Income Shocks and HIV in Africa*

Marshall Burke†, Erick Gong‡, and Kelly Jones§

9th April 2013

Abstract

We examine how variation in local economic conditions has shaped the AIDS epidemic in Africa. Using data from over 200,000 individuals across 19 countries, we match biomarker data on individuals' HIV status to information on local rainfall shocks, a large source of variation in income for rural households. We estimate that infection rates in HIV-endemic rural areas increase by 11% for every recent drought, an effect that is statistically and economically significant. Income shocks explain up to 20% of the variation in HIV prevalence across African countries, suggesting policy approaches for HIV prevention that are distinct from existing efforts.

JEL Codes: I15, O12, O55

Keywords: Income Shocks, HIV/AIDS, sub-Saharan Africa

---

*For useful comments, we thank Ted Miguel, Elisabeth Sadoulet, Jeremy Magruder, Nancy Padian, Ethan Ligon, and seminar participants at UC Berkeley and Middlebury College, and conference participants at Northeast Universities Development Consortium, Pacific Conference for Development Economics, Population and Poverty Network, and Population Association of America. All errors remain our own.

†Dept of Agricultural & Resource Economics, UC Berkeley: marshall.burke@berkeley.edu

‡Department of Economics, Middlebury College: egong@middlebury.edu

§International Food Policy Research Institute : k.jones@cgiar.org

1 Introduction

The relationship between income and health has long been of interest to economists, and a lengthy literature documents strong linkages between economic conditions and many important health outcomes (e.g. Currie, 2009). There has been much less progress, however, in understanding the economic foundations of the HIV/AIDS epidemic, one of the most important global health challenges. Such an understanding might yield particular dividends in sub-Saharan Africa (SSA), where over a million people continue to become newly infected with the disease each year (UNAIDS, 2010).

In this paper we explore the role of negative income shocks in shaping the evolution of the HIV/AIDS epidemic in Africa. Such shocks represent a well-documented challenge to poor households around the world. Lacking access to formal savings and insurance, income shortfalls often force poor households to make difficult tradeoffs between short-run consumption and longer-run earnings and human capital accumulation (Rosenzweig and Wolpin, 1993; Ferreira and Schady, 2009; Maccini and Yang, 2009). Recent indirect evidence suggests that variation in income could also affect important disease outcomes, either by altering individual sexual behavior (Baird et al., 2012; Kohler and Thornton, 2011; Robinson and Yeh, 2011b), or by affecting other phenomena such as migration or marriage timing that play a documented role in disease transmission (Lurie et al., 2003; Clark, 2004; Oster, 2012). Were income variation to play an important role in HIV outcomes through any of these mechanisms, it would suggest policy approaches for HIV prevention quite distinct from current focuses on behavior change campaigns, voluntary counseling and testing, male circumcision, and expansion of access to antiretroviral drugs (ARVs).

Using one of the most widespread sources of income variation in the developing world - rainfall-related shocks to agriculture - we directly assess the effect of negative income shocks on HIV outcomes across the African continent. We use the exogenous timing of rainfall events to develop an annual measure of shocks that is orthogonal to time-invariant determinants of disease outcomes. Using data on roughly two-hundred thousand individuals across nineteen African countries, we compare the HIV status of individuals randomly exposed to a higher number of recent shocks to the status of nearby individuals exposed to fewer recent shocks.

We find that exposure to recent negative rainfall shocks substantially increases HIV infection rates in rural areas with high baseline HIV prevalence. Exposure to a single additional shock leads to a significant 11% increase in overall HIV infection. These results are robust to a variety of ways of constructing the shock measure, to a variety of controls, and to a set of placebo tests. Consistent with expectations, we find little effect of shocks in urban areas (where incomes should be less sensitive to rainfall) and in low-prevalence regions (where there exists less HIV to be transmitted).

We show that these individual-level results are mirrored in the broader cross-country patterns of HIV prevalence observed in SSA. Using country-level data from UNAIDS, we show that exposure to shocks at the country level is also associated with significantly higher levels of HIV infection, and that our shock measure explains 14-21% of the cross-country variation in HIV prevalence across SSA. This provides somewhat independent evidence on the role of shocks in shaping HIV outcomes, and implies that meteorological bad luck earlier on in the AIDS epidemic could have played a substantial role in shaping how the epidemic progressed over the following decades.

While these reduced form results provide direct causal evidence that negative shocks substantially increase equilibrium HIV infection rates, they provide limited insight into the many channels through which shocks might shape HIV risk. For instance, adults may respond to shocks by temporarily migrating in search for work (Skoufias, 2003), or school-aged girls may respond by marrying at an earlier age to increase economic security (Jensen and Thornton, 2003), behaviors that are both associated with an increased risk for HIV (Lurie et al., 2003; Clark, 2004). Alternatively, women may increase their sexual activity in response to economic hardship in order to obtain transfers (both monetary and in-kind) from their male partners (LoPiccalo et al., 2012; Swidler and Watkins, 2007; Robinson and Yeh, 2011b; Dinkelman et al., 2008). This "transactional sex" has been documented among women who are not commercial sex workers in numerous African countries and is believed to be a key driver in the AIDS epidemic (UNAIDS, 2010).

While we are unable to definitively isolate the mechanism by which shocks increase HIV, we show that our data are largely inconsistent with either a migration or an early-sexual-debut explanation. In particular, we show that the effect of shocks is not larger for school-age girls or for households who appear to have better access to temporary migration opportunities. Furthermore, we show that the effects of shocks on HIV are particularly large for women working in agriculture (whose income would have likely declined the most following a shock) and men working outside of agriculture (whose purchasing power would have declined the least), providing suggestive evidence that an outward shift in the supply of transactional sex could be driving the increase in HIV following a shock. Finally, under stronger assumptions, we use an epidemiological model to estimate the magnitude of this shift in supply. We find that each shock leads to an increase of 1.42 sexual partners for women, an estimate that is both plausible and consistent with the existing literature.

We make several contributions. First, we contribute to the literature within and outside of economics that seeks to understand why the AIDS epidemic has disproportionately affected sub-Saharan Africa. Our results provide strong evidence that a primary source of income variation for rural Africans – rainfall-related variation in agricultural productivity – could be an important contributing factor to the epidemic. These results suggest that economic conditions play a significant role in the AIDS epidemic in SSA, and are related to previous work using macro-level data to explore the effects of economic growth on the AIDS epidemic (Oster, 2012).

We also contribute to a broader body of work on the health and livelihood consequences of income shocks. A host of papers show that when saving is difficult and insurance incomplete, negative income shocks can have seriously detrimental effects on longer-run livelihood outcomes. In contrast to existing work, we identify behavioral responses that are not only detrimental to an individual's or household's well being but that also generate large negative health externalities for the community. As such, our results add further impetus to the growing effort aimed at increasing access to risk management tools in the developing world, and could suggest a role for public subsidy if the negative health externalities brought on by incomplete insurance are as large as we estimate. Such policy measures are distinct from most current HIV prevention efforts, which primarily focus on medical interventions such as testing, male circumcision, and the provision of ARVs.

Finally, in our exploration of transactional sex as a potentially contributing behavior, we build directly on recent work that explores how the supply of risky sex responds to economic shocks. Most of the studies in this area focus on individual level income variation and risky sex, and provide evidence that decreases in a female's income lead to increases in risky sexual behavior (Robinson and Yeh, 2011a,b; Dinkelman et al., 2008; Baird et al., 2010; Kohler and Thornton, 2011). Other studies examine how aggregate level shocks affect the supply of risky sex, generally suggesting that riskier sex decreases when economic conditions are good and increases when they are bad (Dupas and Robinson, 2011; Wilson, 2011). We help generalise this literature in two ways. First, our study is one of the first to link negative economic shocks to actual disease outcomes rather than to self-reports of sexual behavior.[1] This is important because not only are HIV infections a primary outcome of interest for policy-makers, but biological markers of risky sex are also not subject to the social desirability bias of self-reports on sexual behavior (Padian et al., 2008; Cleland et al., 2004). Second, we estimate the relationship between economic shocks and HIV from pooled nationally representative surveys, helping us overcome concerns about generalizability from smaller-sample, localised studies.

The rest of the paper is organised as follows. In section 2 we present a simple conceptual framework to motivate our empirical approach. Section 3 presents the data and our empirical methods, and Section 4 discusses our main results, robustness checks, and explorations of contributing behaviors. Section 5 explores how these effects scale up to the country level. Section 6 discusses policy implications and concludes.

2 Conceptual Framework

The goal of this paper is to understand how economic conditions shape HIV risk. Our empirical approach examines how a plausibly exogenous source of income variation – exceptionally low rainfall realizations at a given location relative to

[1]There are three recent studies which study the effects of *positive* income transfers on risky sexual behavior using biological markers as primary outcomes. Baird et al. (2012) examining the effects of cash transfers on HIV and HSV-2 outcomes for teenage girls in Malawi, de Walque et al. (2012) study the effects of cash transfers conditioned on testing negative for STIs in rural Tanzania, and Duflo et al. (2011) examine the effects of educational interventions on HSV-2 outcomes in Kenya. Both Baird et al. (2012) and de Walque et al. (2012) find that positive income transfers lead to reductions in STIs, while Duflo et al. (2011) find that an educational subsidy combined with HIV-prevention information lowers STI rates for girls.

long-term averages ("shocks") – affects local HIV outcomes. Our primary result establishes a strong positive relationship between these shocks and local HIV prevalence. We argue that this is a causal relationship because our shock measure is, by construction, uncorrelated with other time-invariant factors that might also affect disease outcomes (see further discussion in section 3.2). Here we discuss why rainfall-related shocks might matter for HIV, and use this discussion to generate predictions of where and for whom the reduced form relationship between drought and HIV should be largest. We then discuss possible behaviors linking drought to HIV, and describe the patterns in the data that each behavior should generate.

Our empirical analysis begins by examining the reduced-form relationship between drought-related shocks (S) and HIV, or $\frac{\partial HIV}{\partial S}$. The link between these shocks and HIV can be seen in three stylized facts. First, in our Sub-Saharan African setting, heterosexual sex is the primary driver of the epidemic (UNAIDS, 2010), and so deviations in the path of the epidemic are driven largely by changes in sexual behavior. Second, a growing literature documents the importance of economic factors in shaping sexual behavior (Baird et al., 2012; Kohler and Thornton, 2011; Robinson and Yeh, 2011b) in Africa. Third, as is frequently recognised in the literature and as we demonstrate below, variation in rainfall generates substantial variation in both agricultural productivity and broader income measures in Africa.

We begin by taking these facts as given. Denoting sexual risk as p and income as z, the reduced form relationship between shocks and HIV can then be written as:

$$\frac{\partial HIV}{\partial S} = \frac{\partial HIV}{\partial p}\frac{\partial p}{\partial z}\frac{\partial z}{\partial S} \tag{1}$$

The three terms on the right hand side are the following:

- $\frac{\partial HIV}{\partial p}$ represents the relationship between HIV infection and sexual risk. Clearly the risk of HIV infection is increasing in risky sexual behavior $\left(\frac{\partial HIV}{\partial p} > 0\right)$ (Halperin and Epstein, 2008; Potts et al., 2008; Stoneburner and Low-Beer, 2004; Epstein, 2007). Importantly, this relationship also depends on the prevalence of HIV in an area (λ). Regions with higher HIV prevalence will have a stronger relationship between sexual behavior and new infections than regions with low prevalence $\frac{\partial HIV}{\partial p \partial \lambda} > 0$.

- $\frac{\partial p}{\partial z}$ represents the impact of a deviation in income on sexual risk (p). Sexual risk can be measured as the number of partners and/or number of unprotected sexual acts, but can also be measured by how likely a partner is infected with HIV. Below we discuss three main ways identified by the literature in which shortfalls in income might alter sexual behavior, all of which suggest a negative relationship between an income deviation and sexual risk for at least some subset of the population (i.e. $\frac{\partial p}{\partial z} < 0$). All three mechanisms can broadly be considered coping behaviors in response to income shocks, and will be operative for different subsets of the population depending on the coping mechanism in question.

- Finally, $\frac{\partial z}{\partial S}$ is the relationship between negative rainfall shocks and income shocks. Some economic activities are more sensitive to variation in rainfall than others, and we expect that in rural areas (r), where most income is generated from rain-fed agriculture, rainfall shocks will have a larger (negative) effect on income than in urban areas where agriculture is less important for the local economy $\left(\frac{\partial z_r}{\partial S_r} < \frac{\partial z_u}{\partial S_u} \leq 0\right)$.

Because there is little disagreement in the literature on the signs of the first and third terms in Equation 1, the overall sign of $\frac{\partial HIV}{\partial S}$ will depend on how sexual risk responds to variation in income. If we assume that this term is non-zero, then two immediate predictions are generated from Equation 1.

- The effect of shocks on HIV will be larger (in absolute value) where baseline prevalence λ is higher. Intuitively, if shocks increase HIV through changes in sexual behavior, the effect of shocks will be amplified in places where there is more HIV to transmit.

- The effect of shocks on HIV will be larger (in absolute value) in rural areas where income is more dependent on agriculture (and therefore on rainfall).

The sign of $\frac{\partial p}{\partial z}$ will determine the overall sign of $\frac{\partial HIV}{\partial S}$. If some segment of the population copes with negative income shocks in a way that increases sexual risk, as is suggested by the literature, then Equation 1 indicates that the overall relationship between HIV and shocks for these populations would be positive, $\frac{\partial HIV}{\partial S} > 0$.

2.1 Behavioral responses to shocks

How might changes in income induce behavioral changes that increase HIV risk? We discuss three risk-coping behaviors that are responses to income shocks and can lead to higher HIV risk: temporary migration, dropping out of school or marrying early, and transactional sex.

One response to drought-induced income shocks is to migrate from rural to urban areas in search of employment (Skoufias, 2003; Ellis, 2000). Migration is associated with greater levels of risky sexual activity and higher rates of HIV (Lurie et al., 2003; Brockerhoff and Biddlecom, 1999). Individuals may temporarily migrate to urban areas in response to droughts, acquire HIV due to additional partnerships or high-risk partners, and then infect others when returning to their rural communities.[2] If income shocks induce temporary migration, then $\frac{\partial p}{\partial z} < 0$ for both men and women, as both the migrant and his/her partner in the rural village would face increased risk. Because men are most often the temporary migrants, we might expect the effects on HIV to be larger for men if temporary migration is a behavior linking shocks and HIV.

A second set of coping behaviors that may affect sexual risk are changes in schooling and marriage behavior. In SSA, a common response to a negative income shock is to withdraw children from school (Ferreira and Schady, 2009), which appears particularly true for girls (Bjorkman, 2006). Once a girl has withdrawn from school she is much more likely to be sexually active and to marry (Osili and Long, 2008; Duflo et al., 2011; Ozier, 2010), both of which are risk factors for HIV (Clark, 2004; Baird et al., 2011). Furthermore, households may marry-off daughters earlier in response to a shock, especially in regions where bride payment is customary (Hoogeveen et al., 2011; Jensen and Thornton, 2003). If income shocks induce early drop-out and early marriage, which result in earlier sexual activity, then $\frac{\partial p}{\partial z} < 0$. While this could apply to both men and women, young women would be most affected through this channel.

A third coping mechanism is engaging in transactional sex. Transactional sex is thought to be common in sub-Saharan Africa, and is broadly defined to include both prostitution as well as exchanges within casual relationships and long-term

[2] Note that, if the migration is of a permanent nature, this should not affect HIV in the rural area, though it may affect *our estimation* of rural HIV, due to sample selection. We directly address this in section 4.2.

partnerships (Luke, 2006; Swidler and Watkins, 2007; Béné and Merten, 2008; Hunter, 2002; Maganja et al., 2007; Leclerc-Madlala, 2002).[3] While there are many factors affecting the HIV/AIDS epidemic, transactional sex is thought to be a major driver within SSA (Alary and Lowndes, 2004; Dunkle et al., 2004; Côté et al., 2004), and a growing empirical literature suggests that economic conditions affect the market for transactional sex (Robinson and Yeh, 2011b; Dupas and Robinson, 2011). Women may respond to income shocks by increasing their supply of transactional sex, either by taking on additional partnerships or engaging in more frequent or riskier sexual activity (i.e. unprotected sex). Both types of behaviors have been documented throughout sub-Saharan Africa, with women in rural Malawi engaging in multiple partnerships in response to income insecurity (Swidler and Watkins, 2007), and women in South Africa and Western Kenya more likely to engage in unprotected sex as a response to negative income shocks (Dinkelman et al., 2008; Robinson and Yeh, 2011b; Dupas and Robinson, 2011).

An interesting feature of this third mechanism is that, given the women are almost exclusively sellers and men almost exclusively buyers of transactional sex (e.g. Edlund and Korn, 2002), a negative income shock will increase a woman's supply of transactional sex while a man facing a shock will decrease his demand. In this sense, with respect to transactional sex, $\frac{\partial p}{\partial z} < 0$ for women and $\frac{\partial p}{\partial z} \geq 0$ for men. Therefore, if transactional sex is a key element in the link between rainfall and HIV, the effect of shocks should have differential effects by gender. However, because the equilibrium rate of HIV for a given subgroup (e.g. men) will depend on both their own behavior and the behavior of their partners, the fact that $\frac{\partial p}{\partial z} \geq 0$ for men does not automatically imply that $\frac{\partial HIV}{\partial S} < 0$ for men, because increases in HIV for their (female) partners could raise overall HIV rates for men as well. The prediction from this mechanism is instead that the effect of shocks on HIV will be larger for females (f) than for males (m) $\left(\frac{\partial HIV_f}{\partial S} > \frac{\partial HIV_m}{\partial S}\right)$.

In the empirical section, our main finding is that individuals exposed to drought events are more likely to be infected with HIV $\left(\frac{\partial HIV}{\partial S} > 0\right)$. Given the

[3] One could argue that early marriage as a response to an income shock may also be considered transactional sex in some form. We argue that these are conceptually distinct as early marriage would be an increase in sexual activity at the extensive, rather than the intensive margin. Further, these are distinct from a policy perspective.

strong evidence of both the relationship between droughts and income $\left(\frac{\partial z}{\partial S} > 0\right)$ and risky sexual behavior and HIV $\left(\frac{\partial HIV}{\partial p} > 0\right)$, this suggests that the underlying mechanism connecting droughts and HIV is a behavioral response to income shocks that is leading to increased sexual risk $\left(\frac{\partial p}{\partial z} < 0\right)$. We provide suggestive evidence that transactional sex is the mechanism that is most consistent with our results, but we cannot conclusively establish the primary behavior driving this result, nor can we rule out any single behavior as a contributing factor.

2.2 Other pathways

Each type of behavior discussed above - early sexual activity, migration, transactional sex - has a well-documented connection to HIV risk and a plausible link to community-level income shocks. However, droughts also have documented effects on other important factors in rural areas, such as nutrition and civil conflict. We argue that the evidence linking these factors to HIV outcomes is, at best, inconclusive, and that they are unlikely to be pathways that link shocks to HIV.

For HIV infected individuals, malnutrition is associated with higher mortality rates and higher viral loads (John et al., 1997; Weiser et al., 2009). Thus the effect that malnourished HIV-positive individuals will have on the epidemic is ambiguous; higher mortality rates would lead to fewer HIV-positive individuals but higher viral loads would make them more infectious. For HIV-negative individuals, little is known about the relationship between malnutrition and susceptibility to HIV infection (Mock et al., 2004). Though malnutrition may lead to a compromised immune system which could play a role in susceptibility (Schaible and Stefan, 2007), to the best of our knowledge there is no work that demonstrates an increase susceptibility to HIV infection for malnourished HIV-negative individuals. While we cannot rule out that this is a contributing pathway, given the existing evidence it does not appear to play a primary role in the HIV/AIDS epidemic.

In Miguel et al. (2004), the findings suggest that negative rainfall deviations are associated with higher incidence of civil conflict. This could suggest another pathway between rainfall and HIV if civil conflict has a direct effect on disease outcomes, for instance due to conflict-related sexual violence. While we again cannot directly rule out this possibility in our data, the existing evidence also suggests that rainfall-induced conflict is unlikely to be driving our results: recent existing studies find no clear link between conflict and HIV in either the observational data from Africa (Spiegel et al., 2007), or using epidemiological models that attempt to explain observed HIV prevalence with reported rates of sexual violence (Anema et al., 2008). We thus focus our empirical exploration of pathways on the the three coping behaviors described above.

3 Empirical Methods

3.1 Individual HIV-status data

Our individual-level data are taken from 21 Demographic and Health Surveys (DHS) conducted in 19 different Sub-Saharan countries.[4] Of the existing DHS surveys available in early 2011, we employ all those that include results from individual-level HIV-tests as well as longitude and latitude information on the individual's location, allowing us to map households to data on shocks.[5] For two countries (Kenya and Tanzania), two survey rounds matched these criteria; however, these are separate cross-sections and creation of panel data at the individual or cluster level is not possible. Nonetheless, for each country both rounds are included in the analysis as entirely separate surveys.

Each of these surveys randomly samples clusters of households from stratified regions and then randomly samples households within each cluster. In each sampled household, every woman aged 15-49 is asked questions regarding health, fertility, and sexual behavior.[6] A men's sample is composed of all men within a specified age range within households selected for the men's sample.[7] Depending on the survey, this is either all sampled households, or a random half (or third) of households within each cluster. Details regarding survey-specific sampling are presented in Appendix Table A.1. In all households selected for the men's sample, all surveyed men and women are asked to provide a finger-prick blood smear for

[4]A map of these countries can be found in Appendix A.

[5]The one exception is the Mali 2001 survey. We must exclude this survey as it is not possible to link the HIV results to individuals in the GIS-marked clusters.

[6]Mozambique 2009 samples women up to age 64.

[7]The age range for men is 15 to either 49, 54, 59 or 64, depending on the survey. See appendix A for details.

HIV-testing.[8] By employing cluster-specific inverse-probability sampling weights, the HIV prevalence rates estimated with this data are representative at the national level.[9]

Table 1 gives the list of included surveys along with basic survey information. The compiled data contains over 8,000 clusters. On average, there are 25 surveyed individuals per cluster, and 90% of clusters contain between 10 and 50 surveyed individuals. In total, there are over 200,000 individuals in the pooled data. Table 1 also shows HIV prevalence rates for each survey. Overall, women's prevalence is 9.2% and men's is 6.2%. However, these numbers mask a range that varies widely from over 30% prevalence for women in Swaziland to less than 1% prevalence in Senegal. Given that the sexual behavior response to income shocks will have different implications depending on HIV prevalence, we classify countries into two groups: low prevalence countries with less than 5%; and high prevalence countries with over 5% prevalence.[10]

Since the DHS surveys in each country were conducted in different years, we include survey fixed effects in all of our analysis. This controls for any effects that national policies might have on the HIV/AIDS epidemic as well as any time trends of the epidemic. Our analysis is thus focused on making comparisons within country in a given year.

3.2 Weather data and construction of shocks

To understand how economic shocks shape HIV outcomes, we seek a shock measure that satisfies three criteria: derived shocks are economically meaningful, they are orthogonal to other factors that might also shape disease outcomes, and they capture the potential disjoint between when HIV is acquired and when the individual is observed in the DHS. Because we do not directly observe variation

[8]Testing success rates for each survey are shown by sex in Appendix table A.2. Refusal rates are 10%, on average. Mishra et al. (2006) examine test refusal rates in DHS testing, which are between 1% to 22%, depending on the country. They conclude that although those refusing are more likely to be positive, the DHS testing accurately represents national prevalence. In this study, individuals exposed to shocks do not differentially refuse a test (see Appendix Table A.3) so non-response does not induce bias in our results.

[9]For details regarding construction of the weights, see Appendix A.

[10]This categorization follows UNAIDS (2010). Appendix Figure A.2 shows that with the exception of Cameroon, the prevalence classifications for each country remains stable for the ten years preceding the survey year.

in economic performance at a disaggregated level, and because such variation is likely endogenous to disease outcomes, we adopt an approach that is common in the literature and use variation in weather as a proxy for variation in economic productivity.[11] For the largely agrarian societies of Africa, variation in weather directly shapes the economic productivity of the majority of the population that continues to depend on agriculture for their livelihoods (Davis et al., 2010). As we show below, particularly negative rainfall realizations substantially depress agricultural productivity across the region.

Our weather data are derived from the "UDel" (University of Delaware) data set, a 0.5 x 0.5 degree gridded monthly temperature and precipitation data set (Matsuura and Willmott, 2009). These gridded data are based on interpolated weather station data and have global coverage over land areas from 1900-2008.[12] Using the latitude/longitude data in the DHS, we match each DHS cluster to the weather grid cell in which it falls. Because lat/lon data in the DHS are recorded at the cluster level, all individuals within a given cluster are assigned the same weather. Our DHS data match to 1701 distinct grid cells in the UDel data. To capture the seasonality of agriculture, we construct grid-level estimates of "crop year" rainfall, where the crop year is defined as the twelve months following planting for the main growing season in a region.[13] Annual crop year rainfall estimates are generated by summing monthly rainfall across these twelve "crop year" months at a given location.

To capture shocks to economic productivity that are both meaningful and orthogonal to potential confounders, we identify years in which accumulated rainfall was unusually low *relative to what is normally experienced in a particular location.*

[11]Other studies using rainfall variation in sub-Saharan Africa as an exogenous shock to income include Miguel (2005), who examines the relationship between income and witch-killings, Miguel et al. (2004), who instrument income with rainfall in a study of the effects of income shocks on civil conflict, and Hoddinott and Kinsey (2001), who examine the relationship between income and child health outcomes.

[12]0.5 degrees is roughly 50 kilometers at the equator. The UDel data are popular in economic applications (recent papers include Jones and Olken (2010); Dell et al. (2008); Bruckner and Ciccone (2011)). Other rainfall data sets are available, but none were sufficient for our needs, lacking either sufficient temporal coverage or spatial resolution.

[13]Estimates of planting dates are derived from gridded maps in Sacks et al. (2010); planting of staple cereal crops for the primary growing season typically occurs in the boreal (northern hemisphere) spring across most of West and Central Africa, and in the boreal autumn across most of Southern Africa.

In particular, for each of our 1701 grid cells, we fit the history of crop-year rainfall realizations to a grid-specific gamma distribution and assign each grid-year to its corresponding percentile in that distribution.[14] A "shock" is then defined as a realization below a pre-determined percentile in the location-specific distribution. The literature does not provide definitive estimates of the percentile below which a shock becomes meaningful, and unfortunately disaggregated (e.g. grid) measures of economic productivity over time are unavailable. To make progress, we construct an analogous measure of rainfall shocks at the country level and assess how country-level agricultural productivity and GDP growth respond to these shocks.[15] Resulting estimates from panel regressions of country level maize yields or GDP growth on percentile rainfall realizations (purged of country- and time-fixed effects) are shown in Figure 1. Maize is the continent's primary staple crop, the crop grown by the majority of smallholder farmers, and thus perhaps the best direct measure of rural incomes, and point estimates from these panel regressions suggest that realizations below about the 15th percentile are the most harmful to maize yields (Figure 1, left panel). A similar pattern is found in GDP growth (right panel). We thus adopt this 15% threshold as our initial measure of a "shock" - i.e. we define a shock as a crop-year rainfall realization below the 15% quantile of the local rainfall distribution - and show that our results are insensitive to other threshold choices in the neighborhood of 15%.

Finally, because the DHS only observes the disease status of a particular individual at one point in time, and an HIV+ individual could have become infected at any time over the previous decade or longer (median survival time at infection with HIV in sub-Saharan Africa, if untreated, is 9.8 years (Morgan et al., 2002)), our main independent variable is the number of these shocks that

[14] The gamma distribution was selected for its considerable flexibility in both shape and scale. Our results do not depend on the choice of gamma, or the estimation of the distribution more generally. Similar findings result from defining shocks as 1.5 standard deviations below the grid mean. We use the history of rainfall over the period 1970-2008, which was chosen to be a long enough period to be relatively insensitive to the recent shocks of interest, but short enough to capture relatively recent averages if long run means are changing (e.g. with climate change).

[15] That is, we aggregate crop year rainfall over all cells in a given country (weighting by crop area) to get a time-series of rainfall realizations for each country; we fit a separate gamma distribution to each country's time series; and within each country each year is assigned it's corresponding percentile in its gamma distribution. Crop yield data are from FAO (2011), and data on real per capita economic growth is from the Penn World Tables 7.0 (Heston et al., 2011).

have occurred over the 10 years prior to the survey year at a given location. For instance, if an individual was surveyed in the DHS in 2007, the shock variable takes on a value of between 0 and 10 corresponding to the number of crop-year rainfall realizations in that individual's region between 1997-2006 that fell below the 15% cutoff in the local rainfall distribution. We sum the shocks because acquiring HIV is irreversible – if a shock led to an HIV infection 7 years ago, and that individual is still alive, they will be HIV-positive today – and thus past shocks should have a demonstrable effect on current HIV infection. We note that using a more continuous measure of rainfall - e.g. deviations from average rainfall in levels - would tend to obscure past shocks: the sum of a very bad year and a very good year would be similar to the sum of two normal years. The mean and standard deviation of shocks by cluster are shown in Table 2.

By construction, this shock measure should be orthogonal to other confounding variables. Because shocks at a given location are defined relative to that location's historical rainfall distribution, and the same *percentile* cutoff is used in each location to define a shock (instead of the same absolute cutoff), all locations have the same expected number of shocks over any given 10 year period: each year any location has a 15% chance of experiencing a shock. But because rainfall in a given location varies over time, some 10-year time windows will accumulate more shocks than other windows, and it is this plausibly random variation that we exploit. We confirm in Appendix B that accumulated rainfall shocks are orthogonal to the first three moments of the rainfall distribution, providing additional confidence that our shock measure is uncorrelated with other time-invariant unobservables that might also affect HIV outcomes.

This definition of shocks assumes that relative (rather than absolute) deviations in rainfall are what matter for income and HIV outcomes. This construction is necessary for identification – using an absolute threshold for a shock would mean that areas with lower or more variable rainfall would expect more shocks, and these areas could differ in other unobserved ways that matter for HIV – but it is also plausibly captures what is important in our setting. Farmers choose crops that are adapted to the conditions under which they are grown, with farmers in drought-prone regions in Africa sowing crops (such as millet and sorghum) that can withstand low rainfall realizations, and farmers in areas with higher average rainfall sowing crops that are generally higher yielding but less tolerate of drought (e.g. maize). The results in Figure 1, which are constructed using this relative shock measure, confirm that relative deviations matter for both agricultural outcomes and broader economic performance.

3.3 Estimation

To explore the effects of negative income shocks on individual HIV rates, we estimate the following:

$$HIV_{ijk} = \alpha + \beta_1 S_j^t + X_i'\delta + \gamma r_j + \omega_k + \varepsilon_{ijk} \tag{2}$$

where HIV_{ijk} is an indicator for whether individual i in cluster j tested HIV-positive in survey k. S_j^t is the number of rainfall shocks that cluster j has experienced in the t years before the survey. The default indicator for S_j^t is the number of crop-years with rainfall at or below the 15% quantile in the last 10 years for a given cluster. Note again that by construction, no one cluster is any more shock prone than another, i.e. $E(S_m^t) = E(S_n^t)\ \ \forall j = m, n$. All clusters expect the same total number of shocks over the 38 years in our rainfall data, and our identifying variation comes from the random timing of these shocks: some clusters happen to receive more of their shocks in the decade immediately before we observe them, and others receive fewer. Both t and the definition of S are varied over a range to test the robustness of results.

The vector X_i contains characteristics of individual i that are not affected by shocks, specifically, gender and age. r_j indicates that cluster j is rural. The survey fixed effect is ω_k and ε_{ijk} is a mean-zero error term.[16] We estimate linear probability models, allowing for correlation of error terms across individuals in the same weather grid. Survey specific sampling weights are used to make the results representative of individuals living in these 19 countries in Sub-Saharan Africa (see Appendix A).

[16]There are a host of reasons for including survey fixed-effects. Innumerable differences across countries exist that we cannot observe, including social norms of sexual behavior, male circumcision rates, access to health services, and the national response to the AIDS epidemic. Such unobservable differences may also apply to different time periods within the same country, thus motivating a within-survey estimation.

4 Results

4.1 Main results

Table 3 shows estimations of Equation 2, employing various samples and interaction terms. The overall effect of shocks on HIV rates using the full sample is 0.3 percentage points (ppt) and is statistically significant at the 10% level (Column 1). Our simple conceptual framework predicts differential effects depending on whether an individual lives in an urban or rural area, and in line with this prediction we find that the effects are concentrated in rural areas. The point estimate in urban areas is zero and not statistically significant (Column 2; Linear combination), and the difference between estimates for rural and urban areas is borderline significant ($p - value = 0.102$). Focusing our analysis on rural areas (Column 3), we find that shocks have a meaningful effect: we estimate that each shock leads to a 0.3 ppt increase in HIV prevalence, an effect that is significant at the 5% level and that corresponds to a 7.3% increase in HIV rates given a mean of 4.1%.

The second prediction from our framework is that increases in risky behavior as a result of an income shock would result in little change in HIV infection rates if existing HIV prevalence is very low. To capture differential effects by baseline prevalence, we focus on the rural sample and include an interaction between shocks and an indicator for low-prevalence countries. In countries with low prevalence (less than 5%), shocks have an approximately zero effect on HIV (Column 4; Linear combination), and we reject equality across low and high prevalence countries with 95% confidence (Column 4; Shocks x Low Prevalence). Column 5 presents the estimation for the rural sample in high prevalence countries only. In these areas, each shock increases HIV by 0.8 ppt, an 11% increase based on overall prevalence of 7%.

Finally, column 6 disaggregates the impact by gender. We find that shocks increase the probability of infection by 0.9 ppt for women and 0.6 ppt for men, both of which are statistically significant at the 5% level. Given that HIV prevalence is 8.3% for women and 5.6% for men in high prevalence rural areas, these estimates represent large effect sizes of 11% increases in HIV per shock for both women and men. We cannot reject that the effect size is the same across genders

(Column 6; Interaction term).

4.2 Robustness of results

We have thus far argued that the timing of rainfall shocks is plausibly exogenous and that shocks are leading to significant increases in HIV rates. In this section, we examine whether our primary result – the large response of HIV to shocks in rural, high prevalence areas shown in Table 3, column 5 – is robust to various issues of specification, sample selection, variable definition, or omitted variables.

Specification

We first examine whether our results are sensitive to the specification or sample used. We begin by limiting the sample to individuals who are at least 15 years of age when a shock occurred. If shocks are leading to changes in sexual behavior, we would not expect there to be any effect on individuals who are not yet sexually active. We find results that are similar to our main specification (Table 4; Column 1). We sequentially remove all individual level controls, remove population weights, and replace survey-year-fixed effects with country- and year-fixed effects and our results remain stable (Columns 2-4). We also vary the sample used, removing hyper-endemic countries such as Swaziland and Lesotho where HIV-prevalence exceeds 20%, and our results remain stable (Column 5). Finally, within each DHS cluster (i.e. village), we remove all visitors from the sample, defined as those who have lived in the area for less than a year at the time of the survey. We do this for two reasons. First, we want to identify the effect of shocks on HIV for those who were actually living in the area at the time of the shock and removing visitors helps us establish this. Second, it may be that rainfall shocks are inducing NGO and government workers to migrate in drought afflicted areas, and if these types are more likely to be HIV+, than this could potentially explain our results. Removal of these visitors from the sample does not change our results (Column 6).

Shock definition

We also examine the sensitivity of our results to the definition of a shock. While our primary specification defines a shock as a crop-year rainfall realization below the 15th percentile of local realizations, the choice of 15th percentile is somewhat arbitrary. We vary the cut-off for shock definition in increments of

1 percent between the 5th and 40th percentile. The estimated coefficients for each percentile are presented in Figure 2. Overall, the point estimate is relatively stable around our default 15th percentile shock measure, and as the definition of a shock becomes less (more) severe the point estimates generally decrease (increase). Shocks in the neighborhood between between the 10th percentile to 20th percentile generate similar results (regression results shown in appendix Table C.1) For rainfall at or above the 40th percentile, there is no effect on HIV. This corresponds to the estimated relationships between rainfall and maize yields, and rainfall and GDP growth, shown in Figure 1. Both maize yields and GDP growth are unaffected by rainfall realizations above about the 40th percentile, and consistent with this we find that HIV becomes similarly unaffected by rainfall around when realizations cross this threshold.

As an alternative to the quantile-based definition, we also define shocks as rainfall that is 1.5 standard deviations or more below the historical mean for the area. The primary estimation employing this definition of shock is shown in column 6 of Table 4.A, where the estimated coefficient is similar, though slightly larger, and remains statistically significant.

We also vary the period of time over which shocks are summed, for comparison with our preferred definition of shocks summed over the past ten years. We sum shocks in 5-year bins (e.g. number of shocks 1-5 years before the survey, number of shocks 6-10 years before, etc.) and employ each of these binned variables as the regressor in our main specification. Figure 2 plots the point estimates of these regressors. As we show in Appendix E, this time profile of the effect of shocks on HIV is very much as we would expect, with point estimates for the effect of shocks peaking early within the 10-year window. Intuitively, an earlier shock has more time to reverberate through the population and generate additional infections compared to a more recent shock, but effects are attenuated over time as the earliest infected die; given the observed infection rate and the observed timing of mortality following infection, we show via simulation in the Appendix that the effect of a shock will peak 6-10 years later.

Sample selection

Droughts can also effect other types of behaviors that might explain our results. If shocks induce permanent out-migration and the migrants are disproportionately HIV negative, this could yield a correlation between observed shocks and higher HIV prevalence among the remaining population. In order to test whether selective migration can account for our results we conduct a bounding exercise motivated by Lee (2009). Using national rural and total population figures by country, we estimate that rural areas lose approximately 2% of population per shock (see Appendix D for more details) and conservatively assume that *each one* of these individuals is HIV-negative. We replace these individuals in our sample and re-estimate our main results. This, in effect, stacks the deck against finding a result: communities that experience shocks now have more HIV-negative individuals.

Table 5 first reproduces our primary result based on the rural sample of high-prevalence countries: the probability of infection increases by 0.9 percentage points per shock. We then vary the assumed percentage who migrate when a shock occurs, starting with our estimate of 2% and increasing in increments of 1%. We find that when accounting for estimated out-migration of 2% per shock, the estimated coefficient (0.7 ppt) is only slightly different from our original estimate, and still significant.

Note that if *all* of rural to urban migration were caused *only* by shocks, then a more accurate estimate would be that 4% of the population migrates when a shock occurs (again, see Appendix D for details). Thus, the assumption of 4% loss per shock is an extreme upper bound. When we replace a 4% population loss per shock, our effect remains positive (0.4 ppt) and significant at the 10% level. Though 4% is the upper bound, we nonetheless report estimations under the assumptions of 5% and 6% loss per shock to show that the estimate does not lose significance until we assume 6% loss per shock – three times our best estimation of 2% loss per shock. This suggests that sample selection due to permanent migration is unlikely to explain our results.

Omitted variables

A final concern is that our results might be driven by omitted variables. For example, some aspects of local weather might be correlated with other unobservables (wealth, education, etc) that also affect HIV rates. While this is unlikely to be true for our measure of rainfall shocks – by construction all areas expect the same total number of shocks over time – we confirm that our estimates are robust to controlling for characteristics of the underlying distribution. In Table 4, Panel B, we sequentially control for the first three moments of the rainfall distribution (mean, variance, skew) in our main specification (Columns 6-8), and also include all three moments (Column 9). Our estimate remains stable throughout these various specifications.

We can further test for these potential confounders with a "placebo" test - we check whether shocks in the future can predict present HIV rates or other observable present characteristics. Given that the DHS surveys were conducted between 2003 and 2009, and our weather data ends in crop-years 2007-2008, we are only able to examine shocks up to four years in the future.[17] We find no relationship between HIV rates and shocks 1 to 4 years in the future (Table 6; Columns 1-4). We also find no relationship between current wealth quintile and future shocks (Columns 5-7), nor any relationship between an individual's years of education and future shocks (Columns 8-10).

Finally, during the 2000's, there was increasing access to antiretrovirals (ARVs) for HIV-positive individuals, which may bias our results if access was in any way correlated with shocks. We show that during most of our study time frame, ARV access was relatively low (less than 30% for all but one country) and that there is no evidence that suggests ARV access is correlated with our shock measure (see Appendix F). Taken together these tests provide additional evidence that shocks are picking up meaningful variation in economic conditions prior to the survey year, and that this variation is uncorrelated with other factors that might also explain disease outcomes.

4.3 Exploring Coping Behaviors

Our reduced form results suggest that HIV infections are increasing in recent rainfall shocks. As laid out in section 2, we propose that several behavioral responses to drought could be leading to increases in sexual risk, thereby generating this result. To support this notion, we first examine whether risky sexual behaviors increase in response to recent shocks, using self-reported sexual behavior.

The use of self-reported sexual behavior is subject to a few caveats. There is a large body of evidence that suggests self-reported sexual behavior suffers from

[17] The only 2003 survey which has individual HIV infections (Kenya), does not have data on wealth and education. Therefore correlations with these characteristics can only be estimated using data in years 2004+, so these can only be observed up to three years in the future.

social desirability bias (Cleland et al., 2004) and that women significantly underreport their sexual activity (Minnis et al., 2009).[18] In addition, we only have measures of sexual behavior during the 12 months prior to the survey. It is not immediately clear which time window of shocks should be considered to impact sexual behavior in the past 12 months. Certainly shocks in the current and previous year should, however, given the potential lag between lack of rainfall and lack of income, perhaps droughts two years ago should have a similar impact. Further, more distant shocks that induced the creation of new sexual relationships may have continuing impacts on current behavior if those relationships (or behaviors) are persistent. For this reason, we present the impact on recent sexual behavior of shocks within the past 10 years, shocks within the past 5 years, and having a shock that affected income over the past 12 months. Given these caveats, we interpret results on self-reported sexual behavior with caution.

The outcome variables we examine are whether in the past 12 months the respondent has (i) been sexually active, (ii) had multiple partners, or (iii) had non-spouse partner(s).[19] Table 7 shows results of estimations of Equation 2, separately by gender, with these self-reported sexual behaviors as the dependent variables regressed separately on three categories of independent variables as noted. A strong and consistent finding is that both men and women are significantly more likely to have engaged with a non-spouse partner if exposed to a shock in any of the three time periods considered. For both men and women, shocks affecting the past 12 months increase non-spouse partnership rates by about 15%. Shocks in any of the periods also increase the likelihood of engaging with multiple concurrent partners by 10-15%, though the estimates are not precise in all periods. Point estimates for the impact of shocks on being sexually active at all are positive for men, but not significantly different from zero, and for women are not consistent across the periods considered.

[18]Additional caveats are that data that is available for sexual behavior doesn't capture all aspects of risky behavior that could lead to HIV infection. For example, the type of sexual partner you have (commercial sex worker, individual with multiple partners, etc.) will affect the likelihood of HIV infection, but such data are not available in the DHS. In addition, the questions about sexual behavior are not present in all the employed DHS surveys, and therefore the analysis is performed on a sub sample of our data.

[19]In this data, a monogamous cohabiting union is considered a spousal partner, irrespective of formal marital status. Also, single, sexually active individuals are included in those having non-spouse partners.

Overall, these self-reports of sexual behavior indicate that individuals who have experience recent shocks are more likely to report risky sexual activity. Keeping the caveats discussed earlier in mind, these findings suggest that shocks are indeed changing sexual behavior – and in particular leading to riskier sexual behavior – and that these behavioral changes are what could link rainfall shocks to HIV. In the remainder of this section, we seek evidence for whether any one of the coping behaviors discussed above is primarily responsible for this relationship.

Temporary Migration Rainfall shocks could induce rural individuals – particularly males – to migrate temporarily to cities in search of work. If these individuals become infected with HIV in urban areas and return to the countryside and infect their spouses, then such shock-related migration could increase HIV in rural areas. For some surveys, we have information on the number of times individuals have been away from home in the past 12 months, and whether any time away has lasted more than one month. This information is available for men in 16 (and for women in 8) of our 21 surveys. If temporary migration is a primary coping behavior in this setting, we would expect that recent shocks would significantly increase both indicators. The first four columns of Table 8 show that for both men and women, shocks affecting the past 12 months have a correlation with the number of times away from home and being gone for more than one month in the past year that is either negative or indistinguishable from zero. This suggests that in our rural sample, droughts are not inducing significant temporary migration.

As a further check on the temporary migration pathway, we examine whether shocks have a differential impact on individuals living closer to cities for whom temporary migration is likely more of an option. While proximity to a city is likely correlated with other unobservable factors that could also affect the impact of shocks, distance is nevertheless a meaningful constraint to migration in the many areas of Africa with poor transportation infrastructure. Under the assumption that individuals are more likely to temporarily migrate to urban areas if they reside nearer to one, then if temporary migration is indeed the pathway linking shocks to HIV, we would expect shocks to have a greater effect on HIV in rural areas that are nearer to urban areas. We classify households within 100 kilometers of an urban area to be "near to urban", and we vary the population threshold that qualifies a settlement as being "urban" (250,000 or 500,000 inhabitants).[20] We interact our measure of shocks with an indicator for "near to urban area," with the expectation that the interaction will have a positive coefficient if temporary migration is a pathway. The last four columns of Table 8 show that for both men and women, the coefficient on the interaction is consistently negative, suggesting that for those living near urban centers, the increased risk of HIV with each shock occurrence is mitigated. While we again interpret our distance result with caution, these results provide little evidence that temporary migration is the primary coping behavior that links droughts to HIV infections.

Dropping-out and Early marriage

A second possibility is that income shocks may affect dynamics in the marriage market. In particular, households may marry off their daughters earlier in response to shocks, leading to an earlier sexual debut and a higher risk of acquiring HIV for these women. The mean age at marriage for women in this sample is 18. The first column of Table 9 shows an estimate of the impact of shocks occurring when a woman was potentially subject to early marriage (aged 13 to 18) on her eventual age at marriage. The coefficient reflects an effective zero change in age at marriage when exposed to a shock at this critical age. Column 2 shows the same estimate for shocks occurring when a woman was aged 15 to 20. While the coefficient is negative, it amounts to a 0.1% effect, which is not distinguishable from zero. In short, it seems that shocks do not induce earlier marriage for women in this sample.

Even if youth are not marrying earlier, households may respond to income shocks by withdrawing children from school, especially girls. Girls that drop out early are at higher risk for early sexual activity and HIV transmission (Baird et al., 2010). If this is a contributing factor in the link between rainfall and HIV, we would expect to find two telltale results. First, shocks should reduce total schooling for women who were school-aged when the shock occurred; second, the link between rainfall and HIV should be restricted to women who had not yet completed their schooling when the shock occurred. Columns 3 and 4 of Table

[20] Our measures of distance from the nearest urban area, as well as the population size of settled areas, are derived from the Global Rural-Urban Mapping Project (CIESIN, 2010). Population data are from the year 2000, helping to mitigate concerns that urban population size could have itself responded to the shocks of interest (for most of our sample, the shocks of interest are post-2000).

9 estimate the effect of shocks when aged 13 to 18 (and 15 to 20) on years of education. Both estimates produce a negative coefficient, however, both reflect effect sizes of 0.8% or less and are not statistically different from zero. We do not find evidence that rainfall shocks induce significant dropping out of girls. Finally, columns 5 and 6 replicate our primary estimation, excluding women who were school aged during the past 10 years. We find that the results are robust to this exclusion, suggesting that women who were school-aged at the time of the shock are not driving the results. In sum, we find no evidence that early marriage and dropping out are the primary coping behavior linking rainfall to HIV.

Transactional Sex

A third coping behavior that may increase sexual risk is an increase in the supply of transactional sex by women in response to an income shortfall. We cannot directly examine changes in this behavior, as we lack data on transactions.[21] To make progress, we make a few assumptions on the transactional sex market. First, we follow the literature in assuming that women supply and men demand transactional sex. Second, in keeping with a recent micro literature (Baird et al., 2012; Kohler and Thornton, 2011; Robinson and Yeh, 2011b), we assume that women increase their supply of transactional sex if other sources of income decrease. Finally, we assume that individuals experiencing larger income shocks should have a stronger behavioral response – that is, supply is increasing and demand is decreasing in shock exposure.

While we do not observe individual changes in income, we do observe occupation – in particular, whether or not an individual's primary income source is from agriculture.[22] Under the simple assumption that incomes of individuals

[21]Whether a man has paid for sex in the past year is only queried in four surveys from high prevalence countries. This likely only captures explicit prostitution, rather than all forms of transactional sex, as the reporting is low (3%). Women are not queried regarding payment for sex in any of our surveys.

[22]We are able to classify individuals by their employment type at the time of the survey but not at the time of the shock. Our analysis thus makes the assumption that occupation is fairly persistent: individuals in agriculture at the time of survey are *more likely* to have been in agriculture at the time of the shock, and thus our occupational categories are meaningful. We include only those employed in the last year, as the unemployed do not report an occupation. As such, it is difficult to assume whether the currently unemployed previously worked in agriculture or not. A concern with using occupational category is that it may be endogenous to shocks. We examine the predictive effect of number of shocks in the past 10 years on current employment in rural areas, to check its potential to induce bias. Shocks have no predictive effect for employment in agriculture.

working in agriculture are more sensitive to drought than those working outside agriculture, we expect to see the largest effects of drought on coping behaviors among those employed in agriculture. That is, if transactional sex is a primary coping behavior in response to droughts, we would expect that the response of HIV to shocks would be greater for women working in agriculture than for women working outside agriculture. For men, we expect that those working in agriculture would reduce their quantity demanded, however they will also face increased network risk; in net, the effect of shocks on their HIV status should be dampened but not necessarily reversed, relative to men working outside agriculture.

Table 10 presents the primary estimation for both men and women by occupation and confirms these basic predictions. Women in agriculture experience large increases in HIV risk (Column 3), while the effects of shocks on women employed outside of agriculture are much smaller and not significantly different from zero. For agricultural men, the impact on HIV risk is muted: less than a third of the effect size for other men, and not distinguishable from zero. For both men and women, we cannot reject that the effect is the same across occupation groups. However, the differences in point estimates across groups are consistent with the predictions arising from an increase in the supply of transactional sex as the coping behavior linking rainfall and HIV.

Simulation of behavior change

To further investigate the plausibility of coping behaviors as the link between rainfall and HIV, we estimate the changes in underlying sexual behavior that would be needed to generate the observed increases in HIV. We then examine whether such changes are plausible and consistent with related estimates in the literature. To back out the actual change in underlying sexual behavior as a result of shocks, we follow the methodology developed by Gong (2012) and use a simple epidemiological model to estimate the change in sexual partnerships that would result in the 0.9 and 0.6 ppt increases in HIV infection reported in Table 3, Column 6.[23] The model takes as parameters the HIV transmission rate, condom

[23]We focus on the change in number of sexual partnerships following both Kremer (1996) and Oster (2005), who develop epidemiological models of HIV transmission. In addition, Kaplan (1990) suggests that HIV transmission is best represented by number of partners instead of number of sexual acts. However, our focus on the number of partnerships as opposed to the number of unprotected coital acts should not be interpreted that the number of acts is irrelevant. In order to make this model tractable, we make assumptions and hold certain parameters fixed

usage, and the likelihood of matching with an HIV-infected partner. As with all modeling exercises that involve sexual behavior and HIV infections, the estimates generated are sensitive to the parameter values. For these, we rely on values from the health and epidemiological peer-reviewed literature and are explicit on the assumptions we make (as fully described in Appendix G). We note that given this exercise relies on a number of assumptions, the results warrant caution. In addition, our goal is not to precisely estimate the underlying change in sexual behavior, but to see whether, given a reasonable set of assumptions, the change in sexual behavior that would generate the increase in HIV risk we document is plausible.

For men, we estimate that each shock leads to an increase of .65 partners, which is about one-eighth of the mean number of lifetime partners for men.[24] For women, the model suggests an additional 1.42 partners per shock would be needed to generate their .9 ppt change in HIV infection.[25] In this sample, women report on average 2.2 lifetime partners. Based on clinical trials using prostate-specific antigen, which detects sexual activity in the past 48 hours, Minnis et al. (2009) showed that women in Zimbabwe under-report sexual behavior by about 50%. This suggests an average of 4.4 lifetime partners per woman. Annualizing based on the average woman in the sample (age 28 with sexual debut at 16), this averages to about one partner every three years. The large increase in partnerships for women, relative to that for men, offers further suggestive evidence that the coping behavior linking rainfall to HIV is more likely transactional sex rather than migration or youth issues.

How do these estimated changes in sexual behavior compare to other studies? Robinson and Yeh (2011a) find that an individual level health shock that results in total income loss for one day leads a woman to increase her number of sexual partners the following day by 0.3, an 18% increase in their sample. We find that this is comparable to our findings that a year-long income shock increases a woman's lifetime partnerships by about 33%. Note however that this is a difficult comparison, as their study is based on a sample of 192 women who identify as sex workers and average more than 1.5 partners *per day.* Given the significant

(see Appendix G for full details).

[24]The 95% confidence interval for the increase in partnerships for men is (.61,.69).

[25]The 95% confidence interval for change in partners for females is (1.30,1.53).

difference between the samples, it is difficult to say whether one would expect the supply responses to be similar. Nonetheless, we report their results as they are the nearest comparison, and our results are broadly comparable in magnitude to what they report.

5 Macro level implications

Our results suggest that community-level economic conditions play an important role in an individual's risk of HIV infection. A natural question is the extent to which our results inform broader observed patterns of HIV prevalence on the continent. In other words, can income shocks help explain the striking country-level variation in HIV prevalence across sub-Saharan Africa? Given that our estimation strategy above uses only within-country variation, and that we only have individual-level HIV data for about half of the countries in the Sub-Saharan region spread out over different years, it's not obvious that our estimates should inform these broader patterns.

To address this question, we apply our basic approach to country-level estimates of HIV prevalence provided by UNAIDS. UNAIDS estimates of country level HIV prevalence over time build heavily on HIV surveillance data distinct from what is in the DHS (e.g. data from antenatal testing at designated clinics), and thus provide prevalence estimates that are somewhat independent from the DHS biomarker data we focus on above. We use the same gridded climate data to derive a time series of annual average rainfall for each country, where the observation for a given country-year is a weighted average of all the grid cells in that country, using percent of each cell covered by cropland as weights.[26] Similar to above, we calculate these annual rainfall totals for each country back to 1970, fit a separate gamma distribution to each country's time series, and define a shock as a year in which country-average rainfall fell below the 15th percentile in that country's rainfall distribution. We then seek to explain the cross-sectional prevalence in HIV in a given year as a function of accumulated shocks over the

[26] This provides country-level rainfall estimates that are relevant for agriculture but that are also effectively weighted by rural population density, since areas that are farmed more intensively in rural Africa tend to be areas with higher population density (given very small average farm plot size).

previous decade. This regression uses a different source of variation from our individual specifications (cross-country rather than within-country), uses data that are related but distinct, and includes many countries not in our individual-level data. It thus provides a test of the relationship between shocks and HIV that is substantially distinct from the results presented above.

Figure 3 plots these relationships for the two decades for which UNAIDS reports data. Countries with a higher number of shocks are more likely to have higher levels of HIV-prevalence; this is true both in the 1990s (left plot) when the epidemic was growing rapidly, as well as in the 2000s, when the epidemic has plateaued or started to decline in many countries. These simple cross sectional relationships are statistically significant and explain 14-21% of the cross-sectional variation in HIV prevalence across the continent (see Appendix H for regression results).[27] Again, as with our individual-level results this estimate is not picking up differences in underlying propensity to experience shocks (which could be correlated with other factors affecting HIV), but relies instead on the random timing of recent shock exposure.

We draw three implications from these results. First, the fact that we can replicate our basic micro level results using different sources of variation on both the left- and right-hand side gives us additional confidence that economic conditions exert significant influence on HIV outcomes. Second, our results suggest that bad luck with the weather might have played a surprising role in shaping observed patterns of the AIDS epidemic across the African continent: countries that were hit with large negative shocks during the early years of the epidemic have much higher infection rates many years later. Finally, and somewhat more speculatively, given that many areas in sub Saharan Africa lack social safety nets and depend heavily on rainfed agriculture, recurring droughts may play an important and prominent role in explaining why the AIDS epidemic has disproportionately affected sub-Saharan Africa.

[27]We also explore whether shocks can explain the time-path of the epidemic by looking at cross-country decadal *changes* in HIV prevalence as a function of accumulated shocks. Effect sizes are again large but not always quite significant at conventional levels (p=0.12 on the shock variable for 1990s changes), and we explain somewhat less of the cross-country variance in decadal trends than we do in levels. Nevertheless, results are broadly consistent with cross-sectional results.

6 Conclusion

Ultimately any halt to the AIDS epidemic will require a medical intervention, such as a vaccine or methods approximating one (e.g. the aggressive use of ARVs). However, our results suggest that economic factors, and in particular the ways in which individuals respond to changes in their economic environment, also play an important role in shaping outcomes in the epidemic. As such, our findings unite two widely-held notions among researchers in the HIV/AIDS community: that heterosexual sex is a primary driver of the AIDS epidemic in sub-Saharan Africa, and that economic conditions play some role in sexual behavior in these countries.

Our paper provides compelling evidence that a deterioration in economic conditions, in the form of rainfall-related income shocks, contributes significantly to both village- and country-level rates of HIV infection in sub-Saharan Africa. While there are several possible pathways linking shocks to HIV, the available evidence suggests that shock-induced increases in transactional sex could be a primary mechanism. Nonetheless, we cannot fully rule out that other risk coping mechanisms, such as early marriage, school drop-out, or migration, are also contributing factors.

Regardless of the pathway, the policy implications of these findings are substantial. If income shocks lead households to smooth income in ways that contribute to the epidemic, policies that prevent the need for these coping mechanisms would appear to yield large positive returns. Comprehensive social safety nets may unfortunately be an unrealistic short-run goal for many revenue and capacity-constrained governments on the continent. However, more targeted interventions such as access to credit and savings, weather-indexed crop insurance or the development of drought-resistant crop varieties could have an indirect affect on the spread of HIV by reducing the sensitivity of incomes to rainfall shocks. Our results suggest that the social returns to investments in these and related interventions could be much larger than previously thought, particularly in countries where HIV prevalence remains high.

References

Alary, M. and C. M. Lowndes (2004). The central role of clients of female sex workers in the dynamics of heterosexual HIV transmission in sub-Saharan Africa. *AIDS 18*(6), –.

Anema, A., M. Joffres, E. Mills, and P. Spiegel (2008). Widespread rape does not directly appear to increase the overall HIV prevalence in conflict-affected countries: So now what? *Emerging Themes in Epidemiology 5*(11), 1742–1776.

Baird, S., E. Chirwa, C. McIntosh, and B. Özler (2010). The short-term impacts of a schooling conditional cash transfer program on the sexual behavior of young women. *Health economics 19*(S1), 55–68.

Baird, S., R. Garfein, C. McIntosh, and B. Özler (2012). Effect of a cash transfer programme for schooling on prevalence of HIV and herpes simplex type 2 in Malawi: a cluster randomised trial. *The Lancet*.

Baird, S., C. McIntosh, and B. Özler (2011). Cash or condition? evidence from a cash transfer experiment. *The Quarterly Journal of Economics 126*(4), 1709–1753.

Béné, C. and S. Merten (2008, May). Women and fish-for-sex: Transactional sex, HIV/AIDS and gender in African fisheries. *World Development 36*(5), 875–899.

Bjorkman, M. (2006). Income shocks and gender gaps in education: Evidence from uganda. *Institute for International Economic Studies, University of Stockholm*.

Boily, M., R. Baggaley, L. Wang, B. Masse, R. White, R. Hayes, and M. Alary (2009). Heterosexual risk of HIV-1 infection per sexual act: systematic review and meta-analysis of observational studies. *The Lancet Infectious Diseases 9*(2), 118–129.

Brockerhoff, M. and A. Biddlecom (1999). Migration, sexual behavior and the risk of hiv in kenya. *International Migration Review*, 833–856.

Bruckner, M. and A. Ciccone (2011). Rain and the democratic window of opportunity. *Econometrica 79*(3), 923–947.

CIESIN (2010). Global rural urban mapping project. *http://sedac.ciesin.columbia.edu/gpw/*.

Clark, S. (2004). Early marriage and hiv risks in sub-saharan africa. *Studies in family planning 35*(3), 149–160.

Cleland, J., J. Boerma, M. Caraël, and S. Weir (2004). Monitoring sexual behaviour in general populations: a synthesis of lessons of the past decade. *Sexually Transmitted Infections 80*(suppl 2), ii1.

Côté, A.-M., F. Sobela, A. Dzokoto, K. Nzambi, C. Asamoah-Adu, A.-C. Labbé, B. Mâsse, J. Mensah, E. Frost, and J. Pépin (2004). Transactional sex is the driving force in the dynamics of HIV in Accra, Ghana. *AIDS 18*(6), –.

Currie, J. (2009). Healthy, wealthy, and wise: Socioeconomic status, poor health in childhood, and human capital development. *Journal of Economic Literature 47*(1), 87–122.

Davis, B., P. Winters, G. Carletto, K. Covarrubias, E. Quiñones, A. Zezza, K. Stamoulis, C. Azzarri, and S. DiGiuseppe (2010). A cross-country comparison of rural income generating activities. *World Development 38*(1), 48–63.

de Walque, D., W. Dow, R. Nathan, R. Abdul, F. Abilahi, E. Gong, Z. Isdahl, J. Jamison, B. Jullu, S. Krishnan, et al. (2012). Incentivising safe sex: a randomised trial of conditional cash transfers for HIV and sexually transmitted infection prevention in rural Tanzania. *BMJ Open 2*(1).

Dell, M., B. Jones, and B. Olken (2008). Climate change and economic growth: evidence from the last half century. *NBER Working Paper 14132*.

Dinkelman, T., D. Lam, and M. Leibbrandt (2008). Linking poverty and income shocks to risky sexual behaviour: evidence from a panel study of young adults in Cape Town. *South African Journal of Economics 76*, S52–S74.

Duflo, E., P. Dupas, and M. Kremer (2011). Education, HIV and early fertility: experimental evidence from Kenya. *Manuscript*.

Dunkle, K. L., R. K. Jewkes, H. C. Brown, G. E. Gray, J. A. McIntryre, and S. D. Harlow (2004, October). Transactional sex among women in Soweto, South Africa: prevalence, risk factors and association with HIV infection. *Social Science & Medicine 59*(8), 1581–1592.

Dupas, P. and J. Robinson (2011). The (hidden) costs of political instability: Evidence from Kenya's 2007 election crisis. *Working Paper*.

Edlund, L. and E. Korn (2002). A theory of prostitution. *Journal of Political Economy 110*(1), 181–214.

Ellis, F. (2000). The determinants of rural livelihood diversification in developing countries. *Journal of Agricultural Economics 51*(2), 289–302.

Epstein, H. (2007). *The invisible cure: Africa, the West, and the fight against AIDS*. Viking.

FAO (2011). Faostat database. *UN Food and Agricultural Organization*.

Ferreira, F. and N. Schady (2009). Aggregate economic shocks, child schooling, and child health. *The World Bank Research Observer 24*(2), 147–181.

Friedman, W. (2012). Antiretroviral drug access and behavior change. *Working Paper*.

Gong, E. (2012). HIV testing and risky sexual behavior. *Unpublis*.

Groh, K., C. Audet, A. Baptista, M. Sidat, A. Vergara, S. Vermund, and T. Moon (2011). Barriers to antiretroviral therapy adherence in rural mozambique. *BMC Public Health 11:650*.

Halperin, D. and H. Epstein (2008). Opinion: Why is HIV prevalence so severe in southern Africa? the role of multiple concurrent partnerships and lack of male circumcision-implications for HIV prevention. *Southern African Journal of HIV Medicine 8*(1), 19.

Heston, A., R. Summers, and B. Aten (2011, June). Penn world table version 7.0. Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania.

Hoddinott, J. and B. Kinsey (2001). Child growth in the time of drought. *Oxford Bulletin of Economics and Statistics 63*(4), 409–436.

Hollingsworth, T., R. Anderson, and C. Fraser (2008). HIV-1 transmission, by stage of infection. *Journal of Infectious Diseases 198*(5), 687.

Hoogeveen, J., B. van der Klaauw, and G. van Lomwel (2011). On the timing of marriage, cattle, and shocks. *Economic Development and Cultural Change 60*(1), 121–154.

Hunter, M. (2002). The materiality of everyday sex: thinking beyond 'prostitution'. *African Studies 61*(1), 99–120.

Jensen, R. and R. Thornton (2003). Early female marriage in the developing world. *Gender and Development 11*(2), 9–19.

John, G., R. Nduati, D. Mbori-Ngacha, J. Overbaugh, M. Welch, B. Richardson, J. Ndinya-Achola, J. Bwayo, J. Krieger, F. Onyango, et al. (1997). Genital shedding of human immunodeficiency virus type 1 DNA during pregnancy: association with immunosuppression, abnormal cervical or vaginal discharge, and severe vitamin A deficiency. *Journal of Infectious Diseases 175*(1), 57.

Jones, B. and B. Olken (2010). Climate shocks and exports. *The American Economic Review 100*(2), 454–459.

Kaplan, E. (1990). Modeling hiv infectivity: must sex acts be counted? *JAIDS Journal of Acquired Immune Deficiency Syndromes 3*(1), 55.

Kohler, H. and R. Thornton (2011). Conditional cash transfers and HIV/AIDS prevention: Unconditonally promising? *World Bank Economic Review Advance Access Online*.

Kremer, M. (1996). Integrating behavioral choice into epidemiological models of aids. *The Quarterly Journal of Economics 111*(2), 549–573.

Leclerc-Madlala, S. (2002). Youth, HIV/AIDS and the importance of sexual culture and context. *Social Dynamics 28*(1), 20–41.

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

LoPiccalo, K., J. Robinson, and E. Yeh (2012). Income, income shocks, and transactional sex. *Handbook of the Economics of Prostitution*.

Luke, N. (2006). Exchange and condom use in informal sexual relationships in urban Kenya. *Economic Development and Cultural Change 54*(2), 319–48.

Lurie, M., B. Williams, K. Zuma, D. Mkaya-Mwamburi, G. Garnett, M. Sweat, J. Gittelsohn, and S. Karim (2003). Who infects whom? HIV-1 concordance and discordance among migrant and non-migrant couples in South Africa. *AIDS 17*(15), 2245.

Maccini, S. and D. Yang (2009). Under the weather: Health, schooling, and economic consequences of early-life rainfall. *American Economic Review 99*(3), 1006–1026.

Maganja, R. K., S. Maman, A. Groves, and J. K. Mbwambo (2007). Skinning the goat and pulling the load: transactional sex among youth in Dar es Salaam, Tanzania. *AIDS Care 19*(8), 974–981.

Magruder, J. (2011). Marital shopping and epidemic AIDS. *Demography 48*(4), 1401–1428.

Matsuura, K. and C. Willmott (2009). *Terrestrial Temperature and Precipitation: 1900-2008 Gridded Monthly Time series, version 2.1.* University of Delaware.

Miguel, E. (2005). Poverty and witch killing. *Review of Economic Studies 72*(4), 1153–1172.

Miguel, E., S. Satyanath, and E. Sergenti (2004). Economic shocks and civil conflict: An instrumental variables approach. *Journal of Political Economy 112(4)*, 725–752.

Minnis, A. M., M. J. Steiner, M. F. Gallo, L. Warner, M. M. Hobbs, A. van der Straten, T. Chipato, M. Macaluso, and N. S. Padian (2009, Oct). Biomarker validation of reports of recent sexual activity: results of a randomized controlled study in Zimbabwe. *American Journal of Epidemiology 170*(7), 918–924.

Mishra, V., M. Vaessen, J. Boerma, F. Arnold, A. Way, B. Barrere, A. Cross, R. Hong, and J. Sangha (2006). HIV testing in national population-based surveys: experience from the Demographic and Health Surveys. *Bulletin of the World Health Organization 84*(7), 537–545.

Mock, N., S. Duale, L. Brown, E. Mathys, H. O'Maonaigh, N. Abul-Husn, S. Elliott, et al. (2004). Conflict and hiv: A framework for risk assessment to prevent hiv in conflict-affected settings in africa. *Emerg Themes Epidemiol 1*(1), 6.

Morgan, D., C. Mahe, B. Mayanja, J. Okongo, R. Lubega, and J. Whitworth (2002). HIV-1 infection in rural Africa: is there a difference in median time to AIDS and survival compared with that in industrialized countries? *AIDS 16*(4), 597.

Ojikutu, B. (2007). The realities of antiretroviral therapy rollout: overcoming challenges to successful programmatic implementation. *Journal of Infectious Diseases 196*(Supplement 3), S445–S448.

Osili, U. and B. Long (2008). Does female schooling reduce fertility? evidence from nigeria. *Journal of Development Economics 87*(1), 57–75.

Oster, E. (2005). Sexually Transmitted Infections, Sexual Behavior, and The HIV/AIDS Epidemic. *Quarterly Journal of Economics 120*(2), 467–515.

Oster, E. (2012). Routes of infection: Exports and HIV Incidence in Sub-Saharan Africa. *Journal of the European Economic Association 10*(5).

Ozier, O. (2010). The impact of secondary schooling in Kenya: a regression discontinuity analysis. *University of California at Berkeley, mimeo*.

Padian, N., A. Buvé, J. Balkus, D. Serwadda, and W. Cates Jr (2008). Biomedical interventions to prevent HIV infection: evidence, challenges, and way forward. *The Lancet 372*(9638), 585–599.

Pinkerton, S. D. (2008, Sep). Probability of HIV transmission during acute infection in Rakai, Uganda. *AIDS Behav 12*(5), 677–684.

Potts, M., D. Halperin, D. Kirby, A. Swidler, E. Marseille, J. Klausner, N. Hearst, R. Wamai, J. Kahn, and J. Walsh (2008). Reassessing HIV prevention. *Science 320*(5877), 749.

Powers, K., A. Ghani, W. Miller, I. Hoffman, A. Pettifor, G. Kamanga, F. Martinson, and M. Cohen (2011). The role of acute and early HIV infection in the spread of HIV and implications for transmission prevention strategies in Lilongwe, Malawi: a modelling study. *The Lancet 378*(9787), 256–268.

Powers, K., C. Poole, A. Pettifor, and M. Cohen (2008). Rethinking the heterosexual infectivity of HIV-1: a systematic review and meta-analysis. *The Lancet Infectious Diseases 8*(9), 553–563.

Rehle, T., T. Saidel, S. Hassig, P. Bouey, E. Gaillard, and D. Sokal (1998). Avert: a user-friendly model to estimate the impact of HIV/sexually transmitted disease prevention intervetions on HIV transmission. *AIDS 12(suppl 2)*, S27–S35.

Robinson, J. and E. Yeh (2011a). Risk-coping through sexual networks: evidence from client transfers in Kenya. *Journal of Human Resources forthcoming.*

Robinson, J. and E. Yeh (2011b). Transactional sex as a response to risk in Western Kenya. *American Economic Journal: Applied Economics 3*(1), 35–64.

Rosenzweig, M. and K. Wolpin (1993). Credit market constraints, consumption smoothing, and the accumulation of durable production assets in low-income countries: Investments in bullocks in india. *Journal of Political Economy*, 223–244.

Sacks, W., D. Deryng, J. Foley, and N. Ramankutty (2010). Crop planting dates: an analysis of global patterns. *Global Ecology and Biogeography 19*(5), 607–620.

Schaible, U. and H. Stefan (2007). Malnutrition and infection: complex mechanisms and global impacts. *PLoS Medicine 4*(5), e115.

Schlenker, W. and D. Lobell (2010). Robust negative impacts of climate change on African agriculture. *Environmental Research Letters 5*, 014010.

Skoufias, E. (2003). Economic crises and natural disasters: Coping strategies and policy implications. *World Development 31*(7), 1087–1102.

Spiegel, P., A. Bennedsen, J. Claass, L. Bruns, N. Patterson, D. Yiweza, and M. Schilperoord (2007). Prevalence of HIV infection in conflict-affected and displaced people in seven sub-Saharan African countries: a systematic review. *The Lancet 369*(9580), 2187–2195.

Stoneburner, R. and D. Low-Beer (2004). Population-level HIV declines and behavioral risk avoidance in Uganda. *Science 304*(5671), 714.

Swidler, A. and S. Watkins (2007). Ties of dependence: AIDS and transactional sex in rural Malawi. *Studies in family planning 38*(3), 147–162.

Todd, J., J. Glynn, M. Marston, T. Lutalo, S. Biraro, W. Mwita, V. Suriyanon, R. Rangsin, K. Nelson, P. Sonnenberg, D. Fitzgerald, E. Karita, and B. Zaba (2007). Time from HIV seroconversion to death: a collaborative analysis of eight studies in six low and middle-income countries before highly active antiretroviral therapy. *AIDS 21*, S55–S63.

UNAIDS (2010). *Report on the Global AIDS Epidemic.* UNAIDS.

van Dijk, H., D. Foeken, and K. van Til (2001). *Mobile Africa: Changing Patterns of Movement in Africa and Beyond.* Brill. Chapter 2: Population Mobility in Africa, An Overview.

van Dijk, J., C. Sutcliffe, B. Munsanje, F. Hamangaba, P. Thuma, and W. Moss (2009). Barriers to the care of HIV-infected children in rural Zambia: a cross-sectional analysis. *BMC Infectious diseases 9*(1), 169.

Weiser, S., K. Fernandes, E. Brandson, V. Lima, A. Anema, D. Bangsberg, J. Montaner, and R. Hogg (2009). The association between food insecurity and mortality among HIV-infected individuals on HAART. *JAIDS Journal of Acquired Immune Deficiency Syndromes 52*(3), 342.

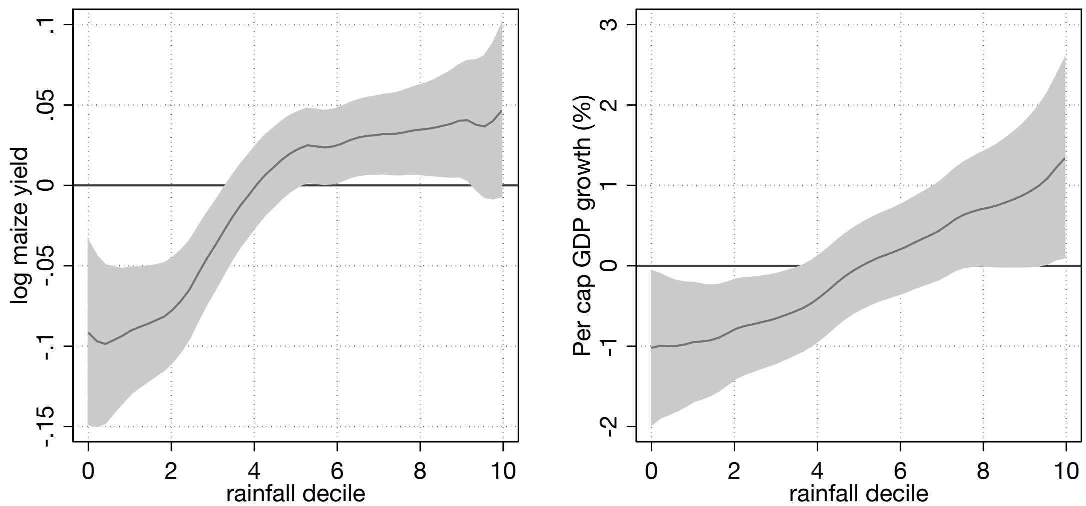
Weller, S. and K. Davis (2002). Condom effectiveness in reducing heterosexual HIV transmission. *Cochrane Database Syst Rev 1*.

WHO (2006). Progress on global access to hiv antiretroviral therapy: A report on "3 by 5" and beyond.

Wilson, N. (2011). Economic growth and the HIV/AIDS pandemic: Evidence from the early 21st century copper boom. *Working Paper*.

Figures

Figure 1: Effect of rainfall shocks on African maize yields (left panel) and per capita GDP growth (right panel)



Data are at the country level over the period 1970-2008, and include all sub-Saharan African countries. Dark lines display point estimates from kernel-weighted local polynomial regressions of the outcome on rainfall percentiles, after removing country and year fixed effects. Grey areas represent 95% confidence intervals. Data sources are given in the text.

Figure 2: Effect of rainfall shocks on HIV, by severity and timing

Figure 2.A: Shock severity

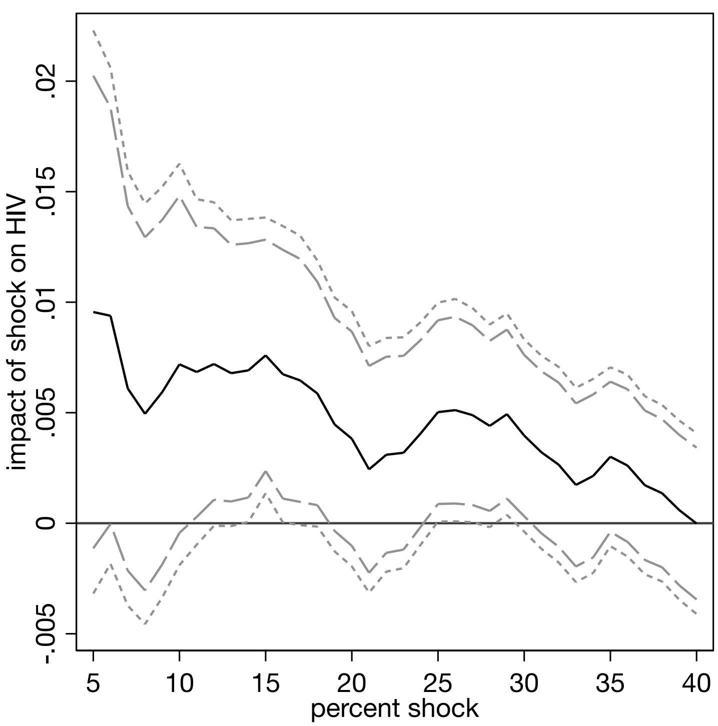
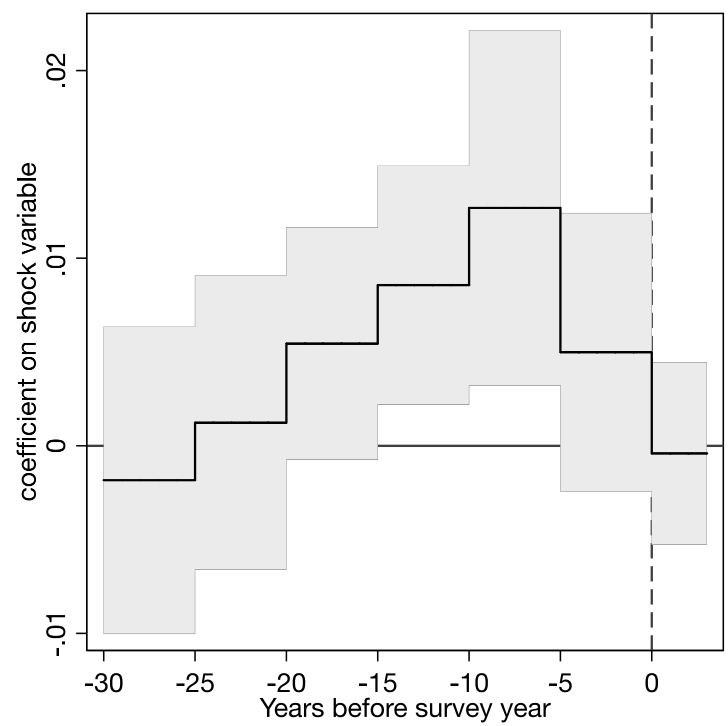
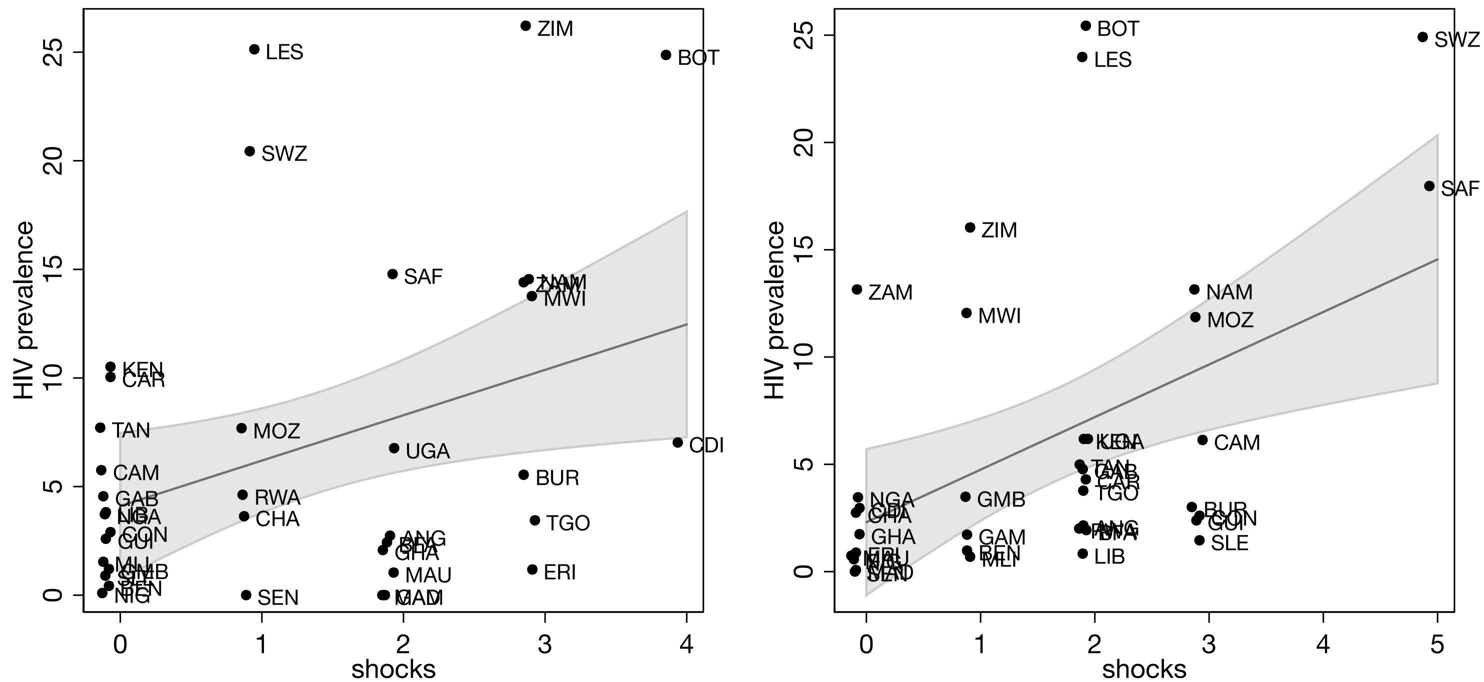


Figure 2.B: Shock timing



The black line represents the coefficient point estimates of the impact of a particular rainfall shock on HIV, using (A) various definitions of "shock" accumulated over the previous 10 years and (B) 15% shocks accumulated over different time periods, including placebo future shocks up to 3 years past the survey date. Plot A: dotted lines represent the 95% confidence intervals; dashed lines represent the 90% confidence intervals. Plot B: shaded area represents 95% confidence interval.

Figure 3: Country-level HIV prevalence & Shocks



The left panel presents results for HIV prevalence in 1999 (y-axis) and accumulated shocks over the previous decade (x-axis). The right panel presents results for HIV prevalence in 2008 and accumulated shocks since 2000. HIV data are from UNAIDS (2010). Dark lines are linear least squares fits, with gray areas representing the 95% confidence interval. Data are jittered to make country labels more legible.

Tables

Table 1: DHS Survey Information

				Prevalence			
	Country	Year	Individuals	Female	Male	Overall	Category
1	Swaziland	2007	8,186	31.1%	19.7%	25.9%	High
2	Lesotho	2004	5,254	26.4%	18.9%	23.2%	High
3	Zambia	2007	26,098	21.1%	14.8%	18.1%	High
4	Zimbabwe	2006	10,874	16.1%	12.3%	14.2%	High
5	Malawi	2004	5,268	13.3%	10.2%	11.8%	High
6	Mozambique	2009	10,305	12.7%	9.0%	11.1%	High
7	Tanzania	2008	10,743	7.7%	6.3%	7.0%	High
8	Kenya	2003	6,188	8.7%	4.6%	6.7%	High
9	Kenya	2009	6,906	8.0%	4.6%	6.4%	High
10	Tanzania	2004	15,044	6.6%	4.6%	5.7%	High
11	Cameroon	2004	10,195	6.6%	3.9%	5.3%	High
12	Rwanda	2005	10,391	3.6%	2.2%	3.0%	Low
13	Ghana	2003	9,554	2.7%	1.6%	2.2%	Low
14	Burkina Faso	2003	7,530	1.8%	1.9%	1.9%	Low
15	Liberia	2007	11,688	1.9%	1.2%	1.6%	Low
16	Guinea	2005	6,767	1.9%	1.1%	1.5%	Low
17	Sierra Leone	2008	6,475	1.7%	1.2%	1.5%	Low
18	Ethiopia	2005	11,049	1.9%	0.9%	1.4%	Low
19	Mali	2006	8,629	1.5%	1.1%	1.3%	Low
20	Congo DR	2007	8,936	1.6%	0.9%	1.3%	Low
21	Senegal	2005	7,716	0.9%	0.4%	0.7%	Low
	Total		203,796	9.2%	6.2%	7.8%	

Prevalence estimates are weighted to be representative at the national level.

Table 2: Shock Prevalence by Country

Prevalence Rank	Country	Survey Year	Mean Shocks	SD Shocks	Number of Clusters	Weather Grids
1	Swaziland	2007	2.90	0.46	275	13
2	Lesotho	2004	1.89	0.44	405	18
3	Zambia	2007	0.84	0.75	319	146
4	Zimbabwe	2006	1.28	0.76	398	122
5	Malawi	2004	1.04	0.75	521	53
6	Mozambique	2009	2.54	1.51	270	115
7	Tanzania	2008	0.77	0.82	345	167
8	Kenya	2003	1.17	0.62	400	81
9	Kenya	2009	1.22	0.78	398	93
10	Tanzania	2004	1.92	0.93	475	178
11	Cameroon	2004	1.59	1.06	466	112
12	Rwanda	2005	2.37	0.61	462	14
13	Ghana	2003	1.31	0.80	412	71
14	Burkina Faso	2003	1.28	0.90	400	88
15	Liberia	2007	1.35	1.05	298	37
16	Guinea	2005	1.34	0.75	295	72
17	Sierra Leone	2008	3.00	0.00	353	27
18	Ethiopia	2005	1.12	1.12	535	167
19	Mali	2006	1.00	0.71	407	149
20	Congo DR	2007	1.89	1.06	300	168
21	Senegal	2005	0.70	0.69	376	61
	Total		1.51	1.04	8110	1701

Table 3: Effect of Shocks on HIV

	All	All	Rural	Rural	Rural Hi Prevalence	Rural Hi Prevalence
	(1)	(2)	(3)	(4)	(5)	(6)
Num. shocks past 10 yrs.	.003*	.004**	.003**	.007**	.008**	.009**
	(.001)	(.002)	(.002)	(.003)	(.003)	(.004)
Shocks * Urban		-.004				
		(.002)				
Shocks * Low Prevalence Co.				-.008**		
				(.003)		
Shocks * Male						-.003
						(.003)
Interaction p-value		.104		.016		.243
Linear combination		-.000		-.000		.006**
		(.002)		(.001)		(.003)
Observations	202216	202216	134874	134874	77760	77760
R^2	.053	.053	.046	.046	.030	.030
Mean Dependent Var	.050	.050	.041	.041	.070	.070

Column headers indicate sample employed. Specifications include controls for gender and age, rural/urban designation (where applicable), and survey fixed effects. Estimations are weighted to be representative of the 19 countries. Robust standard errors are shown in parentheses clustered at the grid level. "Interaction p-value" is the p-value for Shocks X Urban (column 2), Shocks X Low Prevalence Co. (column 4), and Shocks X Male (column 6). "Linear combination" is the sum of coefficients on the number of shocks and the interaction term in each specification. For column 2, the linear combination is (Num shocks past 10 years) + (Shocks X Urban), column 4 is (Num shocks past 10 years) + (Shocks X Low Pow Prevalence Co.), and column 6 is (Num shocks past 10 years) + (Shocks X Male).

Table 4: Robustness to Specifications

Panel A: Robustness to specifications and sample

	Age 15+ at shock (1)	No Controls (2)	No Wgts (3)	CoYrFE (4)	No Hypr (5)	No Visitors (6)	SD Shock (7)
Num. shocks past 10 yrs.	.010**	.007**	.005*	.008**	.008**	.011***	
	(.004)	(.003)	(.003)	(.003)	(.003)	(.003)	
Precip 1.5 SD Shock Past 10 Years							.014**
							(.007)
Observations	42669	77760	77760	77760	68287	53596	77760
R^2	.032	.020	.068	.030	.025	.035	.030
Mean Dependent Var	.095	.070	.110	.070	.068	.063	.070

Panel B: Robustness to controlling for moments of rainfall distribution

	Mean (6)	Variance (7)	Skew (8)	All Moments (9)
Num. shocks past 10 yrs	.009***	.007**	.008**	.008***
	(.003)	(.003)	(.003)	(.003)
Observations	77760	77760	77760	77760
R^2	.031	.030	.030	.031
Mean Dependent Var	.070	.070	.070	.070

Rural sample from high-prevalence countries. All specifications include controls for gender, age and survey fixed effects, except as noted. Estimations are weighted to be representative of the 19 countries, except as noted. Robust standard errors are shown in parentheses clustered at the grid level.

Table 5: Robustness to sample selection from permanent migration

		Replacing lost population share ***per shock***				
	Observed	2%	3%	4%	5%	6%
	(1)	(2)	(3)	(4)	(5)	(6)
Num. shocks past 10 yrs.	.008***	.007**	.006**	.005*	.004*	.004
	(.003)	(.003)	(.003)	(.003)	(.003)	(.002)
Observations	77760	81792	84191	86523	88775	91330
R^2	.030	.022	.023	.023	.024	.024

Rural sample from high-prevalence countries. Column headers denote the population share added to the sample to account for out-migration, assuming all out-migrants are HIV negative. All specifications include controls for gender, age and survey fixed effects. Estimations are weighted to be representative of the 19 countries. Robust standard errors are shown in parentheses clustered at the grid level.

Table 6: Placebo Tests

Dependent Variable	HIV				Wealth quintile			Yrs of Education		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Num. shocks in future 1 yr	.006				.005			-.152		
	(.006)				(.110)			(.232)		
Num. shocks in future 2 yrs		-.002				.113			-.135	
		(.007)				(.095)			(.194)	
Num. shocks in future 3 yrs			-.004				.104			-.146
			(.007)				(.096)			(.194)
Num.shocks in future 4 yrs				-.006						
				(.006)						
Observations	49523	43881	26059	12434	49523	43881	26059	49489	43861	26039
R^2	.044	.040	.025	.010	.031	.033	.029	.119	.118	.102

Rural sample from high-prevalence countries. Note that the only survey in 2003 (Kenya) does not contain information on wealth and education; therefore, the correlations of these characteristics with shocks can only be calculated up to three years in the future, as weather data ends in 2007-2008 crop years. All specifications include controls for gender, age and survey fixed effects. Estimations are weighted to be representative of the 19 countries. Robust standard errors are shown in parentheses clustered at the grid level.

Table 7: Exploring Behaviors: Increasing risky sexual behavior

	Women			Men		
	Sexually Active	Multiple Partners	Nonspouse Partner	Sexually Active	Multiple Partners	Nonspouse Partner
	(1)	(2)	(3)	(4)	(5)	(6)
Num. shocks past 10 yrs.	.013**	.002	.007*	.008	.018***	.015**
	(.005)	(.002)	(.004)	(.006)	(.005)	(.006)
Observations	43145	43119	43147	34607	34563	34613
R^2	.060	.011	.018	.223	.034	.051
Num shocks past 5 yrs	.022***	.004*	.013**	.008	.015**	.021**
	(.007)	(.002)	(.005)	(.009)	(.006)	(.009)
Observations	43145	43119	43147	34607	34563	34613
R^2	.060	.011	.018	.223	.033	.050
Y/N shock affecting past 12mo	-.020*	.004	.019**	.010	.007	.045***
	(.011)	(.004)	(.009)	(.010)	(.011)	(.015)
Observations	43145	43119	43147	34607	34563	34613
R^2	.059	.011	.018	.223	.032	.051
Mean of Dep Var.	.759	.024	.120	.738	.154	.269

Rural sample from high-prevalence countries. Dependent variables are sexual behaviors in the past year. "Non-spouse" indicates sex with a non-spouse partner; this includes all sex for single individuals. All specifications include controls for age and survey fixed effects. Estimations are weighted to be representative of the 19 countries. Robust standard errors are shown in parentheses clustered at the grid level.

Table 8: Exploring Behaviors: Temporary migration

Dependent Variable:	Times away in past yr		Away month+ in past yr		HIV (where urban is defined as...)			
					Pop>250K		Pop>500K	
	Men	Women	Men	Women	Men	Women	Men	Women
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Y/N shock in past year	-.429***	-.295**	-.013	.016				
	(.135)	(.117)	(.021)	(.016)				
Num. shocks in past 10 yrs					.007**	.016***	.006*	.012***
					(.004)	(.005)	(.003)	(.004)
Near Urban * 10 yrs shocks					-.008*	-.019***	-.008	-.018*
					(.005)	(.006)	(.007)	(.009)
Near Urban					.012	.032***	.004	.019
					(.009)	(.012)	(.012)	(.015)
Observations	23802	22990	26133	26300	34613	43147	34613	43147
R^2	.025	.116	.004	.016	.032	.029	.031	.028
Mean of Dep Var	2.064	.990	.151	.130				
Mean of Near Urban					.264	.256	.163	.155

Rural sample from high-prevalence countries. Variables on being away are not available for all countries (see text). The “Near urban” variable indicates whether a given cluster is within 100km of an urban area (defined as the population size in the column header). Urban populations are from the Global Rural-Urban Mapping Project.All specifications include controls for age and survey fixed effects. Estimations are weighted to be representative of the 19 countries. Robust standard errors are shown in parentheses clustered at the grid level.

Table 9: Exploring Behaviors: Early school drop-out and marriage

Dependent Variable:	Age at marriage		Years of Educ		HIV Status	
					Aged 25+	Aged 30+
	(1)	(2)	(3)	(4)	(5)	(6)
Num. shocks when aged 13 to 18	.001		-.021			
	(.047)		(.055)			
Num. shocks when aged 15 to 20		-.029		-.042		
		(.055)		(.054)		
Num. shocks in past 10 yrs.					.010*	.015**
					(.006)	(.006)
Observations	24491	23839	28167	26088	12280	3845
R^2	.033	.022	.222	.220	.031	.022
Mean Dep Var	18.1	18.3	5.5	5.4	.091	.067

Female, rural sample from high-prevalence countries. The first four columns examine the impacts of shocks that occurred when woman was in the noted age range. The last two columns examine the impact of shocks in the past 10 years on HIV for women who were above a minimum age during all of the past 10 years. All specifications include controls for age and survey fixed effects. Estimations are weighted to be representative of the 19 countries. Robust standard errors are shown in parentheses clustered at the grid level.

Table 10: Exploring Behaviors: Impact on HIV by exposure to drought-induced income shock

	Men		Women	
	Ag	Non-Ag	Ag	Non-Ag
	(1)	(2)	(3)	(4)
Num. shocks in past 10 yrs	.002	.009*	.011***	.004
	(.003)	(.004)	(.004)	(.006)
Observations	18845	18740	20586	16901
R^2	.021	.043	.019	.035
Mean Dep Var	.056	.096	.077	.148

Employed sample from high-prevalence countries. "Ag" indicates that the individual works in agriculture. All specifications include controls for age and survey fixed effects. Estimations are weighted to be representative of the 19 countries. Robust standard errors are shown in parentheses clustered at the grid level.

Appendices (for online publication)

A DHS Data

Weighting

Sampling weights are used in this paper so that estimated effects represent the average effect of the population of interest (the population of 19 sub-Saharan African countries). The sampling weights are constructed as follows.

- Each individual is assigned an inflation factor that is $\rho = {}^{N_c}/_{n_c}$ where n_c is the sample size for survey in which he appears, and N_c is the population of his country in the year of that survey.

- Further, each individual has a survey-specific inflation factor h that is provided in the DHS data. h is the inverse probability of his HIV test results being present in the data. MEASURE DHS calculates h based on an individual's probability of being sampled for HIV testing (based on stratification of the survey) and his probability of providing a blood sample if requested, based on observable characteristics.

- A composite weight that is the product of ρ and h is employed in all specifications. A robustness check shows that the primary results of this work are not dependent on the use of sampling weights.

Figure A.1: Countries included in the study. Darker shades corresponding to higher HIV prevalence.

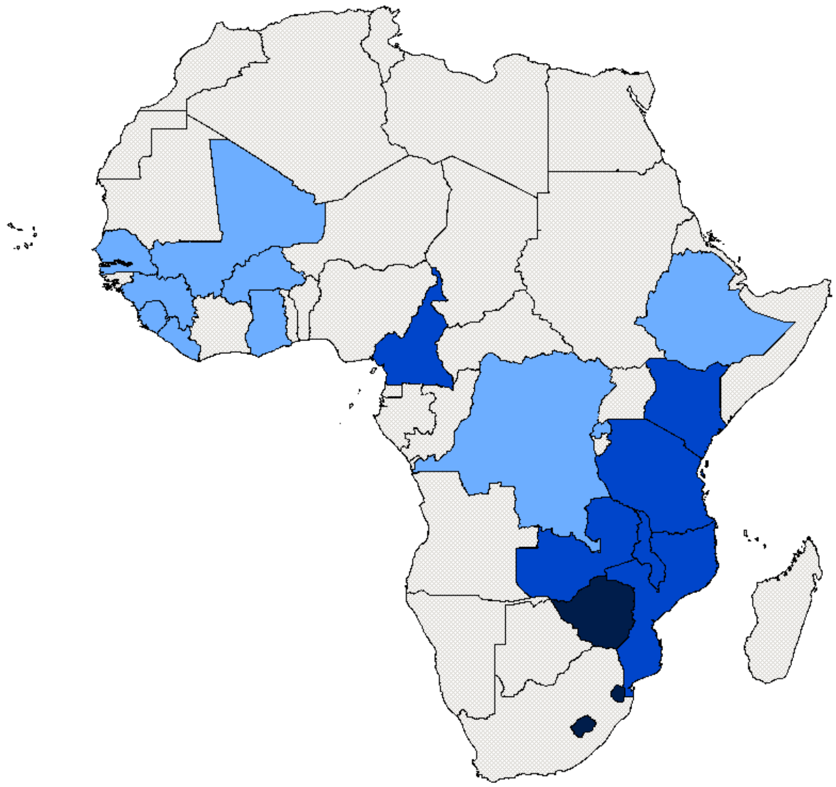
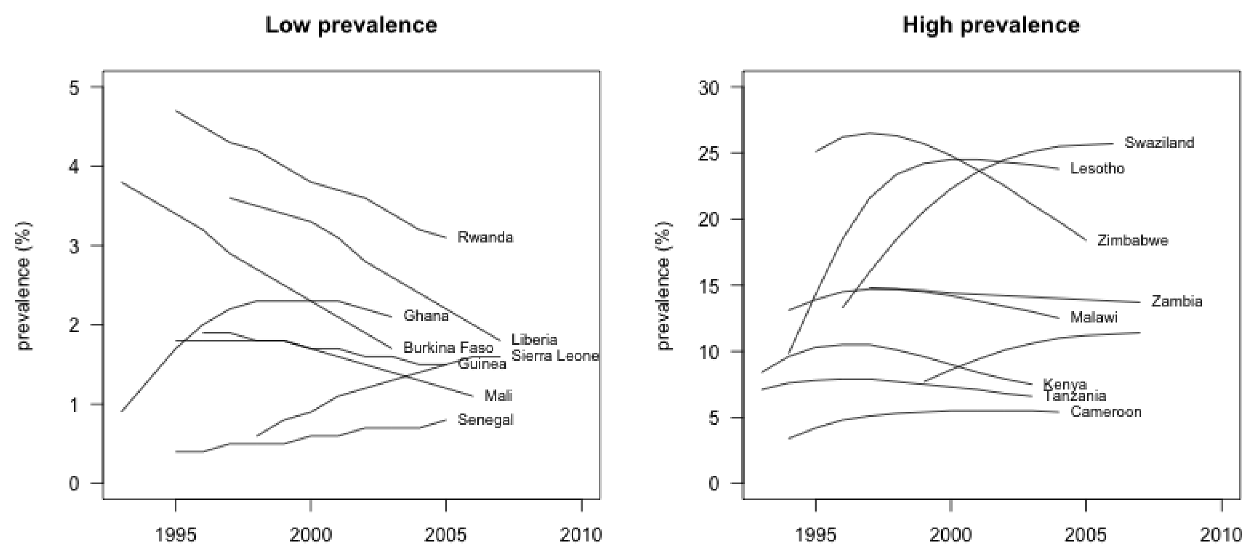


Figure A.2: Pre-survey 10-year HIV trends, Low and High Prevalence Countries.



For each country, we take the ten years preceding the survey year and plot yearly estimates of HIV prevalence from UNAIDS (2010). Ethiopia and Democratic Republic of Congo are not included in the figures as UNAIDS does not have historical estimates of HIV-prevalence for either country. We assume that both countries remained in the low prevalence category over the past ten years.

Table A.1: DHS Sampling for Serostatus Testing

Country	Year	Men Aged	Women Aged
Testing in all sampled households			
Mozambique	2009	12-64	12-64
Swaziland*	2007	15-49	15-49
Tanzania	2004, 2008	15-49	15-49
Liberia	2007	15-49	15-49
Zimbabwe	2006	15-54	15-49
Zambia	2007	15-59	15-49
Ghana	2003	15-59	15-49
Testing in random 50% of sampled households			
Sierra Leone**	2008	6-59	6-59
Kenya	2003, 2009	15-49	15-49
Lesotho	2004	15-59	15-49
Cameroon	2004	15-59	15-49
Congo DR	2007	15-59	15-49
Ethiopia	2005	15-59	15-49
Guinea	2005	15-59	15-49
Rwanda	2005	15-59	15-49
Testing in random 33% of sampled households			
Malawi	2004	15-54	15-49
Burkina Faso	2003	15-59	15-49
Mali	2006	15-59	15-49
Senegal	2005	15-59	15-49

* Swaziland: additional HIV testing for those aged 12-14 and 50+ in a random 50% of sampled households. ** Sierra Leone: Individual questionnaires were administered only to those aged 15-49 (59 for men)

Table A.2: Non-response for Serostatus Testing

		Men		Women	
Country	Year	Tested	Refused	Tested	Refused
Lesotho	2004	68%	16.6%	81%	12.0%
Swaziland	2007	78%	16.6%	87%	9.5%
Zimbabwe	2006	63%	17.4%	76%	13.2%
Malawi	2004	63%	21.9%	70%	22.5%
Mozambique	2009	92%	6.1%	92%	6.1%
Zambia	2007	72%	17.6%	77%	18.4%
Cameroon	2004	90%	5.6%	92%	5.4%
Kenya	2003	70%	13.0%	76%	14.4%
Kenya	2009	79%	7.8%	86%	8.2%
Tanzania	2008	80%	8.0%	90%	6.3%
Tanzania	2004	77%	13.9%	84%	12.3%
Burkina Faso	2003	86%	6.6%	92%	4.4%
Congo DR	2007	86%	5.7%	90%	4.4%
Ethiopia	2005	75%	12.6%	83%	11.2%
Ghana	2003	80%	10.7%	89%	5.7%
Guinea	2005	88%	8.5%	93%	5.0%
Liberia	2007	80%	11.3%	87%	7.3%
Mali	2006	84%	4.8%	92%	3.2%
Rwanda	2005	96%	1.9%	97%	1.1%
Sierra Leone	2008	85%	5.5%	88%	4.7%
Senegal	2005	76%	16.0%	85%	9.9%
Average		79%	11%	86%	9%

Rates are for the full HIV testing sample, with the exception of Mozambique. Rates for MZ are for the 15-49 sample. Rates are those reported in the DHS final reports for each survey, as the outcome of HIV measuring at the individual level is not included as an indicator in most data sets.

Table A.3: Non-response is not correlated with Shocks

Dependent Variable –>	Refused	Refused	Selected but not tested	Selected but not tested
	(1)	(2)	(3)	(4)
Num. shocks past 10 yrs.	-.002	-.002	.001	.001
	(.005)	(.004)	(.003)	(.003)
Indiv. controls	No	Yes	No	Yes
Mean of Dep. Var	.100	.100	.118	.118
95% CI for coeff	(-.011, .007)	(-.009, .006)	(-.006, .008)	(-.005, .006)
Observations	70547	70547	190794	190794
R^2	.026	.034	.032	.045

Note: Whether or not a selected individual refused an HIV test is the dependent variable in columns 1 and 2. The outcome of the test request, including refusal and failure to test for other reasons is given only for women in the following surveys: 2005 (ZW), 2006 (ML, SZ), 2007 (DRC, LB, ZM), 2008 (SL, KE), is given for men and women in 2007 TZ, and is not given at all in the remaining 12 surveys. For this reason, columns 3 and 4 employ all surveys and use lack of HIV test result as the dependent variable. Recall that only a sub-sample of households were selected for the men's survey and HIV testing (see section 3.1), and we endeavor to include only individuals from these households in this analysis. Selection into the sub-sample is not indicated in the data, and thus these households are only identifiable by the existence of an interview with, or a test result from, a male in the household. For households without data on a male, we are not able to identify the selected households (some households were selected but had no male present, for example). The sample employed in columns 3 and 4 includes all individuals in households that have data on a male in any of the surveys. These individuals were definitely selected for HIV testing, though they are not *all* of the individuals selected for testing.

The estimates suggest a fairly precise zero effect of shocks on test refusal or non-response. Based on the 95% confidence intervals, we can reject that a shock affects testing rates by more than one percentage point.

B Weather data and Impact of drought on crop yields

To help confirm that our measure of recent rainfall shocks is plausibly exogenous and not correlated with other moments of the rainfall distribution, we regress the number of rainfall shocks in the past 10 years on the mean, variance, and skewness of each grid's rainfall distribution. Table B.1 presents the results. In all specifications, these correlations are not significant. In other words, when we estimate across grids, recent rainfall shocks are orthogonal to all three moments of the historical distribution.

While we cannot directly show the importance of rainfall shocks for household income (as noted, the DHS do not include income or consumption measures), aggregate data suggest that these shocks are economically important. To demonstrate this, we construct a country-level on maize yields, real per capita economic growth, and country-level rainfall (Data are from FAO (2011), Heston et al. (2011) and Matsuura and Willmott (2009), respectively). Maize is the most widely grown crop in Africa, and we have data for 41 Sub-Saharan African countries for 1961-2008.[28] We similarly have data on real per capita income growth for these same countries across 1961-2008.

The first four columns of Table B.2 shows the impact of rainfall dropping below the 10th or 15th percentile on (log) country-level maize yields across Sub-Saharan African countries, based on panel regressions using country and year fixed effects. Annual maize yields are strongly affected by precipitation: yields are about 12% lower in a year with rainfall at or below the 15th percentile, and 18% lower in a year with rainfall below the 10th percentile. Results are robust to including temperature shocks in the regression. With 60-80% of rural African incomes derived directly from agriculture, these productivity impacts likely represent significant shocks to household incomes (Davis et al., 2010).[29] We repeat

[28]The included countries are: Angola, Benin, Botswana, Burkina Faso, Burundi, Cameroon, Central African Republic, Chad, Congo, Côte d'Ivoire, Democratic Republic of the Congo, Eritrea, Ethiopia, Gabon, Gambia, Ghana, Guinea, Guinea-Bissau, Kenya, Lesotho, Liberia, Madagascar, Malawi, Mali, Mauritania, Mozambique, Namibia, Niger, Nigeria, Rwanda, Senegal, Sierra Leone, South Africa, Swaziland, Togo, Uganda, United Republic of Tanzania, Zambia, and Zimbabwe.

[29]Schlenker and Lobell (2010) demonstrate that these strong negative impacts of weather

the same regressions using growth in real per capita GDP as the outcome variable (Columns 5-8). Negative rainfall shocks again reduce growth rates dramatically, with bigger shocks leading to larger declines in growth rates. We estimate that a 15% shock reduces the economic growth rate in that year by 1.8 percentage points, and a 10% shock by 1.9 percentage points. This demonstrates again that rainfall shocks exert substantial influence on economic productivity in Africa.

Table B.1: Rainfall Shocks and Overall Variability

Dependent Variable: Number of 15% rainfall shocks in past 10 years.

	(1)	(2)	(3)	(4)
mean	.000			.000
	(.000)			(.000)
variance		-.000		-.000
		(.000)		(.000)
skew			-.181	-.128
			(.143)	(.167)
Observations	1701	1701	1701	1701
R^2	.181	.182	.185	.194

Estimation at the grid level, with country fixed effects. Robust standard errors are shown in parentheses, clustered at the country level.

shocks generalize to other African staples, not just maize.

Table B.2: Impact of precipitation shocks on maize yields and per capita GDP growth.

	(1) yield	(2) yield	(3) yield	(4) yield	(5) GDP	(6) GDP	(7) GDP	(8) GDP
10 PCT shock	-0.180***				-1.880***			
	(0.016)				(0.475)			
15 PCT shock		-0.118***				-1.821***		
		(0.023)				(0.631)		
20 PCT shock			-0.148***				-1.694***	
			(0.025)				(0.600)	
30 PCT shock				-0.055				-1.155
				(0.039)				(0.817)
Constant	0.015	0.003	0.009	-0.005	5.490*	5.505*	5.495*	5.506*
	(0.078)	(0.077)	(0.078)	(0.084)	(2.810)	(2.812)	(2.830)	(2.768)
Observations	1537	1537	1537	1537	1533	1533	1533	1533
R squared	0.270	0.259	0.271	0.249	0.156	0.157	0.157	0.154

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Dependent variable is the log of country-level maize yield (columns 1-4) or real per capita GDP growth (columns 5-8). Regressions cover years 1970-2008 and include country fixed effects, year fixed effects, and a constant, and are weighted by country average maize area (maize regressions) or country population (GDP regressions). Errors are clustered at the country level. Yield data are from FAO (2011), GDP data are from the Penn World Tables (Version 7.0), and weather data are from UDel.

C Robustness to Shock Definition

Table C.1: Vary Shock Definition: 10 to 20%

Dependent Variable: HIV Infection						
Percent Shocks	**10%**	**11%**	**12%**	**13%**	**14%**	**15%**
Num Shocks past 10 yrs	0.007	0.007*	0.007*	0.007*	0.007**	0.008**
	(0.005)	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)
Number of observations	77,760	77,760	77,760	77,760	77,760	77,760
Percent Shocks	**16%**	**17%**	**18%**	**19%**	**20%**	
Num Shocks past 10 yrs	0.007**	0.006*	0.006*	0.004	0.004	
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	
Number of observations	77,760	77,760	77,760	77,760	77,760	

Rural sample from high-prevalence countries. All specifications include controls for gender, age and survey fixed effects. Estimations are weighted to be representative of the 19 countries. Robust standard errors are shown in parentheses clustered at the grid level.

D Estimating sample selection due to out-migration

In order to estimate the rate at which shocks affect permanent out-migration rates, we begin with an estimate of rural-to-urban migration. For each country in our sample, we calculate the reduction in rural population (as a share of total population) over a recent 10-year period, based on data from the World Bank.[30] On average, the rural share of the populations of these countries is reduced by 5.8% over ten years. Given the wide range of reasons for migrating to urban areas, migration in response to a shock likely accounts for no more than 20%-30% of this total (van Dijk et al., 2001). Nonetheless, we conservatively assume that low-rainfall shocks account for as much as half of this migration, and therefore induce a rural population loss of approximately 2.9% *over ten years*. Note that this is the accumulated loss from all shocks occurring during a ten year period.

We use these estimates to back-out the share of population that leaves during each shock. The column headers in Table D.1 show several possible assumptions of population loss *per shock* ranging from 1% to 5%. A bit of algebra reveals that if, for example, 3% of the population leaves *during each shock*, a village with three shocks over the past ten years has lost 8.73% of its population in that time. The calculation of lost population by number of shocks and assumption maintained are shown in the body of table D.1. By applying these calculations to the rural clusters in our data according to each cluster's number of shocks, we calculate the total population lost in our rural sample over the ten years preceding the survey. The bottom row of table D.1 shows these estimates of total population loss over a ten year period.

The second column, which assumes that 2% of the population leaves per shock predicts that the rural sample has lost 2.91% of the population over the ten year period. This prediction aligns the best with the estimate that rural areas lose 2.9% of population to drought-induced migration over a ten year period. Therefore 2% loss per shock is the assumption maintained. Notice that, if we assume that *all* out-migration is shock induced (i.e. 5.8% loss over 10 years), this would suggest 4% population loss per shock. We therefore take 4% loss per shock as our extreme

[30]Figures from World Bank Development Indicators, 1990-2000.

upper bound.

Table D.1: Potential Loss in Rural Populations due to Shock-induced Migration

Out-migration Per Shock –>	1%	**2%**	3%	**4%**	5%
Shocks over 10 yrs	10-yr population loss				
0	0.00%	0.00%	0.00%	0.00%	0.00%
1	1.00%	2.00%	3.00%	4.00%	5.00%
2	1.99%	3.96%	5.91%	7.84%	9.75%
3	2.97%	5.88%	8.73%	11.53%	14.26%
4	3.94%	7.76%	11.47%	15.07%	18.55%
5	4.90%	9.61%	14.13%	18.46%	22.62%
6	5.85%	11.42%	16.70%	21.72%	26.49%
7	6.79%	13.19%	19.20%	24.86%	30.17%
Estimate of **10-yr reduction** in population based on number of shocks observed in our data	1.46%	**2.91%**	4.33%	**5.74%**	7.13%

Each cell represents the ten-year population loss in a cluster that has occurrences of shocks as given by the row, and population loss *per shock* as given by the column. The last row represents the assumed *total 10-year loss* from the rural sample as a whole based on the shocks observed in the data. The highlighted columns best match our rural-to-urban migration estimates that (col. 2) rural areas lose approximately 2.9% of population over the course of ten years due to drought-induced migration and (col. 4) villages lose 5.8% of population over the course of ten years due to total (all-cause) migration.

E Considering shock timing

In this appendix, we consider whether shocks occurring within the past ten years differ in their impacts on the HIV epidemic according to whether they occurred relatively early or late during that period. We begin by simulating a model of the epidemic and then observe the impact of simulated shocks at various points in time.

Simulating the epidemic

We simulate an epidemic broadly representative of the high prevalence countries in our sample in the following way. In 1950, one person is infected in a country with a population of 25 million.[31] Each year, the newly infected individuals infect, on average, 1.2 other individuals (Pinkerton, 2008). Those infected more than 1 year ago infect another with an annual probability of 0.1 (Pinkerton, 2008). Each cohort of new infections dies at a rate that is specific to the years since their seroconversion (Fig. E.1).

At the end of each year, the number of infections is given by the number of infections at the end of the previous year, minus the deaths in the current year (from previous cohorts), plus the year's cohort of new infections.

In 1990, when prevention efforts began *en masse*, the annualised probability of transmission drops to 0.6 for individuals with an acute infection, and 0.04 for individuals with a chronic infection (>1 yrs since seroconversion). The "current year" (or year of observation) is 2005, which is the mean and median of our data years. The simulated epidemic is shown from 1970 to 2005 by the black line (labeled "Model") in Fig. E.2. Notice the dashed line at 1995; this serves as visual reference for comparing the post-1995 trend to the graphs of country prevalence in Fig. A.2.

When the simulation ends in 2005, there are 3.8m People Living with HIV (PLWH), yielding a prevalence of 16.5%. Of these, the number that were infected in each of the previous 18 years is calculated (note that none survives year 19 in the model). The CDF in Fig. E.3 shows that for PLWH, more than 80%

[31] The first documented case of HIV in SSA was in 1959, suggesting that the virus mutated between 1910 and 1950 (Zhu et al, 1998; Worobey et al, 2008). Our survey countries' average population is 22 million.

were infected in the past 10 years. One might consider extending the time period of analysis to 18 years to capture 100% of infections, but notice that for those infected 11-18 years ago, only 20% are currently alive (see Fig. E.1).

Role of shock timing

New infections in each year are a function of new infections in previous years. Therefore, a shock that increases incidence in 1995 will indirectly increase incidence in each of the following 10 years (when observing in 2005). In contrast, a shock in 2000 will only affect the incidence of the following 5 years until 2005. Shocks further in the past should have greater impacts on current prevalence (conditional on the shock being within the past 10 years).

The red and blue lines in Fig. E.2 depict simulations of shocks to incidence that occur 8 years prior ("early") and 2 years prior ("late") to the end-point of 2005, respectively. In each shock, the annual transmission probabilities increase to 0.7 and 0.04 for acute and chronic infections, respectively, during the year of the shock and the following year. It is clear that the path of prevalence between 1996 and 2005 differs significantly in the two scenarios. Given a shock of the same size and duration, the earlier shock increases 2005 prevalence by 50% more than does the later shock (1.7 percentage points vs. 1.1 ppts).

This suggests that shocks occurring 6 to 10 years ago will exhibit a stronger impact on current prevalence than will shocks that have occurred in the past 5 years.

Figure E.1: Survival Following Seroconversion (East African population without ARV)

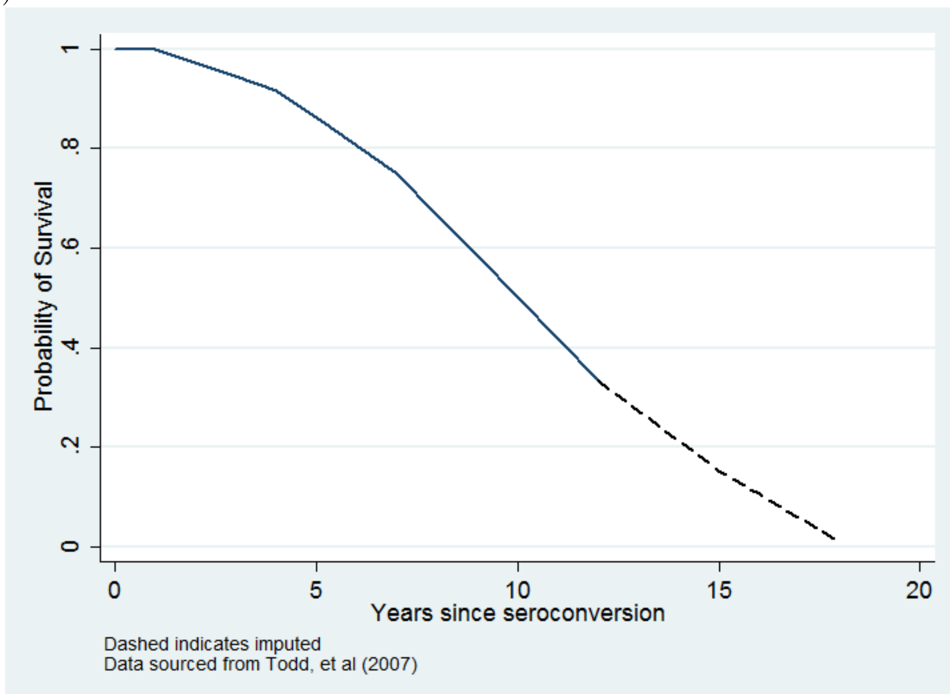


Figure E.2: Epidemic Curve

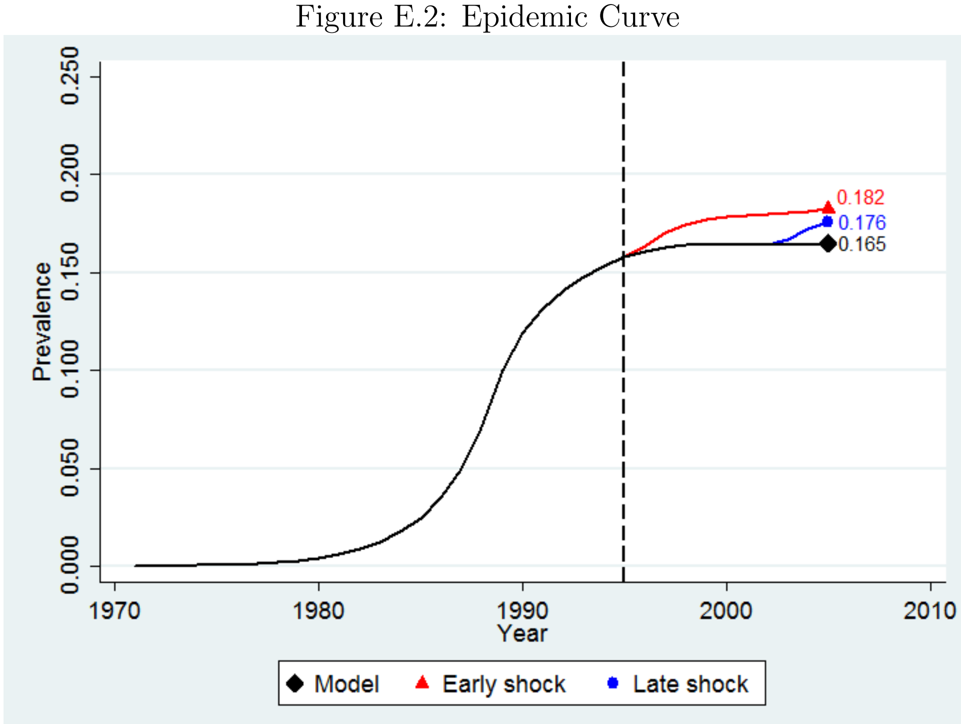
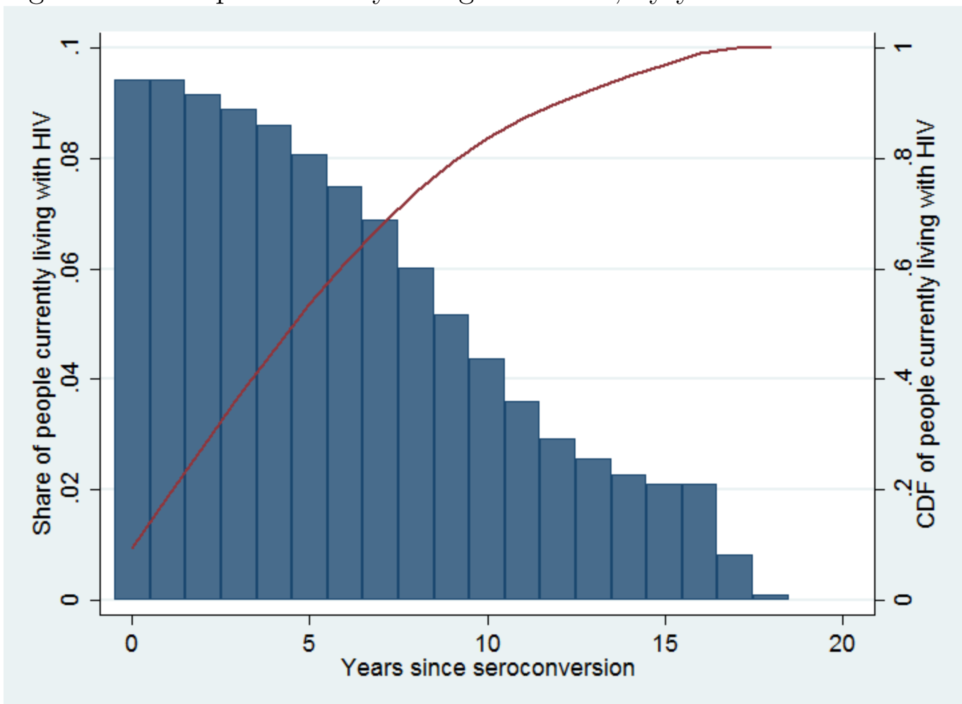


Figure E.3: People currently living with HIV, by year of seroconversion

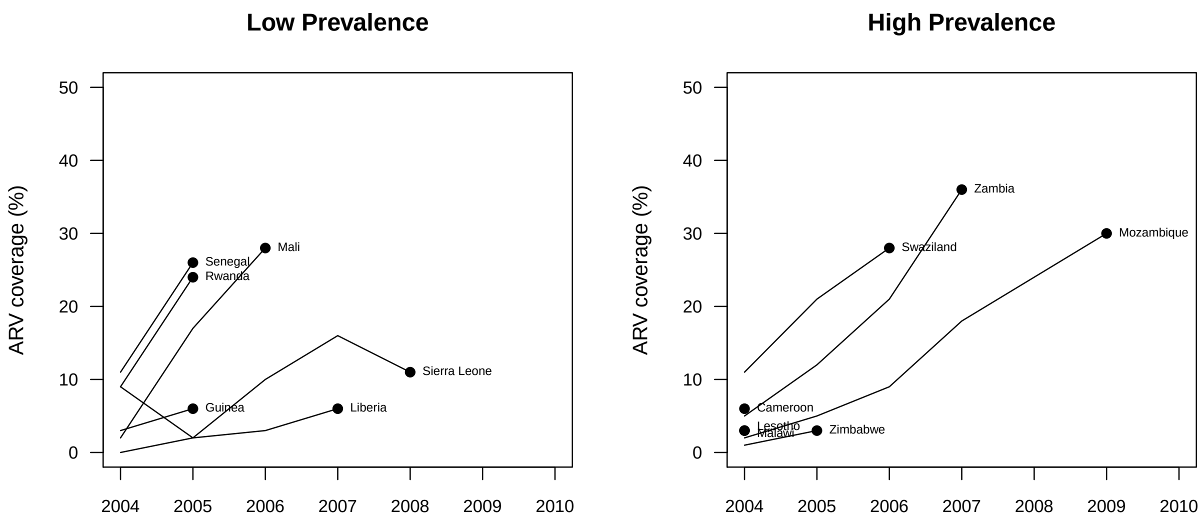


F The role of ARV Access

The number of people in sub-Saharan Africa being treated with ARVs has increased dramatically during the 2000 to 2009 period. This is due to both increased funding by PEPFAR and the Global Fund as well as substantial reductions in procurement costs (WHO, 2006; Friedman, 2012). Figure F.4 presents ARV coverage rates on a country level for when data is first available (2004) up until the year of the DHS survey. ARV coverage rates (the percentage of those receiving ARVs who are in need) are below 40% for all countries in our sample time frame. In rural areas, where we identify the effects that shocks have on HIV rates, ARV coverage rates are even lower. Access to ARVs is more limited for individuals living in rural areas due to both resource constraints and fewer trained medical professionals (van Dijk et al., 2009), as well as the greater distances that rural individuals may have to travel to access ARVs at clinics (Ojikutu, 2007). A number of country-specific studies substantiate this claim. In Kenya, ARVs were first targeted at urban areas and regions with high HIV prevalence (Friedman, 2012), longer travel times make it more difficult to access ARVs in rural areas in Zambia (van Dijk et al., 2009), and overall access to health services (including ARV access) is much more limited in rural than urban Mozambique (Groh et al., 2011).

One concern might be that shocks are correlated with ARV access which could lead to a different interpretation of our main results. For example, if ARVs are more readily available in areas with shocks then our results might be explained by more HIV+ individuals living longer in areas with shocks. Unfortunately subnational data on ARV availability is not available for our sample. The DHS does ask "Have you heard of drugs to help infected people to live longer?" in select countries, which we use as a crude proxy for ARV accessibility. We find that our shock measure is not correlated with ARV awareness which suggests that ARV access and shocks are not correlated (Table F.1).

Figure F.4: ARV Coverage Rates (2004-2009)



ARV Coverage rates from the World Bank Development Indicators. Black dots indicate the year of the DHS survey used for each country.

Table F.1: ARV Awareness and Shocks

	(1) All	(2) Rural	(3) Rural& High Prevalence
Num. shocks past 10 yrs.	-.002	-.001	.007
	(.006)	(.006)	(.006)
Observations	89208	55336	43082
R^2	.304	.337	.101
Mean of Dep. Var	.658	.633	.819

Sample includes: Congo DR, Liberia, Malawi, Mozambique, Sierra Leone, Swaziland, Tanzania (2008), Zambia, and Zimbabwe. Column headers indicate sample employed. Specifications include controls for gender and age, rural/urban designation (where applicable), and survey fixed effects. Estimations are weighted to be representative of the 19 countries. Robust standard errors are shown in parentheses clustered at the grid level.

G Estimating Changes in Sexual Behavior

The epidemiological model used to estimate changes in sexual behavior uses a similar technique employed by Gong (2012) which does a simple transformation of the AVERT model (Rehle et al., 1998). The model is expressed as:

$$M = \frac{log(1 - \mathbb{P}(HIV\ Infection))}{log(W[1 - R(1 - FE))^N + (1 - W))} \tag{3}$$

where $\mathbb{P}(HIV\ Infection)$ is the likelihood of HIV infection, W =HIV prevalence, R = HIV transmission per unprotected coital act, F =fraction of sexual acts where a condom is used, E = effectiveness of condoms at reducing HIV transmission, N = Number of sex acts per partner, and M= Number of sexual partners. Parameter estimates are in Table G.1.

One of the key parameters is the HIV transmission rate, or the likelihood of HIV infection per unprotected coital act. The major factor in determining this is the HIV infected partner's stage of infection. There are three stages of infection with the acute stage lasting approximately six months after the initial infection occurs, and the asymptomatic stage lasting approximately 8.5 years, until the onset of AIDS which marks the late stage. The HIV transmission rate in the asymptomatic phase is estimated to be approximated .07% (Boily et al., 2009; Powers et al., 2008), while the acute phase amplifies the transmission rate 26-fold (Hollingsworth et al., 2008). We use a single HIV transmission rate, by taking a weighted average of these HIV transmission rates. We estimate that 8% of those infected with HIV are in the acute phase, while the remainder are in the asymptomatic phase. Those with late stage infections are no longer sexually active (Powers et al., 2011). Finally, the condom parameters (F and E) come either from the DHS data or the epidemiological literature, while the number of sexual acts per partner assumes a one year duration for each partnership with 65 sexual acts per year (Powers et al., 2011).[32] Confidence intervals are calculated by drawing epidemiological parameter values for HIV-prevalence and HIV-transmission following (Powers et al., 2011) and estimating the number of partners (M) 1000 times.

[32] Powers et al. (2011) assumes 5 sexual acts per month, we increase this assumption slightly to 8.33 acts per month.

Table G.1: Parameter Values

Parameter	Value	Source
W (Prevalence Men) in Rural Areas in High Prevalence Countries	5.65%	DHS
W (Prevalence Women) in Rural Areas in High Prevalence Countries	8.26%	DHS
R (HIV Transmission)	.20%	(Powers et al., 2008; Boily et al., 2009; Hollingsworth et al., 2008; Magruder, 2011)
F (Fraction of Acts Condom Used)	11%	DHS
E (Condom Effectiveness)	80%	(Weller and Davis, 2002)
N (Sex Acts per Partner)	65	(Powers et al., 2011)

H Country-level prevalence

To explore the relevance of shocks for the broader patterns of HIV prevalence across Sub-Saharan Africa, we run simple cross-sectional regressions relating prevalence at the end of each decade to accumulated shocks over the previous decade. Country-level HIV prevalence is estimated as a functions of number of shocks over previous 10 years for the 38 countries in Sub-Saharan Africa with HIV data.[33] Details on the AIDS data are provided in UNAIDS (2010). Shocks are the sum of rainfall realizations below the 15th percentile, based on annual country-average rainfall (weighted by crop area). Country HIV prevalence data are from UNAIDS.

Table H.1: Shocks predict country-level HIV prevalence

	(1) levels 1990s	(2) levels 2000s	(3) change 1990s	(4) change 2000s
Num. shocks in past 10 yrs	2.089**	2.450***	1.250	0.408*
	(0.887)	(0.788)	(0.798)	(0.234)
Observations	37	37	36	37
R^2	0.140	0.216	0.064	0.068
Mean dep. var.	7.0	6.3	4.6	-0.7

Regressions marked “levels” have HIV prevalence in either 1999 (model 1) or 2008 (model 2) as the dependent variable; Regressions marked “changes” have as the dependent variable the change in HIV prevalence over the previous decade, with the end year either 1999 (model 3) or 2008 (model 4). Regressions include the 38 Sub-Saharan African countries with data in the UNAIDS database.

[33]The countries included in these regressions are: Angola, Benin, Botswana, Burkina Faso, Burundi, Cameroon, Central African Republic, Chad, Congo, Côte d'Ivoire, Eritrea, Gabon, Gambia, Ghana, Guinea, Guinea-Bissau, Kenya, Lesotho, Liberia, Madagascar, Malawi, Mali, Mauritania, Mozambique, Namibia, Niger, Nigeria, Rwanda, Senegal, Sierra Leone, South Africa, Swaziland, Togo, Uganda, United Republic of Tanzania, Zambia, and Zimbabwe.