beegle et al. (2008), Bell et al. (2004), Hunter and Williamson (2000)), and reduced economic growth (Cuddington and Hancock (1994), Corrigan et al. (2005)). The introduction and rapid expansion in access to Antiretroviral Drugs (ARVs), which can dramatically extend the lives of HIV-positive individuals, 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 empirically estimates the increase in unprotected sex in response to ARV access and uses this to predict the impact of ARV provision on new HIV infections. The analysis links 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 variety of difference-in-differences strategies with geographically identified survey clusters linked across rounds, I estimate a ranges of responses, ranging from small and insignificant to quite large increases in pregnancies and self-reported risky sexual behavior in the previous 4 weeks. New infections resulting from ARV access cannot yet be directly identified. To fill this gap, 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 a substantial increase in risk-taking, even with a conservative estimate of the reduction in transmission probability.

☆ I am extremely grateful to Edward Miguel, Ernesto Dal Bo, Frederico Finan, and Justin McCrary for helpful advice. The following individuals provided incredibly 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 am also indebted to Dan Bennett, Josh Blumenstock, Mariana Carrera, Aimee Chin, Damien de Walque, Lauren Falcao, Jane Fortson, Tadeja Gracner, Erick Gong, Jonas Hjort, Ashley Langer, Elaine Liu, Jeremy Magruder, Vikram Maheshri, Jamie McCasland, Zoe McLaren, 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 at UC Berkeley.

E-mail address: whfriedman@uh.edu.

https://doi.org/10.1016/j.jdeveco.2018.07.011
Received 15 November 2016; Received in revised form 17 July 2018; Accepted 21 July 2018
Available online 13 August 2018

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, 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, whereby individuals with greater expected access to ARVs would be more likely to risk HIV infection than those without. 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)), treated individuals are less likely to transmit HIV. 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.[1] 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 this paper is able to address to obtain credible estimates. The first challenge to address is to define access to ARVs. Self-reported measures of awareness of ARVs introduce endogenous variation in individual characteristics. But proximity to an ARV facility presents variation in treatment access that is independent of idiosyncratic individual characteristics. I exploit detailed geographic information to use the location of respondents relative to health facilities providing ARVs as a proxy for access to and information about treatment. Two measures of proximity are used: distance and being within the borders of the same administrative division as a health facility providing ARVs.[2]

The second challenge is to define a reasonable comparison group to serve as a counterfactual for those with access. I use a difference-in-differences identification strategy with geographic linking to deal with unobserved time-invariant differences across areas. As different villages were surveyed in each wave of the DHS, I use location to link observations across rounds. In the main specification, I link clusters of observations from each wave with those from the nearest clusters from the other wave. With multiple links, this presents a reasonable counterfactual with which to estimate the treatment effect. This will be explained in more detail in Section 3. 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 could raise 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 reasonably parallel for the years before ARVs were introduced. In the main analysis, I also restrict the sample of control areas to those where the pre-ARV pregnancy rates mirrored the levels in the treatment areas, although results will be presented for the full sample as well.

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 (2011), Minnis et al. (2009), Hewett et al. (2008)). 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, 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. (2015), Dupas (2011)). I will also report impacts on self-reported recent sexual activity, and the results are consistent.

I focus on estimating the behavioral response in pregnancy and self-reported sexual activity in areas where ARVs were introduced among women aged 15–18. This age group is used because they are the least likely to be in stable partnerships, and therefore the most likely to change their decisions about unprotected sex in response to changes in the threat of HIV infection. As a result, their childbearing - unlike that of older and married women - is more likely a reflection of HIV-related risk-taking. The data show that married women abstaining from unprotected sex is a rare category. This is discussed in more detail in Section 4. The estimated effect sizes across different outcome measures and data sources yield a wide range of estimates, which makes it difficult to conclusively pin down a precise point estimate. The highest estimates reflect a relative increase in pregnancy rates of 87%, while the lowest are close to zero. In particular, using retrospective data from 2014 provides a much lower estimate. An additional concern with the identification strategy that will be shown and discussed below is that the trend in comparison areas does not follow the trajectory from before ARV distribution began, compromising its validity in providing a measure of the counterfactual behavior. I address this by limiting the sample of comparison areas to those where the pre-distribution pregnancy levels are comparable to those in treated areas, however, this threat cannot be wiped out.

A set of additional specifications address potential alternative mechanisms that could explain the observed relationship between ARV access and fertility. A change in fertility preferences from an increase in life expectancy could generate the observed changes in pregnancy. 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. Sero-sorting - matching among individuals with the same HIV infection status - facilitated by an increase in HIV testing is another alternative mechanism. I show that the results are similar for the full sample and for a simple limited to only those who have not been tested. The size of the population that could sero-sort is also sufficiently small that this cannot drive the primary empirical results. While the outcome measures are limited, I also show an increase in self-reported recent sexual activity among men.

It is not currently possible to empirically estimate the impact of ARV provision on new HIV infections with a purely quasi-experimental approach. First, the full change in new infections will not be realized immediately. With relatively low transmission rates, any behavioral impacts on risk-taking will take some time to generate new HIV infections. 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. Distinguishing between new and old infections is infeasible in the absence of high frequency and nationally representative panel data or a test of the timing of infection.

[1] Someone can believe that the likelihood of contracting HIV has gone down without knowledge of research about ARVs reducing HIV transmission. For example, this would be the case if fewer people are visibly sick, because of treatment. This is explained in more detail in the Theoretical Framework section.

[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 two DHS clusters is 2007 km^2.

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. The medical literature has provided a range of estimates of the effect of ARVs on reducing transmission probabilities. To be conservative, the simulation combines these estimates with a high estimate of an 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 even a large increase 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)) and exercise and healthy nutrition in response to access to and advertising for cholesterol drugs (Kaestner et al., 2014; Kaplan, 2010; Mancino and Kuchler, 2009; Iizuka and Jin, 2005).

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. (2016), Wilson et al. (2014)), although at least one study has found evidence of risk compensation among women (Maughan-Brown and Venkataramani, 2012).[5] A recent study found no evidence of risk compensation with pre-exposure prophylaxis for HIV (Marcus et al., 2013). Estimates of the behavioral response to HIV risk generally have found small or no impacts on sexual behavior (Oster (2012)) or fertility (Fortson (2011), Juhn et al. (2008), Kalemli-Ozcan and Turan (2011)), although Young (2005) and Young (2007) do find a reduction in childbearing associated with HIV prevalence. Most recently, Godlonton and Thornton (2013) find an increase in risk-taking when individuals learn the test results of their acquaintances and adjust their beliefs of overall HIV risk downward.

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)). One other paper (de Walque et al., 2012) studies the impacts of beliefs about ARV effectiveness on risk-taking in sub-Saharan Africa, but the authors do not have an exogenous measure of ARV access, relying instead on self-reported beliefs about ARVs, introducing important concerns about endogeneity. Other papers have estimated the impact of ARV access on those who are HIV positive (Bor et al. (2012), Lakdawalla et al. (2006), Thirumurthy et al. (2008), Thirumurthy et al. (2012), Nikolov (2011)), and the impacts on other outcomes including employment (McLaren (2012) and Wagner et al. (2015)), mortality risk-perceptions and productivity (Baranov et al. (2015)), human capital investments (Baranov and Kohler (2018)), child health (Lucas and Wilson (2013)), HIV testing (Wilson (2011)), and mortality (Bendavid et al. (2012)). 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 how ARV provision will change the course of the epidemic.

This paper proceeds as follows. In Section 1, 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 (2015) 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)) or increase risk behavior when their beliefs about their own probability of HIV infection go down (Paula et al., 2014). In a more recent model of the HIV epidemic, Greenwood et al. (2013) demonstrate that the consideration of behavioral responses can change not only the magnitude of the effectiveness of interventions to reduce HIV spread, but the direction as well.

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, most of whom know that they are or recently were definitely not infected with HIV because they have

[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)).

[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)).

[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 especially plausible if proximity brings with it information about the existence of ARVs. I discuss the empirical implications of this assumption in section 2.

not yet had sex. 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. Throughout this section, the impact of weakening the assumption of full information about own status will be directly addressed.

- Individuals do not observe the status of any particular potential partner, but they do know the distribution of other types among potential sexual partners.[7]

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:

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

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.[8] 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:

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

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.[9,10] Altruism, morbidity, fatalism, desire for children, or any other channel through 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

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

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:

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

This assumes that Types 2 and 3 (those who are HIV positive) will not change their behavior in response to treatment access *of others*.[11]

This model also assumes that individuals know the distribution of types, yet there is evidence that in many contexts, individuals overestimate the fraction of the population infected with HIV (eg: Kerwin (2012); Kaler (2003)). In this model, that would mean that the *perceived* fractions of the population would be different, so $\widehat{A_1}$ would be smaller, and the sum of $\widehat{A_2}$ and $\widehat{A_3}$ would be larger. There is nothing in the model that relies on the assumption that $A_1 = \widehat{A_1}$, but they are not distinguished to simplify the notation. While this would change the magnitudes of the thresholds, it will not change the direction of any of the mechanisms through which the presence of ARVs change behavior in the model.

Another concern with this formulation is that many do not know their status.[12] 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.

[7] Weakening the second part of this assumption is discussed below.

[8] 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), Robinson and Yeh (2011), and Shah (2013)).

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

[10] For more thorough studies of the impact of ARVs on risk-taking behavior among those who are HIV positive, see Nikolov (2011) and Lakdawalla et al. (2006).

[11] This assumption depends on the claim that while HIV positive individuals may 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 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.

[12] 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.

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.[13] 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 $\hat{q} = 1$.[14] 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 $r \cdot q$.[15]

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

$$p = r \cdot \frac{A_2N_2 + A_3N_3q}{A_1N_1 + A_2N_2 + A_3N_3} \tag{6}$$

and by analogy, the perceived likelihood of infection is:

$$\hat{p} = r \cdot \frac{A_2N_2 + A_3N_3\hat{q}}{A_1N_1 + A_2N_2 + A_3N_3} \tag{7}$$

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 at the beginning of the current period. Besides the possibility of different behavioral parameters, γ_i and ω_i, as outlined above, individuals of Types 2 and 3 have different death rates (d_2 and d_3 respectively). The size of each population can be written as:

N_1 is fixed from the previous period.

$$N_2 = M \cdot (1 - D) \cdot (1 - d_2) \tag{8}$$

$$N_3 = M \cdot D \cdot (1 - d_3) \tag{9}$$

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 drug provision moves p moves it in the same direction through all channels. 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.[16]

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 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 had 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

[13] 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, and $\hat{q}$, which could be anywhere between 0 and 1. A belief that HIV risk has decreased after ARVs are introduced does not depend on knowledge of the medical research. Instead, if people see fewer visibly sick people, they see fewer people dying of AIDS, and they see people who were previously sick getting healthy, their perceptions about the likelihood of transmission from those who may have been sick can change.

[14] 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.

[15] Estimates of the likelihood of transmission from a single act are quite low, although other factors - including age, STI infections, circumcision status - can change it dramatically. A meta-analysis concluded that the risk of transmission from vaginal intercourse with an infected partner was 0.04% for men and 0.08% for women (Boily et al., 2009). The same estimate for low-income countries were 0.38% and 0.30%.

[16] 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 $\frac{d\hat{p}}{dD}$

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

Year	Number Facilities with ARVs	HIV Prevalence (WHO)	DHS survey	
			Female Respondents	Clusters
2003	1	7.5	8195	400
2004	7	7.1		
2005	153	6.8		
2006	188	6.6		
2007	263	6.4		
2008	336	6.3	8444	398
2009	392	6.3		
2010	610			

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

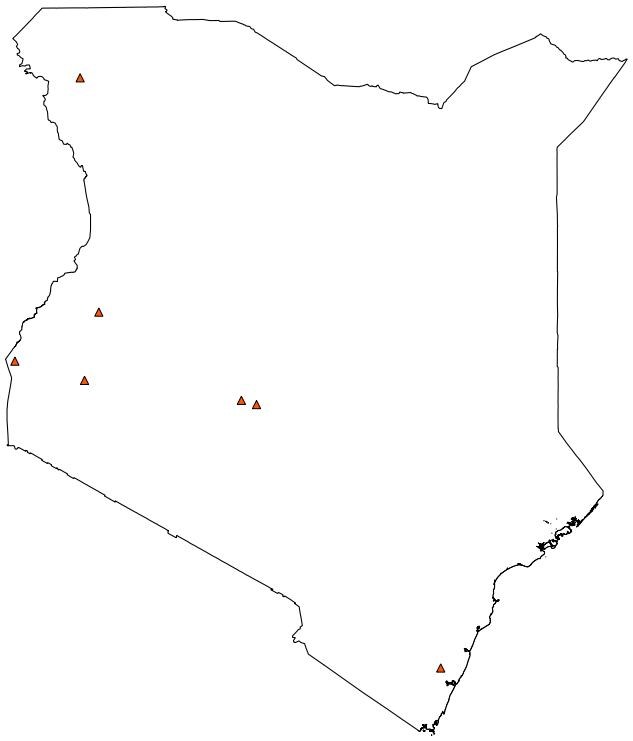

Fig. 1. ARV distribution sites in Kenya: 2004.

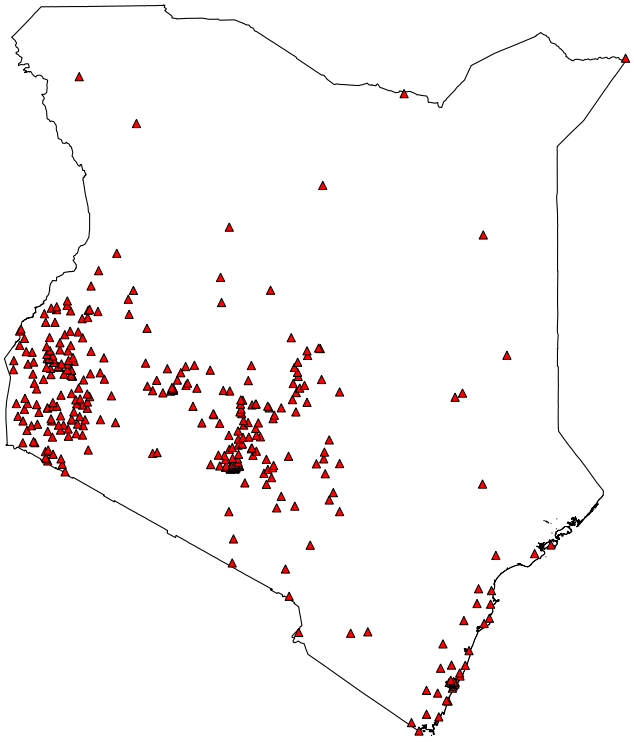

Fig. 2. ARV distribution sites in Kenya: 2008.

this increased substantially to 336 in 2008 (Figs. 1 and 2). Treatment is free for those who are HIV positive and eligible.[17]

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. 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 GPS location of all health facilities comes from the Kenya Open Data Initiative,[18] 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. The first instance in which a health facility appears in any records is used as the year in which treatment became available.[19] Table 1 shows the number of health facilities providing treatment in each year.

[17] 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 recommended threshold to 350. In Kenya, the official guidelines in 2005 stated that all patients with CD4 counts under 200 should be offered treatment. The next updated guidelines (in 2011) instructed providers to start treatment for anyone with a CD4 count below 350. In both cases, those with other risk factors (for example, more opportunistic infections and pregnant women) were to be started with higher CD4 counts.

[18] See opendata.go.ke.

[19] 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.

Table 2
Summary statistics.

	2003		2008/2009	
	No ARVs in 8 kms by 2008	ARVs in 8 kms by 2008	No ARVs in 8 kms by 2008	ARVs in 8 kms by 2008
HIV positive	.014	.036	.01	.056
	(.119)	(.187)	(.098)	(.231)
Years of education	5.922	6.863	6.94	7.731
	(2.898)	(2.515)	(2.765)	(2.296)
Married	.153	.063	.069	.082
	(.361)	(.242)	(.254)	(.275)
Heard of AIDS	.963	.995	.979	.994
	(.189)	(.071)	(.142)	(.078)
Knows someone who has or died of AIDS	.658	.673	.609	.715
	(.475)	(.47)	(.488)	(.452)
Ever been tested for AIDS	.033	.053	.213	.286
	(.179)	(.223)	(.41)	(.452)
Ever had sex	.359	.352	.304	.326
	(.48)	(.478)	(.46)	(.469)
Had sex in the last 4 weeks	.144	.117	.072	.117
	(.352)	(.322)	(.259)	(.322)
Currently Pregnant/Miscarried	.082	.043	.037	.047
	(.275)	(.202)	(.189)	(.212)
Current unwanted pregnancy/miscarriage	.037	.027	.014	.034
	(.189)	(.162)	(.119)	(.182)
Ever Pregnant	.216	.154	.122	.162
	(.412)	(.362)	(.327)	(.369)
Number of Pregnancies	.285	.182	.162	.192
	(.601)	(.461)	(.473)	(.469)
Ideal number of children	3.82	3.317	3.764	3.219
	(2.479)	(1.731)	(2.258)	(1.627)
Used any birth control method	.034	.056	.033	.059
	(.18)	(.23)	(.179)	(.235)
Used any birth control if had sex	.081	.143	.11	.178
	(.274)	(.351)	(.313)	(.383)
Has at least two sexual partners	.009	.021	.007	.011
	(.095)	(.143)	(.085)	(.105)
Had any STD in last 12 months	.002	.006	.003	.003
	(.047)	(.078)	(.054)	(.055)
Had STD symptoms in last 12 mos.	.009	.014	.01	.016
	(.095)	(.116)	(.099)	(.127)
Lives in urban area	.046	.205	.024	.254
	(.209)	(.404)	(.153)	(.436)
Within 10 km of large hospital	.034	.127	.011	.091
	(.182)	(.333)	(.103)	(.288)
No health facility within 10 km	.047	0	.092	0
	(.212)	(0)	(.289)	(0)
Division HIV prevalence	.054	.079	.047	.083
	(.079)	(.082)	(.051)	(.08)
Observations	530	716	681	642

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

For the main analysis, the data on individual behaviors for the main analysis come from two waves of geocoded Demographic and Health Surveys (DHS) from 2003 to 2008/2009,[20] 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. This will be supplemented with the recently-released 2014 DHS data for some of the analysis, for which retrospective reports can be used.

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. As they are less likely to be in stable relationships (particularly with partners whose HIV status they know), their sexual activity and childbearing are more likely to reflect decisions related to HIV-risk taking. Those who are already married are less likely to change their behavior in measurable ways as a result of HIV risks and concerns. 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. Only 3.5% of those who are married report using a condom the last time they had sex, demonstrating that abstinence from unprotected sex within marriages is rare. Premarital sex is widely discouraged, but it is quite common. Of respondents in the full sample, 58% reported that the age when they first had sex was younger than the age when they were first married. This younger group is also less likely to be already infected with HIV and thus any patterns largely reflect the behavior of those who are HIV negative. 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 percent-

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

ages 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 baseline prevalence rates, which will be addressed in the analysis. This shows that the HIV prevalence - within the sample - increased in treated areas across waves. While the goal of this paper is to look at impacts on HIV infection and it would be tempting to use HIV as an outcome, there are a number of reasons why this is not justified. First, the fraction who are HIV positive at this age is low enough that this analysis lacks sufficient power to detect an effect. Second, any estimated effect would be biased because this group has a particularly high incentive to move. Of the small number of respondents who are HIV positive in the analysis sample, 55% of those who are HIV positive report that they are visiting or have lived in their current location fewer than 5 years. In wave 2, this is true for 63% of those within 8 km of an ARV-providing facility and only 38% of those more than 8 km from a facility. The same pattern of mobility does not appear in Wave 1 or for those who are not HIV positive. Finally, it is too early to capture any full impact on HIV infections among individuals who are still taking risks through exposure. Thus even perfect data would yield an under-estimate of the effect, hence the need to look at behavioral outcomes.

Looking at other variables in Table 2, we see that 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 (2011). Among both groups, only a very small fraction report STD symptoms or multiple partnerships. Among the subset of 15–18 year-olds with sexual partners, the median age gap is 6.5 years.

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. Jamison and Karlan (2011), Minnis et al. (2009), Gersovitz et al. (1998)). 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 risky sexual behavior (e.g. Duflo et al. (2015), 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.[21] 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 ARVs and access to ARVs with the proximity to a facility providing ARVs. Proximity changes both the ease of obtaining ARVs through variation in travel costs (see for example Pinto et al. (2013)) and the likelihood of knowing about them. This change in awareness can happen through several channels, including deliberate information campaigns, posters, and billboards announcing the availability of treatment.[22] Unfortunately, the Kenya data does not include information about knowledge or use of ARVs.

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.

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.[23] Finally, each specification includes controls for urban-rural status, proximity to large - provincial and referral - hospitals to any health facilities, and each of these interacted with wave 2 to allow different trends.

The basic equation I estimate is:

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

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 km of a facility with ARVs by 2008.[24] Because the same villages were not sampled across waves, the relevant comparison group across waves is not obvious. To address this, observations are linked 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.[25] 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 linked. Any pair that is more than 100 kms apart

[21] 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.

[22] 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 15–49 who reported that they had heard of ARVs believed that they were a cure for HIV. Among those 15–18, 33% reported believing they were a cure. 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.

[23] This is constructed using both men and women in 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.

[24] 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.

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

is dropped.

Using this expanded and linked sample, I estimate another difference-in-difference estimate with linked-pair fixed effects. Each specification includes a fixed effect for each linked-cluster pair. Each observation is additionally weighted by the minimum of the inverse of the distance and 1/8.[26] Pairs of clusters 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 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 linked, 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 $Treat_j$.

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.[27] 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.[28] Therefore it is logical that the relevant proximity that would determine the spread of information about a new HIV treatment or travel costs to a facility could be within the same division. 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 linked specification is preferred.

Robustness is verified using multiple age and distance 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 nearly identical when excluding this group, as presented in Appendix Table A4. 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.

4.1. Parallel trends and limiting sample

The primary assumption to justify the difference-in-difference specification is that, conditional on the included control variables, the trends in the treatment and control areas would have been the same in the absence of treatment. Unfortunately, this is difficult to test. A standard suggestive test of this assumption relies on checking trends before treatment is introduced. Previous waves of DHS data do not include geographic information, so it isn't possible to know which respondents were in treated or comparison locations. Fortunately, the detailed birth records of women up to age 49 in the DHS data makes it is possible to check trends in pregnancy rates. Other variables of interest - sexual activity and HIV - are only measured for a single point in time.

Fig. 3 plots the percent of those who turn 18 each year who have ever been pregnant, separately for treated and control areas (defined using the 8 km distance threshold), based on the birth registry in the DHS data from 2003, 2008/2009, and 2014. The levels are clearly different, but the trends before ARV introduction appear reasonably similar. For a more formal test, I regress pregnancy on the year, treatment, and the interaction of the two, and the coefficient on the interaction is small and statistically insignificantly different from zero.

Including the data from the time when ARVs began to be introduced allows for a further visual check of the validity of the assumption that the trends in the treatment and control areas would have been similar in the absence of treatment. Even if the trends before introduction appear similar, the increase in pregnancy rates in the control areas just around the time of introduction of ARVs raises some concerns about this assumption. This raises legitimate concerns with the validity of the difference-and-difference estimates, further limiting the ability to pin down a precise magnitude of a behavioral effect, even in this context.

The difference in the levels demonstrate that the control and treatment areas are different. For a more plausible counterfactual, I therefore limit the comparison group to only control areas with pre-2006 pregnancy rates at similar levels to the treated areas. I do this by repeatedly dropping the control cluster with the highest pre-2006 average pregnancy rate until the mean in the comparison areas is within 0.001 of the mean in the treated areas. The pregnancy rates in this limited sample are presented in Fig. 4. This limited sample is used for the main analysis, although additional results with the full sample are also included in the Appendix.

4.2. Retrospective check with 2014 data

Also using retrospective birth histories, I use the 2014 round of DHS data to provide an additional estimate of the effect of ARV introduction on fertility among 15-18 year-olds using a difference-in-difference strategy. The main limitation of this strategy is that the only available information is that which is reported historically, which means that it is not possible to include the same set of control variables as in the main analysis. On the other hand, a key strength of this analysis is that

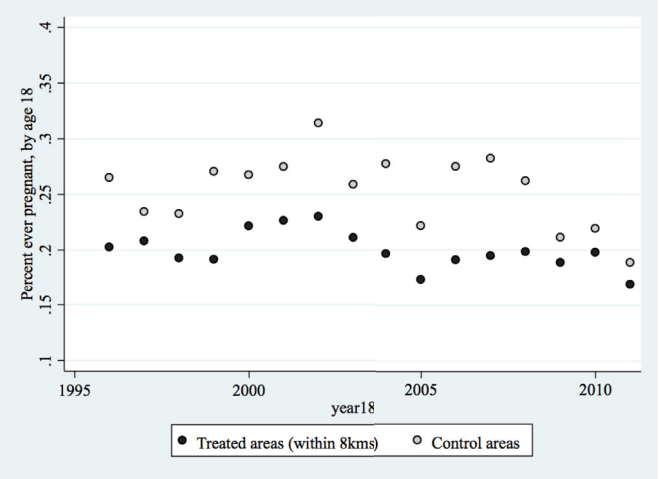

Fig. 3. Percent ever pregnant, by age 18 (2003, 2008/2009, 2014): Full sample.

[26] 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.

[27] During the time between waves, administrative boundaries have shifted. For consistency, I use current borders and place observations within them using their GPS locations. As of 2003, there were 266 divisions in Kenya. Of these, 180 contain at least one cluster from Wave 1 and 193 contain at least one cluster from Wave 2. 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.

[28] In the analysis sample, 69.5% of observations are coded the same across the two measures of treatment. Those that are within 8 kms but not in a division with ARVs represent 4.6% of the sample and those within a division with ARVs but more than 8 kms represent 25.9%. In this latter group, the average distance to a clinic is measured to be 15.7 kms with a median distance of 12.8 kms.

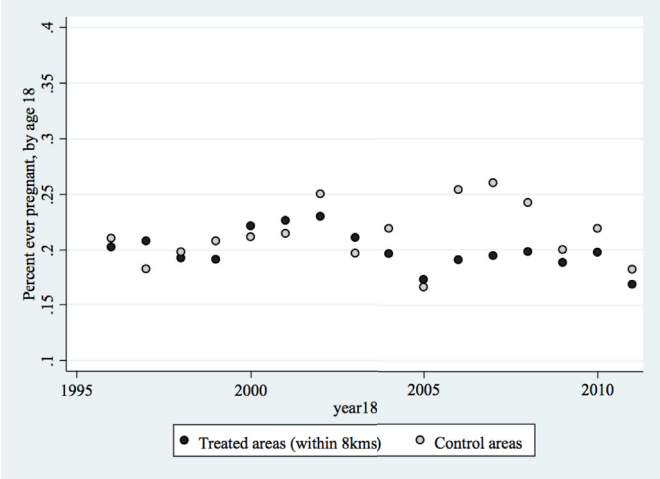

Fig. 4. Percent ever pregnant, by age 18 (2003, 2008/2009, 2014): Limited sample.

since respondents of different ages are interviewed in the same survey clusters, I can include survey cluster fixed effects to control for time-invariant location-specific factors.

Using the age when women reported having children and the age when they first had sex, I construct a retrospective dataset of women who would have been the right age to have been surveyed and included in the analysis during the earlier DHS waves. With a sample of all women who turned 18 in 2004 and 2010 - the years immediately following the earlier survey waves - I construct an indicator of whether each respondent ever gave birth before the age of 18. I use the years after the survey, for timing of birth to better match reported pregnancies in the main analysis. In an additional specification, I also include as many controls as are available to match the original analysis. Unfortunately, HIV status is not reported in the most recent DHS survey, so I am unable to control for that. As discussed above, I limit the sample to only control locations with pregnancy rates between 1995 and 2005 that are similar to the treated areas.

5. Results

The main results are reported in Panel A of Table 3. Columns 1 and 2 present the results using the specifications with linked clusters of observations. In Column 1, this estimation includes all links, 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 4.1 percentage points in Column 1 and 4.5 percentage points in Column 2. 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 9.3 percentage points. In all three specifications, the coefficient of interest is positive and statistically significantly different from zero. These point estimates are large. The standard errors are also large enough, that the true effect may be somewhat smaller.

Panels B and C of Table 3 repeat the same estimation, using other measures of fertility. Panel B presents results with the outcome of the total number of pregnancies reported. The coefficients in Columns 1 and 2 show are 14.1 and 14.5 percentage points, with 11.3 percentage points in Column 3. All are statistically significantly different from zero. The outcome in Panel C is an indicator for whether the respondent has ever been pregnant. The coefficients across the three columns are 9.4, 9.3 and 9.2 percentage points, all statistically significantly different from zero.

Panels D and E report estimates with whether the respondent reports that she has had sex in the last 4 weeks as the outcome. In Panel D, the outcome is reporting having had sex in the last 4 weeks at all, while in Panel E, the outcome is reporting having had sex in the last 4 weeks and also reporting having not used a condom with the most recent sexual partner. In all of these estimates on self-reported behavior, the point estimate of the treatment effect is measured to be between 1.8 and 3.9 percentage points, but the differences are not statistically significantly different from zero. 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 not statistically significantly different from zero.

Panel F presents results on *Ever Pregnant* as in Panel C, but using DHS data from 2014. These estimates are based on the retrospective birth histories of those reporting in 2014 whose ages would have matched those from the earlier DHS waves. In Column 1, the estimate is from a specification with no controls except survey cluster fixed effects, while controls similar to the main specification were included to estimate the coefficient reported in Column 2. The point estimate in Column 1 is 0.3 percentage points, and 3.8 percentage points in Column 2. Neither of these is statistically significantly different from zero.

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 estimation strategy 3. These results show no significant increases in the likelihood of having been visited by a family planning worker or birth control use either conditional on having had sex or unconditionally. The ideal number of children went down disproportionately in treatment areas, further suggesting that any increase in pregnancies could not be due to an increase in desired fertility.

Another way to test whether the impacts on pregnancy reflect changes in general fertility preferences is to look at who changes their behavior. If ARVs changed fertility preferences, 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. This table shows no statistically significant effects among those who are married, those who have been married by more than a year, or those who are cohabiting. 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 7.2 percentage points, more similar to that estimated for young women. These estimates further suggest that any estimated fertility differences are unlikely to reflect changes in fertility preferences.

Columns 1 and 2 of Appendix Table A2 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. 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. Additional specifications

5.2.1. HIV testing

As discussed earlier, Wilson (2011) demonstrates that demand for HIV testing is likely to increase with ARV access, because the instrumental benefit of learning ones status goes up. Similarly, Sood et al. (2015) and Wagner et al. (2014) find that access to health insurance

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

	VARIABLES	(1) Linked, Same Treatment Status	(2) Linked	(3) ART in Division
Panel A: Currently Pregnant	Treat*Wave2	.041**	.045***	.093***
		(.019)	(.019)	(.031)
		[.034]	[.034]	
	Observations	6618	9915	2318
	Clusters	532	582	203
	DepVar Mean	.047	.049	.053
Panel B: Number of Pregnancies	Treat*Wave2	.141***	.145***	.113**
		(.04)	(.039)	(.051)
		[.07]	[.07]	
	Observations	6618	9915	2318
	Clusters	532	582	203
	DepVar Mean	.204	.205	.211
Panel C: Ever Pregnant	Treat*Wave2	.094***	.093***	.092**
		(.03)	(.03)	(.043)
		[.054]	[.054]	
	Observations	6618	9915	2318
	Clusters	532	582	203
	DepVar Mean	.164	.165	.168
Panel D: Sex in the last 4 weeks	Treat*Wave2	.039	.035	.018
		(.024)	(.024)	(.043)
		[.042]	[.042]	
	Observations	6612	9901	2315
	Clusters	532	582	203
	DepVar Mean	.119	.115	.116
Panel E: Unprotected sex in the last 4 weeks	Treat*Wave2	.023	.022	.02
		(.021)	(.021)	(.042)
		[.036]	[.036]	
	Observations	6618	9915	2318
	Clusters	532	582	203
	DepVar Mean	.099	.096	.101
		(Survey Cluster FEs)	(Survey Cluster FEs & Controls)	
Panel F: Ever Pregnant 2014 sample	Treat*Wave2	.003	.038	
		(.056)	(.068)	
	Observations	3206	3206	
	Clusters	1280	1280	
	Dep Var Mean	0.271	0.271	

Note: All estimates include controls for age and education, district and division HIV prevalences, urban-rural status, the presence of large - provincial and referral - hospitals and any health facilities within 10 kms, 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. (2011) 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.

increases HIV testing in the US, particularly among high-risk groups.[29] Such an 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 A2 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.

5.2.2. *Varying cutoffs*

The threshold of 8 km 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 Appendix Table A3 allows the distance threshold to vary from 8 to 12 km. The results are reasonably consistent across these specifications.

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.[30] Panel B of Appendix Table A3 repeats the analysis from Column 3 of Table A6 varying the age cutoff from 19 to 24, and Panel C of Appendix Table A3 repeats the analysis from Column 3 of Panel A of Table A6 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

[29] Derksen et al. (2014) provides experimental evidence that information about ARVs reduces stigma and its role as a barrier to HIV testing.

[30] 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.

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

		(1) Visited by FP worker	(2) Ideal number of children	(3) Used birth control (conditional)	(4) Used any birth control
Panel A: Fertility Preferences	Treat*Wave2	.006	-.203*	.011	.001
		(.011)	(.123)	(.063)	(.021)
		[.019]	[.23]	[.274]	[.036]
	Observations	6598	6618	2193	6618
	Clusters	532	532	359	532
	DepVar Mean	.028	3.462	.163	.057
		(1) Married	(2) Married last year	(3) Cohabiting	(4) Unmarried Over 25
Panel B: Alternate Subsets	Treat*Wave2	-.018	-.021	-.008	.072***
		(.014)	(.014)	(.013)	(.024)
		[.023]	[.023]	[.022]	[.044]
	Observations	19,910	19,031	21,964	5706
	Clusters	564	564	564	524
	DepVar Mean	.113	.103	.114	.035

Note: All estimates include controls for age and education, district and division HIV prevalences, urban-rural status, the presence of large - provincial and referral - hospitals and any health and any health facilities within 10 kms, 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. (2011) 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.

respond to changes in risk of unprotected sex, and this is likely to generate the decline in the estimated effect.

5.2.3. Full sample

5.2.3.1. Dropping Nyanza and Western Kenya. Readers may be concerned that western Kenya has particular features that could confound the analysis. For example, many experiments - including some on HIV infection among young people (Duflo et al., 2015; Dupas, 2011) - have been conducted in and around Busia. In addition, the region has continued to have high rates of HIV, while also being the site of some of the earlier ARV distribution programs. To address such concerns, I repeat the main analysis removing Nyanza and Western Provinces. This is presented in Appendix Table A5. While the sample size is smaller, the magnitudes of the estimated effect sizes are similar to the main analysis, and they remain statistically significant for the three estimates of pregnancies. The estimates for self-reported sex including protected sex are again statistically insignificantly different from zero, with small and positive point estimates.

5.2.4. Men

The nature of heterosexual sex suggests that increased unprotected sex among women should also be observed among men. There are two limits to using DHS data to measure effects on unprotected sex among men. The first is that men are only surveyed in half of all households. This means that the sample of male respondents is limited. Second, pregnancies - particularly out of wedlock - that result from a man having sex are more easily hidden from the numerator and therefore underreported. This weakness means that pregnancy is not a useful proxy for unprotected sex for men as it is for women.

Still, it's possible to use this smaller sample to look at impacts on self-reported recent sexual activity. In the primary analysis, I restricted the sample to young women who were more likely to be unmarried. The relevant age for men is less clear. The median age of the partner of a 15-18 year-old woman is 25. This age gap is larger for those having unprotected sex. The median age of partners for those who had protected sex is only 23. The upper tail is also thicker for those having

Table 5
Impacts of ARV access on self-reported sexual activity among men.

	VARIABLES	(1) Linked, Same Treatment Status	(2) Linked	(3) ART in Division
Panel A: Sex in the last 4 weeks	Treat*Wave2	.06***	.048***	.008
		(.018)	(.018)	(.041)
		[.031]	[.031]	
	Observations	16,029	23,719	5526
	Clusters	565	618	205
	DepVar Mean	.479	.477	.477
Panel B: Unprotected sex in the last 4 weeks	Treat*Wave2	.036**	.022	.019
		(.017)	(.017)	(.042)
		[.028]	[.028]	
	Observations	16,058	23,755	5534
	Clusters	565	618	205
	DepVar Mean	.411	.413	.418

Note: All estimates include controls for age and education, district and division HIV prevalences, urban-rural status, the presence of large - provincial and referral - hospitals and any health facilities within 10 kms, 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. (2011) 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.

unprotected sex. Of those who used condoms, 12% of partners are age 30 or older, while 20% of partners of 15–18 year-olds who last used a condom are 30 or older. For these reasons, the analysis for men includes men of all ages.

In Table 5, I estimate the effects on men reporting that they had sex in the last four weeks (Panel A) and men reporting that they had unprotected sex in the last four weeks (Panel B). In Panel A, Columns 1 and 2 show statistically significant coefficients of 4.8–6 percentage points, while the coefficient in Column 3 is both small and statistically indistinguishable from zero. In Panel B, the coefficients range from 1.9 to 3.6 percentage points, and only the largest in Column 1 is statistically significantly different from zero.

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 1, demonstrating how the sign of the impact of ARVs on new infections is ambiguous and depends on behavior. If the behavioral effect is zero or quite small, then the reduction in transmission risk from treatment will dominate the impact of an expansion in ARV access, decreasing the number of new infections.

The empirical analysis above showed the possibility of a substantial relative increase in risk-taking among those with access to antiretroviral treatment. If the true behavioral impact is large, 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.[31]

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.

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 are summarized in Appendix Table A1.

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 Fig. 5. If the level of distribution at which behavioral effects plateaus is different or if this plateau is either steeper or more gradual, the general patterns of increasing and decreasing incidences will remain the same.

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 unpro-

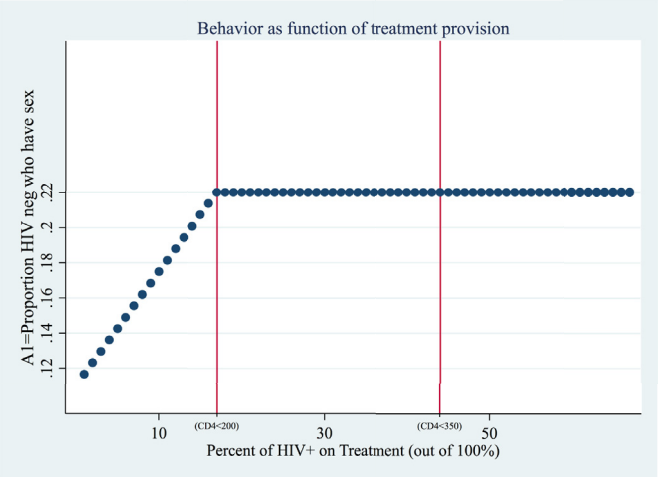

Fig. 5. Simulation assumption of behavior change.

[31] WHO changed the recommended CD4 count threshold to determine ARV eligibility from 200 to 350, however most countries in Sub-Saharan Africa had 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 serodiscordant couples to reduce the likelihood of transmission to the uninfected partner. A 2009 report from the National AIDS Control Council (NACC) and the National AIDS and STI Control Programme (NASCOP) estimated that 46% of the need for ARVs in Kenya was met in 2007.

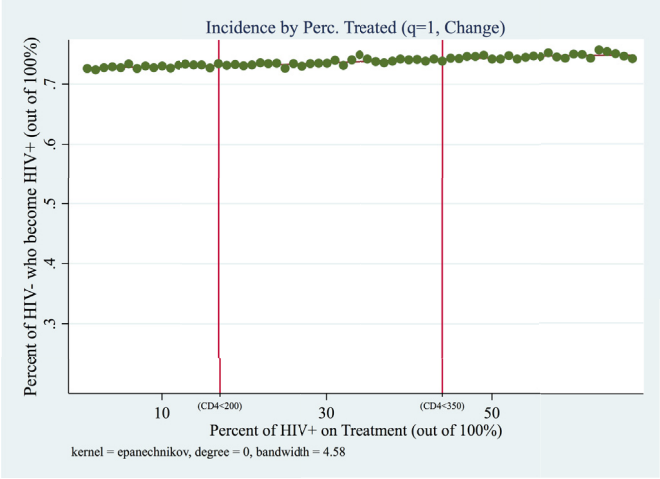

Fig. 6. No reduction in transmission probability (q = 1): No behavior change.

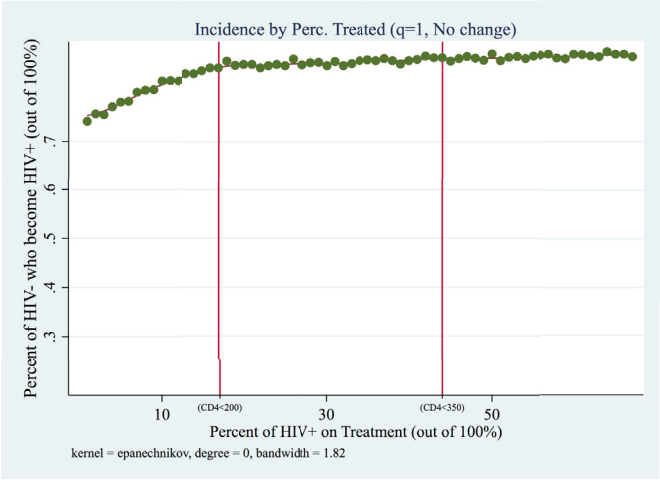

Fig. 7. No reduction in transmission probability (q = 1): With behavior change.

tected sex. Of those who choose to, they are matched randomly.[32] Some become infected. This is repeated for each percentage on treatment from 0 to 67%, 500 times, with and without behavior change, and with q equal to 1, 0.5, and 0.04. These values for q were chosen to reflect the estimates from the medical literature about ARV's reduction in HIV transmission. The most commonly cited estimate for this reduction is 96% (Cohen et al., 2011), corresponding to a q of 0.04. The lowest estimate I found for this reduction was 26% (Jia et al., 2013), although subsequent research has shown this to be an extreme outlier. While many researchers are converging on the estimate of 96%, there is still some disagreement about the true reduction in transmission, especially in real-world scenarios (see, for example Wilson (2012)). A reduction of 50% is still much lower than most estimates found in medical studies, but as will be seen in the simulation results, it could still outweigh a large behavioral response.

The estimates from the simulation suggest that even a moderate reduction in the transmission probability from ARVs can outweigh a large increase in risk-taking, predicting a decrease in new HIV infections, especially as the level of ARV provision increases. Figs. 6–9 present the estimated new infection rates (incidences). The y-axes represent the percent of new infections in the next period, so 0.7 implies 7/1000 become infected. Fig. 6 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. Fig. 7 also presents estimates with no reduction in transmission, but with a change in behavior. This presents a much larger increase in infection rates. Fig. 8 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. Fig. 9 uses

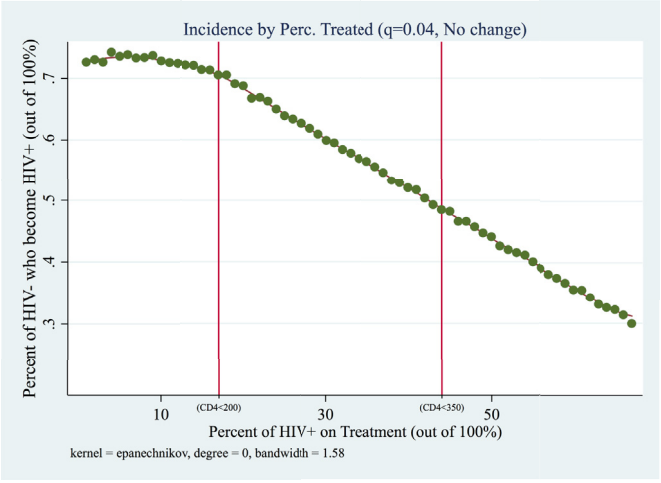

Fig. 8. Behavior change and reductions in transmission probability:Reduction in transmission probability of 96% (q = 0.04).

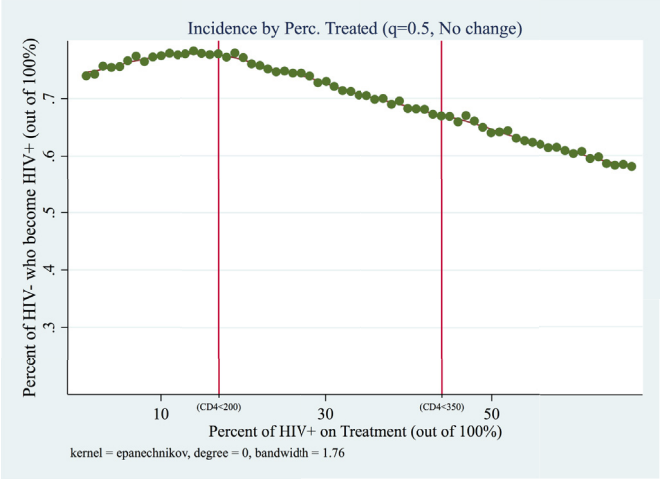

Fig. 9. Behavior change and reductions in transmission probability:Reduction in transmission probability of 50% (q = 0.5).

[32] This random matching may be unrealistic if those who are HIV positive are more likely to partner with those who are also HIV positive. Such sero-sorting would reduce the likelihood of HIV transmission by reducing the likelihood of sero-discordant partnerships. This would not change the shape of the simulated curves. A related possibility is that ARV introduction facilitates sero-sorting as more people get tested for HIV (see (Wilson, 2011) for evidence of an increase in testing). If this is the case, an increase in ARVs further reduces the HIV incidence, exacerbating the reduction from ARVs reducing HIV transmission probabilities. Therefore the possibility of sero-sorting is another factor that reduces the chance that even a large increase in risk-taking spurred by the introduction of ARVs would increase new HIV infections.

$q = 0.5$ 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.

The simulation results suggest that even a very modest decrease in the transmission rate would mean that provision of treatment can decrease new infection rates. The weaker the reduction in transmission, the more important reaching a sufficient threshold of ARV provision becomes in outweighing the behavioral response.

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. Taking this response seriously is necessary for credibly assessing drug provision as developing country governments and international donors 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 range of causally identified estimates of the change in risky behavior due to treatment access in the context of a generalized epidemic. Second, it shows how even the highest of 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 the change in risk-taking in response to treatment access. Among young women, the highest of these estimates corresponds to an increase in pregnancies of 70% and an increase in self-reported sexual behavior of 35%. However, the range of estimates across specifications and data sources suggests that this may be an overestimate of the true behavioral effects, even in this context. Incorporating the largest of the range of estimated behavioral responses 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. The smallest estimated effects are consistent with 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. (2016), Oster (2012), Wilson et al. (2014)), while large effects match evidence of behavioral responses to ARV provision among gay men in the US (Mechoulan (2007), Papageorge (2012)). Previous changes in the risk environment were generated by variation in the likelihood of infection, while 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 in different contexts remains to be done to assess the variation in responses among different populations and across different types of diseases. Beyond HIV, these results suggest that risk-taking with respect to other diseases or health hazards could depend on treatment provision changing the associated costs of the risky behavior. Future assessments of proposed policy changes regarding disease treatments should acknowledge the potential strength and importance of behavioral responses.

While the estimates of the behavioral effects leave open the possibility of a wide range of effect sizes, a key finding of this paper is that even a modest reduction in the likelihood of transmission of HIV through ARV treatment can overwhelm even with a substantial shift in risk-taking resulting from ARV distribution, implying a reduction in new HIV infections resulting from ARV distribution.

Acknowledgments

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 at UC Berkeley.

Appendix

Table A1
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) $= \frac{0.06}{1-(1-0.01)^{78}} = 0.11$

Table A2
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)
Treat* Wave2	0.0557*	0.00852	0.0720**	0.0357** (0.0177)
	(0.0304)	(0.0127)	(0.0319)	
Observations	2318	6618	2198	6202
R-squared	0.110	0.189	0.145	0.222
Clusters	203	532	203	528

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 10 kms, 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 A3
Impacts of ARV access on pregnancy, Varying Cutoffs.

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

	(1) Cutoff: 6 km	(2) Cutoff: 7 km	(3) Cutoff: 8 km	(4) Cutoff: 9 km	(5) Cutoff: 10 km	(6) Cutoff: 11 km
Treat*Wave2	.042**	.03	.041**	.033*	.035*	.038
	(.019)	(.019)	(.019)	(.02)	(.021)	(.025)
	[]	[.034]	[.034]	[.035]	[.037]	[.044]
Observations	6340	6461	6618	7018	7100	7307
Clusters	548	539	532	545	542	532
DepVar Mean	.051	.047	.047	.05	.051	.051

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

	(1) Under 18	(2) Under 19	(3) Under 20	(4) Under 21	(5) Under 22	(6) Under 23
Treat*Wave2	.025*	.041**	.039**	.037**	.022	.005
	(.015)	(.019)	(.018)	(.018)	(.016)	(.016)
	[.026]	[.034]	[.031]	[.031]	[.028]	[.026]
Observations	4980	6618	8181	9838	11,185	12,737
Clusters	509	532	549	558	560	562
DepVar Mean	.033	.047	.056	.064	.068	.076

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
Treat*Wave2	.05	.093***	.093***	.074***	.06**	.045
	(.032)	(.031)	(.03)	(.028)	(.029)	(.031)
Observations	1742	2318	2867	3458	3913	4450
Clusters	203	203	203	204	204	205
DepVar Mean	.037	.053	.062	.071	.077	.083

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 10 kms, 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. (2011) 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.

Table A4
Impacts of ARV access on pregnancy, self-reported sexual activity, exclude HIV positive.

	VARIABLES	(1) Linked, Same Treatment Status	(2) Linked	(3) ART in Division
Panel A: Currently Pregnant	Treat*Wave2	.044**	.048***	.092***
		(.019)	(.019)	(.031)
		[.034]	[.034]	
	Observations	6512	9767	2286
	Clusters	532	581	203
	DepVar Mean	.047	.048	.052

(*continued on next page*)

Table A4 *(continued)*

	VARIABLES	(1) Linked, Same Treatment Status	(2) Linked	(3) ART in Division
Panel B: Number of Pregnancies	Treat*Wave2	.109***	.109***	.112**
		(.04)	(.04)	(.051)
		[.072]	[.072]	
	Observations	6512	9767	2286
	Clusters	532	581	203
	DepVar Mean	.196	.198	.206
Panel C: Ever Pregnant	Treat*Wave2	.077***	.074***	.093**
		(.031)	(.031)	(.043)
		[.055]	[.055]	
	Observations	6512	9767	2286
	Clusters	532	581	203
	DepVar Mean	.159	.16	.164
Panel D: Sex in the last 4 weeks	Treat*Wave2	.039	.035	.017
		(.024)	(.024)	(.043)
		[.043]	[.043]	
	Observations	6506	9753	2283
	Clusters	532	581	203
	DepVar Mean	.115	.112	.113
Panel E: Unprotected sex in the last 4 weeks	Treat*Wave2	.021	.021	.018
		(.022)	(.022)	(.042)
		[.037]	[.037]	
	Observations	6512	9767	2286
	Clusters	532	581	203
	DepVar Mean	.095	.092	.098

Note: All estimates include controls for age and education, district and division HIV prevalences, urban-rural status, the presence of large - provincial and referral - hospitals and any health facilities within 10 kms, 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. (2011) 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 A5
Impacts of ARV access on pregnancy, self-reported sexual activity, exclude Western and Nyanza Provinces.

	VARIABLES	(1) Linked, Same Treatment Status	(2) Linked	(3) ART in Division
Panel A: Currently Pregnant	Treat*Wave2	.03	.047**	.106***
		(.019)	(.021)	(.038)
	Observations	3771	6019	1503
	Clusters	356	390	147
	DepVar Mean		.043	.046
Panel B: Number of Pregnancies	Treat*Wave2	.107**	.117***	.105*
		(.047)	(.046)	(.062)
	Observations	3771	6019	1503
	Clusters	356	390	147
	DepVar Mean		.168	.186
Panel C: Ever Pregnant	Treat*Wave2	.084**	.087***	.069
		(.039)	(.037)	(.051)
		[0]	[0]	
	Observations	3771	6019	1503
	Clusters	356	390	147
	DepVar Mean		.132	.147
Panel D: Sex in the last 4 weeks	Treat*Wave2	.01	.009	.032
		(.031)	(.032)	(.048)
	Observations	3767	6014	1502
	Clusters	356	390	147
	DepVar Mean		.097	.103
Panel E: Unprotected sex in the last 4 weeks	Treat*Wave2	.03	.031	.032
		(.027)	(.027)	(.048)
	Observations	3771	6019	1503
	Clusters	356	390	147
	DepVar Mean		.084	.093

Note: All estimates include controls for age and education, district and division HIV prevalences, urban-rural status, the presence of large - provincial and referral - hospitals and any health facilities within 10 kms, 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. 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 A6
Impacts of ARV access on pregnancy, self-reported sexual activity, full sample.

	VARIABLES	(1) Linked, Same Treatment Status	(2) Linked	(3) ART in Division
Panel A: Currently Pregnant	Treat*Wave2	.057***	.06***	.091***
		(.02)	(.019)	(.033)
		[.033]	[.033]	
	Observations	7797	11,820	2562
	Clusters	602	636	211
	DepVar Mean	.052	.053	.055
Panel B: Number of Pregnancies	Treat*Wave2	.183***	.187***	.129**
		(.04)	(.039)	(.062)
		[.067]	[.067]	
	Observations	7797	11,820	2562
	Clusters	602	636	211
	DepVar Mean	.219	.221	.223
Panel C: Ever Pregnant	Treat*Wave2	.118***	.118***	.116***
		(.029)	(.029)	(.049)
		[.05]	[.05]	
	Observations	7797	11,820	2562
	Clusters	602	636	211
	DepVar Mean	.174	.175	.175
Panel D: Sex in the last 4 weeks	Treat*Wave2	.05**	.045*	.023
		(.024)	(.024)	(.042)
		[.042]	[.042]	
	Observations	7791	11,806	2559
	Clusters	602	636	211
	DepVar Mean	.126	.123	.123
Panel E: Unprotected sex in the last 4 weeks	Treat*Wave2	.043*	.04*	.026
		(.022)	(.021)	(.041)
		[.037]	[.037]	
	Observations	7797	11,820	2562
	Clusters	602	636	211
	DepVar Mean	.108	.105	.108

Note: All estimates include controls for age and education, district and division HIV prevalences, urban-rural status, the presence of large - provincial and referral - hospitals and any health facilities within 10 kms, 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. (2011) are reported in square brackets. All estimates are weightedusing 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 A7
Summary Statistics, using sample with geographically linked observations.

	2003		2008/2009	
	No ARVs in 8 kms by 2008	ARVs in 8 kms by 2008	No ARVs in 8 kms by 2008	ARVs in 8 kms by 2008
HIV positive	.022	.043	.005	.06
	(.147)	(.204)	(.072)	(.238)
Years of education	6.104	6.915	6.845	7.81
	(2.904)	(2.481)	(2.805)	(2.143)
Married	.129	.066	.071	.088
	(.336)	(.248)	(.257)	(.283)
Heard of AIDS	.967	.995	.979	.998
	(.179)	(.068)	(.143)	(.039)
Knows someone who has or died of AIDS	.665	.676	.596	.709
	(.473)	(.468)	(.491)	(.455)
Ever been tested for AIDS	.03	.052	.2	.292
	(.17)	(.222)	(.401)	(.455)
Ever had sex	.346	.367	.302	.332
	(.476)	(.482)	(.46)	(.471)
Had sex in the last 4 weeks	.133	.126	.066	.117
	(.34)	(.332)	(.249)	(.322)
Currently Pregnant/Miscarried	.078	.042	.038	.053
	(.268)	(.202)	(.191)	(.224)
Current unwanted pregnancy/miscarriage	.041	.031	.011	.039
	(.2)	(.173)	(.104)	(.195)
Ideal number of children	3.575	3.273	3.869	3.166
	(2.252)	(1.545)	(2.33)	(1.604)
Used any birth control method	.031	.062	.025	.061
	(.173)	(.241)	(.156)	(.239)
Used any birth control if had sex	.078	.165	.083	.18
	(.27)	(.372)	(.276)	(.385)
Has at least two sexual partners	.008	.019	.006	.015
	(.087)	(.135)	(.074)	(.121)

(*continued on next page*)

Table A7 *(continued)*

	2003		2008/2009	
	No ARVs in 8 kms by 2008	ARVs in 8 kms by 2008	No ARVs in 8 kms by 2008	ARVs in 8 kms by 2008
Had any STD in last 12 months	.004	.007	0	.004
	(.06)	(.082)	(0)	(.063)
Had STD symptoms in last 12 mos.	.009	.012	.008	.021
	(.093)	(.111)	(.091)	(.143)
Lives in urban area	.056	.21	.019	.258
	(.231)	(.408)	(.138)	(.438)
Within 10 km of large hospital	.05	.116	.007	.121
	(.218)	(.321)	(.085)	(.326)
No health facility within 10 km	.035	0	.099	
	(.185)	(0)	(.299)	0
				(0)
Division HIV prevalence	.058	.078	.046	.095
	(.084)	(.074)	(.05)	(.082)
Observations	325	507	553	485

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

References

Auvert, Bertran, Taljaard, Dirk, Lagarde, Emmanuel, Sobngwi-Tambekou, Joelle, Sitta, Rémi, Adrian, Puren, 2005. Randomized, controlled intervention trial of male circumcision for reduction of HIV infection risk: the ANRS 1265 Trial. PLoS Med. 2 (11), e298.

Bailey, Robert C., Moses, Stephen, Parker, Corette B., Agot, Kawango, Maclean, Ian, Krieger, John N., Williams, Carolyn FM., Campbell, Richard T., Ndinya-Achola, Jeckoniah O., 2007. Male circumcision for HIV prevention in young men in Kisumu, Kenya: a randomised controlled trial. Lancet 369 (9562), 643–656.

Baranov, Victoria, Kohler, Hans-Peter, 2018. The impact of AIDS treatment on savings and human capital investment in Malawi. Am. Econ. J. Appl. Econ. 10 (1), 266–306.

Baranov, Victoria, Bennett, Daniel, Kohler, Hans-Peter, 2015. The indirect impact of antiretroviral therapy: mortality risk, mental health, and hiv-negative labor supply. J. Health Econ. 44, 195–211.

Beegle, Kathleen, De Weerdt, Joachim, Dercon, Stefan, 2008. Adult mortality and consumption growth in the age of HIV/AIDS. Econ. Dev. Cult. Change 56 (2), 299–326.

Bell, Clive, Devarajan, Shantayanan, Gersbach, Hans, 2004. Thinking about the long-run economic costs of AIDS. Macroecon. HIV/AIDS 96, 128–129.

Bendavid, Eran, Holmes, Charles B., Bhattacharya, Jay, Miller, Grant, 2012. HIV development assistance and adult mortality in Africa. JAMA 307 (19), 2060–2067.

Blower, S.M., Gershengorn, Hayley B., Grant, R.M., 2000. A tale of two futures: HIV and antiretroviral therapy in San Francisco. Science 287 (5453), 650–654.

Boily, Marie-Claude, Baggaley, Rebecca F., Wang, Lei, Masse, Benoit, White, Richard G., Hayes, Richard J., Alary, Michel, 2009. Heterosexual risk of HIV-1 infection per sexual act: systematic review and meta-analysis of observational studies. Lancet Infect. Dis. 9 (2), 118–129.

Bor, Jacob, Frank, Tanser, Marie-Louise Newell, Bärnighausen, Till, 2012. In a study of a population cohort in South Africa, HIV patients on antiretrovirals had nearly full recovery of employment. Health Aff. 31 (7), 1459–1469.

Cameron, A Colin, Jonah, B Gelbach, Douglas, L Miller, 2011. Robust inference with multiway clustering. J. Bus. Econ. Stat. 29 (2), 238–249.

Cavalcanti Ferreira, Pedro, Pessoa, Samuel, 2003. The Long-run Economic Impact of AIDS. Available at SSRN 411782.

Cohen, Myron S., Chen, Ying Q., McCauley, Marybeth, Gamble, Theresa, Hosseinipour, Mina C., Kumarasamy, Nagalingeswaran, Hakim, James G., Kumwenda, Johnstone, Grinsztejn, Beatriz, Pilotto, Jose HS., et al., 2011. Prevention of HIV-1 infection with early antiretroviral therapy. N. Engl. J. Med. 365 (6), 493–505.

Corrigan, Paul, Glomm, Gerhard, Mendez, Fabio, 2005. AIDS crisis and growth. J. Dev. Econ. 77 (1), 107–124.

Cuddington, John T., Hancock, John D., 1994. Assessing the impact of AIDS on the growth path of the Malawian economy. J. Dev. Econ. 43 (2), 363–368.

Derksen, Laura, Adamson, Muula, Joep van Oosterhout, 2014. Love in the Time of HIV: Theory and Evidence on Social Stigma and Health Seeking Behavior. Working Paper. London School of Economics.

de Walque, Damien, Kazianga, Harounan, Mead, Over, 2012. Antiretroviral therapy perceived efficacy and risky sexual behaviors: evidence from Mozambique. Econ. Dev. Cult. Change 61 (1), 97–126.

Duflo, Esther, Dupas, Pascaline, Kremer, Michael, 2015. Education, HIV, and early fertility: experimental evidence from Kenya. Am. Econ. Rev. 105 (9), 2757–2797.

Dupas, Pascaline, 2011. Do teenagers respond to HIV risk information? Evidence from a field experiment in Kenya. Am. Econ. J. Appl. Econ. 3 (1), 1–34.

Fortson, Jane G., 2011. Mortality risk and human capital investment: the Impact of HIV/AIDS in Sub-Saharan Africa. Rev. Econ. Stat. 93 (1), 1–15.

Gersovitz, Mark, Jacoby, Hanan G., Dedy, F Seri, Tapé, A Gozé, 1998. The balance of self-reported heterosexual activity in KAP surveys and the AIDS epidemic in Africa. J. Am. Stat. Assoc. 93 (443), 875–883.

Gertler, Paul, Shah, Manisha, Bertozzi, Stefano M., 2005. Risky business: the market for unprotected commercial sex. J. Polit. Econ. 113 (3), 518–550.

Godlonton, Susan, Munthali, Alister, Thornton, Rebecca, 2016. Responding to risk: circumcision, information, and HIV prevention. Rev. Econ. Stat. 98 (2), 333–349.

Godlonton, Susan, Thornton, Rebecca L., 2013. Learning from others HIV testing: updating beliefs and responding to risk. Am. Econ. Rev. 103 (3), 439.

Gong, Erick, 2015. HIV testing and risky sexual behaviour. Econ. J. 125 (582), 32–60.

Granich, Reuben M., Gilks, Charles F., Dye, Christopher, De Cock, Kevin M., Williams, Brian G., 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.

Gray, Ronald H., Godfrey, Kigozi, Serwadda, David, Makumbi, Frederick, Stephen, Watya, Nalugoda, Fred, Kiwanuka, Noah, Moulton, Lawrence H., Chaudhary, Mohammad A., Chen, Michael Z., et al., 2007. Male circumcision for HIV prevention in men in Rakai, Uganda: a randomised trial. Lancet 369 (9562), 657–666.

Greenwood, Jeremy, Kircher, Philipp, Santos, Cezar, Tertilt, Michèle, 2013. An Equilibrium Model of the African HIV/AIDS Epidemic. National Bureau of Economic Research.

Hewett, Paul C., Mensch, Barbara S., Ribeiro, Manoel Carlos S. de A., Jones, Heidi E., Lippman, Sheri A., Montgomery, Mark R., van de Wijgert, Janneke HHM., 2008. Using sexually transmitted infection biomarkers to validate reporting of sexual behavior within a randomized, experimental evaluation of interviewing methods. Am. J. Epidemiol. 168 (2), 202–211.

Hunter, Susan, Williamson, John, 2000. Children on the Brink. USAID, Washington DC.

Iizuka, Toshiaki, Jin, Ginger Zhe, 2005. Drug Advertising and Health Habit. National Bureau of Economic Research.

Jamison, Julian, Karlan, Dean, 2011. Measuring Preferences and Predicting Outcomes, pp. 1–3.

Jia, Zhongwei, Mao, Yurong, Zhang, Fujie, Ruan, Yuhua, Ma, Ye, Li, Jian, Guo, Wei, Liu, Enwu, Dou, Zhihui, Zhao, Yan, et al., 2013. Antiretroviral therapy to prevent HIV transmission in serodiscordant couples in China (2003–11): a national observational cohort study. Lancet 382 (9899), 1195–1203.

Juhn, Chinhui, Kalemli-Ozcan, Sebnem, Turan, Belgi, 2008. HIV and fertility in Africa: first evidence from population-based surveys. J. Popul. Econ. 1–19.

Kaestner, Robert, Darden, Michael, Lakdawalla, Darius, 2014. Are investments in disease prevention complements? The case of statins and health behaviors. J. Health Econ. 36, 151–163.

Kalemli-Ozcan, Sebnem, Turan, Belgi, 2011. HIV and fertility revisited. J. Dev. Econ. 96 (1), 61–65.

Kaler, Amy, 2003. My girlfriends could fill a yanu-yanu bus: rural Malawian mens claims about their own serostatus. Demogr. Res. 19.

Kaplan, Cameron, 2010. Risk Compensation and Treatment for High Cholesterol.

Kerwin, Jason T., 2012. Rational Fatalism: Non-monotonic Choices in Response to Risk.

Kremer, Michael, 1996. Integrating behavioral choice into epidemiological models of AIDS. Q. J. Econ. 111 (2), 549–573.

Lakdawalla, Darius, Sood, Neeraj, Goldman, Dana, 2006. HIV breakthroughs and risky sexual behavior. Q. J. Econ. 121 (3), 1063–1102.

Law, Matthew G., Garrett, Prestage, Grulich, Andrew, Van de Ven, Paul, Kippax, Susan, 2001. Modelling the effect of combination antiretroviral treatments on HIV incidence. AIDS 15 (10), 1287–1294.

Lorentzen, Peter, McMillan, John, Wacziarg, Romain, 2008. Death and development. J. Econ. Growth 13 (2), 81–124.

Lucas, Adrienne M., Wilson, Nicholas L., 2013. Adult antiretroviral therapy and child health: evidence from scale-up in Zambia. Am. Econ. Rev. 103 (3), 456–461.

Mancino, Lisa, Kuchler, Fred, 2009. Offsetting behavior in reducing high cholesterol: substitution of medication for diet and lifestyle changes. J. Choice Modell. 2 (1), 51–64.

Marcus, Julia L., Glidden, David V., Mayer, Kenneth H., Liu, Albert Y., Buchbinder, Susan P., Amico, K Rivet, McMahan, Vanessa, Georges Kallas, Esper, Montoya-Herrera, Orlando, Pilotto, Jose, et al., 2013. No Evidence of Sexual Risk Compensation in the IPrEx Trial of Daily Oral HIV Preexposure Prophylaxis.

Maughan-Brown, Brendan, Venkataramani, Atheendar S., 2012. Learning that circumcision is protective against HIV: risk compensation among men and women in Cape Town, South Africa. PLoS One 7 (7), e40753.

McLaren, Zoë M., 2012. The Effect of Access to AIDS Treatment on Employment Outcomes in South Africa.
Mechoulan, Stéphane, 2007. Risky Sexual Behavior, Testing, and HIV Treatments, vol. 10. De Gruyter, pp. 1–51.
Minnis, Alexandra M., Steiner, Markus J., Gallo, Maria F., Lee, Warner, Hobbs, Marcia M., van der Straten, Ariane, Chipato, Tsungai, Macaluso, Maurizio, Padian, Nancy S., 2009. Biomarker validation of reports of recent sexual activity: results of a randomized controlled study in Zimbabwe. Am. J. Epidemiol. 170 (7), 918–924.
Nikolov, Plamen, 2011. Does AIDS Treatment Stimulate Negative Behavioral Response? a Field Experiment in South Africa. Harvard University. Working Paper.
Oster, Emily, 2012. HIV and sexual behavior change: why not Africa? J. Health Econ. 31 (1), 35–49.
Over, Mead, Marseille, Elliot, Sudhakar, Kurapati, Gold, Julian, Gupta, Indrani, Indrayan, Abhaya, Hira, Subhash, Nagelkerke, Nico, Arni, S.R., Rao, Srinivasa, Heywood, Peter, 2006. Antiretroviral therapy and HIV prevention in India: modeling costs and consequences of policy options. Sex. Transm. Dis. 33 (10), S145–S152.
Papageorge, Nick W., 2012. HIV and Incentives for Monogamy Among Gay Men.
Paula, Áureo De, Shapira, Gil, Todd, Petra E., 2014. How beliefs about HIV status affect risky behaviors: evidence from Malawi. J. Appl. Econom. 29 (6), 944–964.
Peltzman, Sam, 1975. The effects of automobile safety regulation. J. Polit. Econ. 677–725.
Pinto, Andrew D., van Lettow, Monique, Rachlis, Beth, Adrienne, K Chan, Sumeet, K Sodhi, 2013. Patient costs associated with accessing HIV/AIDS care in Malawi. J. Int. AIDS Soc. 16 (1).
Rao, Vijayendra, Gupta, Indrani, Lokshin, Michael, Jana, Smarajit, 2003. Sex workers and the cost of safe sex: the compensating differential for condom use among Calcutta prostitutes. J. Dev. Econ. 71 (2), 585–603.
Robinson, Jonathan, Yeh, Ethan, 2011. Transactional sex as a response to risk in Western Kenya. Am. Econ. J. Appl. Econ. 3 (1), 35–64.
Shah, Manisha, 2013. Do sex workers respond to disease? Evidence from the male market for sex. Am. Econ. Rev. 103 (3), 445–450.
Sood, Neeraj, Wagner, Zachary, Wu, Yanyu, 2015. The impact of insurance on HIV testing. Am. J. Health Econ. 1 (4), 515–536.
Thirumurthy, Harsha, Pop-Eleches, Cristian, James, Habyarimana, Goldstein, Markus, Zivin, J Graff, 2012. Behavioral Responses of Patients in AIDS Treatment Programs: Sexual Behavior in Kenya, vol. 15, pp. 1–29.
Thirumurthy, Harsha, Zivin, Joshua Graff, Goldstein, Markus, 2008. The economic impact of AIDS treatment labor supply in Western Kenya. J. Hum. Resour. 43 (3), 511–552.
Wagner, Zachary, Barofsky, Jeremy, Sood, Neeraj, 2015. PEPFAR funding associated with an increase in employment among males in ten Sub-Saharan African countries. Health Aff. 34 (6), 946–953.
Wagner, Zachary, Wu, Yanyu, Sood, Neeraj, 2014. The Affordable Care Act may increase the number of people getting tested for HIV by nearly 500,000 by 2017. Health Aff. 33 (3), 378–385.
Wilson, David P., 2012. HIV treatment as prevention: natural experiments highlight limits of antiretroviral treatment as HIV prevention. PLoS Med. 9 (7), e1001231.
Wilson, Nicholas, 2011. Antiretroviral Therapy and Demand for HIV Testing: Evidence from Zambia. Available at SSRN 1982185.
Wilson, Nicholas L., Xiong, Wentao, Mattson, Christine L., 2014. Is sex like driving? HIV prevention and risk compensation. J. Dev. Econ. 106, 78–91.
Williams, B.G., Korenromp, E.L., Gouws, E., Schmid, G.P., Auvert, B., Dye, C., 2006. HIV infection, antiretroviral therapy, and CD4+ cell count distributions in African populations. J. Infect. Dis. 194 (10), 1450–1458.
Winston, Clifford, Maheshri, Vikram, Mannering, Fred, 2006. An exploration of the offset hypothesis using disaggregate data: the case of airbags and antilock brakes. J. Risk Uncertain. 32 (2), 83–99.
Young, Alwyn, 2005. The gift of the dying: the tragedy of AIDS and the welfare of future African generations. Q. J. Econ. 120 (2), 423–466.
Young, Alwyn, 2007. In sorrow to bring forth children: fertility amidst the plague of HIV. J. Econ. Growth 12 (4), 283–327.