A Social Vaccine? HIV Infection, Fertility, and the Non-Pecuniary Returns to Secondary Schooling in Botswana*

Jacob Bor
Boston University

Jan-Walter De Neve
Harvard University

NEUDC 2014 SUBMISSION - DO NOT CITE OR DISTRIBUTE

August 7, 2014

Abstract

Education has been hailed as a "social vaccine" against HIV infection; but there is little causal evidence to support this claim. A 1996 policy reform in Botswana changed the grade structure of secondary school and led to sharp increases in educational attainment among affected birth cohorts. We use this 'natural experiment' to identify the effect of secondary schooling on HIV infection risk, fertility, sexual behaviors, and labor market outcomes. Data were obtained from the 2004 and 2008 Botswana AIDS Impact Surveys, nationally-representative household surveys with HIV biomarker collection. Each additional year of secondary schooling induced by the policy change decreased the probability of HIV infection by 8.1 percentage points (se=3.1), relative to a baseline prevalence of 25.6%. Effects were particularly large among women, who also saw a 15.8% point (se=5.7) reduction in the probability of having ever given birth. Schooling had no effect on HIV knowledge; however it influenced norms and behaviors, increasing condom use, HIV testing, and reporting that it is acceptable for women to carry condoms. For women, education delayed sexual debut and increased labor force participation. For men, education increased number of partners, but also increased literacy, and discussion about HIV with others. Supply-side measures to expand access to education in developing countries may have large health benefits. Estimates of the returns to schooling that exclude these non-pecuniary benefits may be too low. *JEL Codes: I1, I2, J12, O15*

---

* Correspondence to: jbor@bu.edu, 801 Massachusetts Ave, 3rd Floor, Boston, 02118. Thank you to Gunther Fink, S (Subu) V Subramanian, and Sikhulile Moyo for feedback on this project. Thank you to the field and data management staff of the Botswana AIDS Impact Surveys and to the respondents who gave their time to this research. This work was supported by the Takemi Program in International Health at the Harvard School of Public Health, Belgian American Educational Foundation, and Fernand Lazard Foundation. The sponsors of the study had no role in study design, data collection, analysis, and interpretation, or writing of the report.

I. Introduction

HIV continues to be a major global health challenge with an estimated 2.3 Million new infections each year. [1] Formal education, particularly of girls, has been hailed as a 'social vaccine' to reduce the spread of HIV. [2] However, there is little causal evidence for this claim. [3] Existing cross-sectional and longitudinal studies have found conflicting evidence on the association between education and HIV risk. Early national surveillance surveys found higher rates of HIV among people with more education in a number of sub-Saharan Africa countries. [4-6] However, other studies have found a protective association between higher education and HIV infection, particularly as the epidemic has matured and information on prevention strategies has become more widely available. [7-9]

The effect of education on HIV infection is theoretically ambiguous. Education may reduce HIV risk through: increased exposure to information about HIV and prevention methods [10,11]; improved cognitive skills to make complex decisions regarding HIV risk reduction [12]; higher returns to market labor [13-16], increasing financial independence of women, reducing participation in transactional sex, and increasing bargaining power within relationships [17-19]; assortative mating with lower-risk partners [17,20,21]; less time for sexual relationships leading to increased abstinence during years in school [22,23]; changing fertility preferences, leading to increased contraceptive use and/or less frequent sex [24-28]; and increased future orientation. On the other hand, education may increase the size of one's sexual network; prolong the period of pre-marital sex [29]; increase earnings, enabling men to have more partners and/or engage in riskier sex [30]; and increase one's attractiveness in the "market" for sexual partners, leading (possibly) to more opportunities for unprotected sex.

This paper contributes to a larger literature on the health impacts of education in general. Due to the paucity of randomized trials and natural experiments, there is little consensus in the economics literature on whether a causal relationship exists between education and health, and if so what the mechanisms are. [12,31-36] In addition to its significance for HIV prevention, our study thus has implications for understanding the role of education in the health production function, and its effect on disparities in health outcomes across populations. [37-42]

The challenge in determining the causal effect of schooling on HIV infection risk is that school attainment is closely related to factors such as socioeconomic status, psychological traits, and preferences, which are difficult to control for fully in observational studies and which may also affect HIV risk. Several randomized trials have sought to identify the impact of schooling on HIV risk, but they have been underpowered to look at HIV incidence and/or have been paired with other interventions that make it difficult to attribute any effects to schooling. [43-45] For example, a randomized trial in Zimbabwe provided orphan girls with "comprehensive school support," including school fees, uniforms and a community visitor who monitored school attendance. The intervention prevented school dropout, early marriage, and reduced risk factors associated with infection with HIV. [45] A trial in Kenya randomized school uniforms, which led to an increase in schooling and a decrease in pregnancy; however the study was underpowered to detect changes in HIV incidence and the uniforms may have had an income effect. [44] In a trial in Malawi, girls were randomized to receive cash transfers conditional on school attendance, cash transfers alone, or nothing. Although the cash transfer reduced HIV risk, the authors could not reject the null hypothesis that the schooling condition had no effect on HIV incidence. [17] (Other RCT's are underway to assess the effect of conditional cash transfers for school attendance or achievement on HIV risk, but these studies are not designed to be able to separate the effects of cash from the effects of the condition since they do not have an unconditional cash treatment arm. [46,47]) A challenge for schooling experiments is that the demand-side subsidies and incentives offered in RCTs have generally led to small changes in educational attainment, with the result that their power to identify an effect on HIV infection is low. In contrast, supply-side policy reforms have, in some cases, led to large changes in educational attainment [48-50], and thus suggest a promising approach to assess the causal effect of schooling on HIV risk.

This study exploits plausibly-exogenous variation in educational attainment generated by a policy reform. In 1996, the Government of Botswana changed the grade structure of secondary school nationwide, moving grade 10 from senior secondary school to junior secondary school, dramatically increasing the number of seats available. The result was a sharp increase in average years of schooling by 0.8 years at the population level. The policy change affected specific birth cohorts – e.g., those who would have entered grade ten in 1996 or later – and as a sector-specific supply-side reform was unlikely to have affected HIV risk through mechanisms other than schooling itself. Using multiple survey waves to disentangle age and cohort effects, we use variation in exposure to the reform to identify the causal effect of education on the cumulative risk of HIV infection, fertility, and potential mediating risk factors, including labor force participation and sexual behaviors.

The paper proceeds as follows: Section II describes the policy reform and study context. Section III describes the data used in the analysis. Section IV presents the empirical approach. Section V presents the main results, with robustness checks in Section VI. In Section VII, we explore mediating pathways from education to HIV risk. Section VIII evaluates the cost-effectiveness of secondary schooling as an HIV prevention intervention. Section IX concludes.

II. Policy Reform and Study Context

Botswana is a sparsely-populated, land-locked country in southern Africa, whose economy depends substantially on diamond-mining. Botswana has among the highest rates of HIV in the world, with 25.6% of adults aged 15-49 years infected in 2008 (BAIS 2008).

Botswana's K12 public education system is divided into primary, junior secondary and senior secondary education In 1994, the National Commission on Education (NCE) brought up several problems associated with the existing '7+2+3' grade structure, in which two years of junior and three years of senior secondary school followed seven years of primary education. In particular, two years of junior secondary education was insufficient to prepare students for work or further training and did not offer sufficient time for students to adjust from primary to secondary school. The NCE recommended a switch to a '7+3+2' structure with primary and junior secondary school forming 'Ten Years of Basic Education' (rather than 'Nine Years of Basic Education'). Botswana's public education system is strongly centralized and the policy was implemented rapidly and universally; it also serves the vast majority of the population with less than 1% of secondary school students attending private schools (Government of Botswana 2013). [51]

In January 1996, Botswana shifted the tenth year of education from senior secondary to junior secondary school, with the goal of increasing access to grade ten. [52] The reform may have influenced educational attainment through multiple channels. First, the reform led to a large increase in the supply of grade ten education. There are about nine junior secondary schools for every senior secondary school in Botswana (CSO 2010); moving grade ten to junior secondary school thus increased the number of seats and reduced travel time for students. Second, it increased the number of years of schooling required to obtain a Junior Certificate, raising the benefits of completing grade ten and establishing continuity with grade nine. Third, if education affects preferences for later schooling (e.g., some students may discover that they like school), then increasing grade ten completion could increase progression through later secondary and even tertiary education. A previous study has shown that the reform led to an increase in total years of schooling – with a large increase in grade ten completion.[48] This 'natural experiment' provides an opportunity to estimate the causal impact of schooling on risk of HIV infection, by assessing whether there is decreased risk of HIV infection among birth cohorts exposed to the reform.

III. Data Description

III.A Study Population

The study population included all women and men living in Botswana, at least 18 years old at the time of the surveys, and who were citizens of Botswana born in or after 1975. Respondents younger than 18 years old at the time of the surveys were excluded because they would not have had the opportunity to complete secondary education. Multiple previous school reforms led to rapid changes in the distribution of schooling for cohorts born before 1975, so they were excluded from the analysis, [52] resulting in a study population ages 18 to 32 years. Immigrants to

Botswana were excluded because they would not have been exposed to the schooling intervention if they migrated in adulthood.

III.B Sampling Strategy

Data were obtained from the Botswana AIDS Impact Surveys (BAIS) II (2004) and III (2008), nationally representative population-based household surveys with HIV biomarker collection. BAIS II and III each employed a two-stage probability sample design stratified according to district and major urban centers. For each survey, a representative probability sample of approximately 8,300 households was selected. All household members aged 10-64 who were usual members of the household, and spent the last night in the household, were eligible to be interviewed. For survey year 2004, 8,206 female and 6,656 male respondents were interviewed; in 2008, 7,497 female and 6,055 male respondents were interviewed. Data on demographics, HIV biomarkers, and self-reported sexual behaviors were collected for all respondents. Detailed sampling plans and HIV testing procedures are available from survey final reports. [53,54] Household and individual participation rates were, respectively, 92% and 93% for survey year 2004, and 87% and 82% for survey year 2008. HIV test participation rates were 61% for survey year 2004, and 67% for survey year 2008. A small number of HIV tests (< 2%) produced invalid results. Data on years of schooling completed were available for 99.7% of respondents with an HIV test result.

III.C Measurement of Exposures and Endpoints

Data on educational attainment, age in years, HIV status, fertility, marriage, sexual and HIV testing behaviors, HIV knowledge and attitudes, employment, literacy, gender, citizenship, district of birth, household number, respondent line number, and sample weights were extracted from the BAIS datasets. Year of birth was calculated as Survey Year *minus* 0.5 *minus* Age at Survey. The key exposure in our analysis was the total number of years of schooling completed. As a source of exogenous variation in schooling, we defined an indicator for whether or not a respondent was exposed to the 1996 education policy reform. Historically in Botswana, children were eligible to begin primary school in the year of their seventh birthday. Presuming they would progress on time through school, without repeating or skipping grades, they would enter 10th grade in the year of their 16th birthday. We defined exposure to the reform ("reform cohort") as one if the respondent was aged 16 years or younger in 1996 (born in or after 1981), and zero otherwise.

Our primary endpoint was HIV status, which reflects the cumulative probability that a respondent acquired HIV up to his or her age at the time of the BAIS surveys. As secondary endpoints, we assessed the causal effect of schooling on a range of potential mediating pathways including: age at first intercourse, an important predictor of HIV risk, the number of previous births, marriage, years of premarital sex, current sexual and HIV testing behaviors, HIV knowledge and attitudes, literacy, and labor market outcomes. [55] The text of the specific questions is provided in the Appendix.

IV. Empirical Approach

Because high-risk individuals may self-select for higher (or lower) educational attainment based on unobserved characteristics (confounders), naïve bivariate and covariate-adjusted associations between years of schooling and HIV status should be interpreted as "descriptive," and may not reflect a causal relationship. [4-7,9,11] To obtain causal effects, we exploited exogenous variation in educational attainment resulting from a 1996 policy reform that changed the grade structure of secondary school in Botswana. Using exposure to the reform as an instrumental variable, we identified the effect of additional years of schooling on the cumulative probability of HIV infection.

As a benchmark against which to compare our instrumental variables estimates, we assessed the association between years of schooling and the cumulative probability of HIV infection in descriptive, multivariate OLS (linear probability) regression models. [56] We estimated several specifications, modeling years of schooling as a continuous covariate; allowing for different slopes for 0-9 years and 10-13+ years of schooling; and modeling each additional year of schooling non-parametrically, with separate indicators for additional years of schooling completed. Although logistic regression is typically used for binary outcomes, instrumental variables models with a logistic second stage are not consistent; the coefficients in these descriptive linear probability models can be compared directly with the coefficients in our causal instrumental variable models, as described below. [57]

Our IV analysis proceeded in three steps. First, we assessed whether birth cohorts exposed to the reform ("reform cohorts") had higher educational attainment than birth cohorts not exposed to the reform. We estimated the effect of exposure to the reform on total years of schooling completed in multivariate OLS regression models ("first stage"), adjusting for covariates. We also assessed the effects of the reform on the distribution of years of schooling completed, assessing effects on the probabilities of completing at least 7, 8, 9, 10, 11, 12, and 13+ years of schooling. Second, we assessed the "intention-to-treat" (ITT) (a.k.a. "reduced form") effect of being in a reform cohort on the cumulative probability of HIV infection. ITT estimates were assessed in multivariate linear probability models, adjusting for covariates. Third, similar to an RCT with non-compliance [58], complier causal effects can be obtained by dividing the ITT by the First Stage. To obtain such a causal estimate, while adjusting for covariates, we estimated two-stage least squares (2SLS) regression models, using exposure to the reform as an instrument for total years of schooling.

In all models – both descriptive and causal – we controlled flexibly for age with a full set of single-year age indicators, to account for the non-monotonic pattern of HIV infection across ages in Botswana. [59] We included indicators for survey year (2008 vs. 2004) and for district of birth. Finally, we adjusted for a continuous linear term in year of birth, to account for continuous trends in education or HIV infection risk across birth cohorts. Exposure to the reform was modeled as an intercept shift for cohorts born in or after 1981. We estimated all models first for women and men separately, and then on the pooled sample. When pooling sexes, we additionally included indicators for sex and the interactions of sex with the other covariates**.** In the pooled 2SLS regressions, we did not interact sex with years of schooling, so that the coefficient on schooling reflects a weighted average of the effects for men and women. Identical models were estimated for all secondary outcomes. All IV estimates are interpreted as "local" to the subpopulation who "complied" with their treatment assignment – e.g., persons who increased their years of schooling because of the reform. [60]

Figure 1 displays a causal diagram illustrating the assumptions underpinning our study. Four assumptions are required. First, the instrument (Z) must have had an effect on schooling (E), e.g., a valid first stage; this is testable and we find large effects. Second, the instrument (Z) must be independent of unobserved confounders (U), conditional on observed covariates (X); in our application this implies that people born before and after 1981 were similar, after controlling flexibly for age, survey year, district of birth, and a linear trend in HIV risk across birth cohorts. In robustness checks, we included quadratic terms for year of birth, reduced the window of observation to a narrower set of birth cohorts, and allowed the slope of the trend across birth cohorts to differ before and after 1981. Identification comes from the fact that the policy reform led to a discontinuous change in schooling across cohorts born before and after 1981; it is unlikely that other unobserved factors would have led to a discontinuous change in HIV risk for precisely those cohorts affected by the reform. Importantly, the availability of two survey years enables us to identify these cohort effects, while controlling flexibly for age and period effects. Third, we assume that exposure to the policy reform (Z) affected HIV risk (Y) only through changes in schooling (E) (the exclusion restriction); this is highly plausible given that the reform was a supply-side intervention that would not have specifically affected the reform cohorts except through their increased access to grade 10. Fourth, to interpret our results as complier causal effects (a.k.a. local average treatment effects), we must assume monotonicity; e.g., that exposure to the reform (Z) only caused individuals to obtain more schooling or to have no change in schooling, and did not lead some individuals to obtain less schooling. Violations of this assumption are possible but unlikely (e.g., a person with a very strong preference for small class size might have continued to grade ten pre-reform but dropped out after grade nine post-reform). [61,62]

Stata (version 12.0, StataCorp, College Station, Texas) was used for all statistical analyses.

V. Results

V.A Sample Description

The 2004 and 2008 BAIS surveys included 3,965 women and 3,053 men with valid HIV biomarkers, for a total of 7,018 respondents. Figure S1 in the Appendix shows a participant flow diagram. Table 1 shows summary statistics. Mean age was 22.7 (SD 3.1) for women and 22.6 (SD 3.2) for men in the BAIS II survey and 24.9 (SD 4.2) for women and 24.7 (SD 4.3) for men in the BAIS III survey. Mean years of schooling was similar for men and women, at about ten years. Age at first intercourse was 18.0 (SD 2.0) years for women and 17.8 (SD 2.5) years for men in the BAIS II survey, and 18.2 (SD 2.5) for women and 18.5 (SD 3.0) for men in the BAIS III survey. 28% of women and 11% of men in the BAIS II study sample and 27% of women and 12% of men in the BAIS III study sample were HIV positive.

V.B Descriptive Association between Education and HIV Infection Risk

Figure 2 shows the crude associations between education and HIV infection risk for men and women. The relationship between HIV infection risk and schooling appears non-monotonic. HIV risk peaks for persons completing 8 - 9 years of education, and declines sharply after nine years of schooling. Table S1 in the Appendix shows the covariate-adjusted OLS association between years of schooling and HIV infection risk. On average, one additional year of schooling is associated with 1 - 2 percentage points lower HIV risk. However, each additional year of schooling up to nine years was associated with a 0.3% point *higher* risk for HIV infection ($p = 0.229$); by contrast, each additional year of schooling above nine years of schooling was associated with a -3.6% point lower risk of HIV infection ($p < 0.0001$). Since these associations may be confounded by unobserved characteristics, we used an instrumental variables approach to obtain causal effect estimates.

V.C Effect of the 1996 Grade Reform on Years of Schooling Completed

Figure 3 shows the proportion of respondents who completed at least 7, 8, 9, 10, 11, or 12 years of schooling and how this distribution changed across birth cohorts. The fraction of students completing at least 7, 8, or 9 years of schooling rose gradually and continuously across birth cohorts. However, the share of students with at least ten years of schooling increased sharply for the birth cohorts affected by the reform. Modest increases in completion of 11 and 12 years of schooling were also observed for the reform cohorts. Tables 2 and S2 in the Appendix displays regression estimates of the impact of the reform on educational attainment. The reform induced an increase of 0.79 years of schooling ($p < 0.0001$) - 0.64 for women ($p = 0.004$) and 1.01 for men ($p = 0.002$).

V.D Causal Effect of Education on HIV Infection Risk

Table 2 presents “intention-to-treat” results, in which HIV status was regressed directly on the instrument and covariates. Women who were exposed to the reform period were 7.4 percentage points less likely to be HIV positive ($p = 0.017$); men were 5.0 percentage points less likely to be HIV positive ($p = 0.052$). The pooled coefficient was 6.4 percentage points ($p = 0.002$). Figure 4 plots the predicted proportion of HIV positive by birth cohort, with and without the reform, superimposed over the observed HIV prevalence. Observed HIV prevalence closely matched the model predictions. For the cohorts exposed to the reform (1981-1985), HIV prevalence was lower than it was predicted to be in the absence of the reform. Table 2 shows 2SLS results for the effect of additional years of education on HIV status. These instrumental variable estimates show that additional years of schooling induced by the reform had a protective effect against HIV infection. Batswana who stayed in school for an additional year had an 8.1 percentage point lower risk of HIV infection ($p = 0.008$) – 11.6 percentage points for women ($p = 0.045$) and 5.0 percentage points for men ($p = 0.085$). Since the reform increased completion of grades ten and above, we interpret the instrumental variable estimates as the causal effect of additional secondary school education on HIV risk, for the population that would have otherwise dropped out of school after grade nine. For men, the effect of education appeared to be protective of HIV risk, but with smaller samples, this result did not reach conventional benchmarks for statistical significance.

VI. Sensitivity Analyses and Robustness Checks

We assessed the robustness of our results to the presence of non-linearities in long-run cohort trends in HIV risk, by controlling for a quadratic in year of birth in addition to the linear term included in the main analysis, and find almost identical results (Table 3, columns 4, 5). The assumption that underlying trends are approximately linear is also more plausible the narrower the window of cohorts included. Using narrower window of birth cohorts – 1981 +/- 4, 5 years, results remain similar, although the estimates are less precise (Table 3, columns 7, 8). HIV consent rates were incomplete and it is possible that selective non-response may have biased our estimates. However, exposure to the reform was not significantly associated with consent rates (Table S6). Further, imputing for HIV status based on observables, our results were essentially unchanged (Table 3, columns 10, 11). To rule out the possibility that other national policy changes might have affected HIV risk for the same cohorts, we used an alternative identification strategy exploiting the fact that the reform would be expected to have the biggest impact in districts where a higher proportion of students completed exactly nine years of schooling in the pre-reform period. Interacting district (high vs. low 9^{th} grade completion) with the indicator for exposure to the reform, we implemented a difference-in-differences strategy and found similar results (Table 3, column 12). Further details on robustness checks are provided in the Appendix. Table S5 in the Appendix displays robustness checks for men and women separately, as a companion to Table 3 which presents combined results.

VII. Causal Pathways from Education to HIV Risk

To investigate the causal pathways from education to HIV risk, we assessed the effect of the reform on potential mediators, our secondary endpoints. Table 4 shows 2SLS results for the effect of additional years of schooling on age at first intercourse, marriage, fertility, sexual and

HIV testing behaviors, HIV knowledge and attitudes, literacy, and labor market outcomes for women and men separately. Schooling had no effect on HIV knowledge; however it influenced other HIV risk factors, HIV attitudes, literacy, and labor market outcomes. For women, education delayed sexual debut by 0.76 years ($p = 0.004$), increased labor force participation by 17.2 percentage points ($p = 0.025$), and reduced the proportion that had ever given birth by 16.0 percentage points ($p = 0.006$). For men, education increased the likelihood of having more than one partner by 11.1 percentage points ($p = 0.028$), but also increased literacy by 8.4 percentage points ($p = 0.001$), and discussion about HIV with others by 8.7 percentage points ($p = 0.056$). For both men and women, education increased condom use, HIV testing, and the proportion reporting that is acceptable for women to carry condoms. In interpreting these mediators, we caution that while sexual debut likely occurred prior to HIV infection, contemporary knowledge, attitudes, and behaviors were observed after all of the HIV-infected survey respondents had seroconverted, implying that the coefficients for secondary endpoints may capture behavioral responses downstream from HIV infection. However, for most secondary outcomes, these biases would be expected to run counter to our results: cohorts with higher HIV prevalence would be more likely to have ever tested for HIV, more likely to use condoms and to believe that it is acceptable for women to carry condoms, less likely to bear children, and more likely to talk about HIV.

VIII. Cost-Effectiveness of Secondary Education as an HIV Prevention Strategy

To assess the cost-effectiveness of secondary schooling as an HIV prevention intervention, we estimate the cost per disability-adjusted life year (DALY) averted using estimates of the per-pupil-per-year costs of secondary education published by the UNESCO Institute for Statistics and our own calculations of the treatment costs and DALYs associated with an HIV infection in Botswana. We also compare the cost per HIV infection averted due to secondary schooling vis-à-vis other proven HIV prevention interventions. Based on calculations presented in the Appendix, we estimate that an HIV infection at age 20 would lead to 16.3 lifetime DALYs for someone who did not initiate antiretroviral therapy (ART); and 3.5 lifetime DALYs for someone who initiated ART, with a lifetime treatment cost of $12,400; all costs and DALYs were discounted at

3%. These calculations imply cost-effectiveness ratios (CER) of $4,387/DALY with ART and $1,703/DALY without ART; each of these CERs is lower than Botswana's $5,178 per capita GDP (2009), implying that as an HIV prevention intervention secondary school is "very cost-effective" according to the standard benchmark of 1x per capita GDP. Table 5 compares the cost-effectiveness of secondary school with other proven HIV prevention interventions, such as medical male circumcision, treatment as prevention, and pre-exposure prophylaxis. Secondary schooling is more expensive than circumcision and treatment as prevention, but of similar cost-effectiveness to pre-exposure prophylaxis. [63-66] Importantly, unlike these other interventions, secondary schooling has large benefits beyond the reduction of HIV transmission – benefits that have been excluded from the above calculations.

IX. Conclusion

Using an education policy reform as a natural experiment, we find that secondary schooling has a large protective effect against risk of HIV infection in Botswana. Effects were particularly large for women and were consistent through a wide array of robustness checks. Our IV estimates are somewhat larger, but generally consistent with the strong negative associations we found between secondary schooling and HIV risk in OLS regressions. One explanation for why the 2SLS results are larger than the OLS results could be that unobserved factors, such as personal charisma, may be positively associated with both educational attainment and HIV risk, thereby reducing the magnitude of the OLS coefficient. Another explanation is measurement error, although this is unlikely that noise in reported years of schooling is large enough to account for the difference between OLS and IV coefficients. A third explanation is that – as with all IV estimates – the causal effects that we estimate are "local" to the subpopulation of compliers, e.g., those induced to increase schooling because of the reform. This subpopulation consists of persons who, in the absence of the reform, would have dropped out after ninth grade – a group likely to be at particularly high risk for HIV.

Our effect estimates for a single year of schooling are large; however, there is reason to believe that the later years of secondary school are particularly protective against HIV infection risk. The

OLS regressions provided suggestive evidence that the effect of schooling on HIV risk may be non-monotonic in Botswana: an additional year of schooling was associated with slightly *increased* HIV risk from years of schooling 0 to 9; but with large reductions in HIV risk in years 10-13+. Multiple countervailing pathways from education to HIV risk may be at work. Participation in late primary and early secondary school may increase social (and sexual) network size but confer little in the way of economic opportunity or cognitive skills to navigate a complex risk environment. Although we do not have a natural experiment for primary schooling, it is quite likely that results would differ substantially. The later grades of secondary school may be a critical exposure period in the determination of HIV risk and fertility decisions. [67]

Additional years of schooling had a causal effect on some proximate risk factors for HIV, but not others, providing insights on potential mechanisms. Education increased condom use and improved norms regarding women carrying condoms, suggesting more widespread adoption of this particular prevention technology among the better educated, a phenomenon suggested in other studies. [9,11,12] Education increased HIV testing for both men and women and led to increases in the proportion of men who reported talking with others about HIV, suggesting increased openness about HIV and demand for knowledge about one's own HIV status. Additional schooling led women (but not men) to delay sexual debut and to delay childbearing. The reduction in fertility is generally consistent with reduced unprotected sexual activity and lower risk for HIV, and is also of interest in its own right. Although education led to later sexual debut, we observed no change in entry into marriage, which in general occurs quite late in Botswana. In contrast to work by Case and Paxson, [29] we find the exogenous changes in education led to shorter (not longer) durations of pre-marital sexual activity, which may have reduced exposure to HIV.

Interestingly, education had no effect on abstinence (measured at the time of the surveys) or number of partners for women, and actually increased numbers of partners for men – a finding similar to other literature. [5,30] Although we have limited data on partner characteristics, we find no evidence that education reduced participation in age-disparate relationships – "sugar daddies/sugar mamas", thought to be a driver of HIV risk. [68-71] In summary, whereas education neither reduced sexual activity, nor reduced numbers of sex partners, nor lead people to select younger partners – it did appear to delay sex for women and to reduce HIV risk *within* relationships through increased awareness of HIV status, communication about HIV, and normalization of condom use. These findings contribute to our understanding of the margins on which Batswana with additional schooling successfully adapted to reduce exposure to HIV in a highly endemic setting.

What was it about education that led to these changes in behavior and reductions in HIV infection risk? Perhaps the most obvious hypothesis is that schooling might provide information about HIV that enables people to make safer decisions. In fact, gains in education induced by the reform had zero causal effect on HIV knowledge and misconceptions. Although counterintuitive, this result is not surprising given that the reform cohorts in our study completed secondary school in the 1990s and early 2000s, before Botswana launched a formal HIV curriculum in schools. Many resources have been devoted to HIV-specific education programs, and indeed HIV knowledge was observed to be high in Botswana. However, our results suggest that the effect of secondary schooling on HIV risk is not attributable to knowledge acquisition, but rather to other factors. Scholars have long argued that "knowledge is not enough" to prevent HIV infection. [72] Our results indicate that secondary education provides critical enabling factors that allow knowledge to be utilized and enable people to avoid HIV infection.

Although education had no effect on HIV-specific knowledge, secondary schooling did have large causal effects on other factors that may have mediated the effect of schooling on HIV risk. These pathways differed for men and women. For women, the additional schooling induced by the reform had very large effects on labor market participation – a consequence of the reform that has also been reported elsewhere. [48] In our data, the reform caused over half of those women who would have otherwise been out of the labor force to seek (and in many cases find) employment (per year of schooling induced by the reform). These effects are very large and suggest that the skills learned by women in grade ten are very important in local labor markets. The annual private rate of return for each schooling cycle generally increases by level of schooling in Botswana: rates of return in 1993/94 were the highest for upper (185%) and lower (83%) secondary education, and the lowest for primary education (7%), as evidenced elsewhere. [15,73] Indeed, whereas many of the manual labor jobs traditionally open to men require minimal schooling, labor market opportunities for women (e.g., as teachers, nurses, clerks, or the hospitality sector) typically require secondary education. [15] No significant effects on labor market participation were observed for men, and in fact, the influx of women into the labor market may have slightly increased unemployment for men in the same cohorts (we observed an increase in unemployment for men that was economically, but not statistically significant).

Given the local labor market context, the gains in schooling induced by the 1996 reform led to a large shift in women's labor supply from the home to the market, and led to a shift in labor market opportunities from men to women. These changes in economic opportunity may have enabled women to make explicit choices to reduce HIV risk, e.g. by increasing bargaining power for women within relationships. [74] Indeed we find evidence that education changed norms for both women and men about the acceptability of women carrying condoms, and led to increases in condom use and HIV testing. (As discussed above, we find no evidence that education led women to select less risky partners, at least with respect to age; however, we are limited in our ability to assess partner choice across other dimensions.) In addition to economic empowerment, schooling may have also led women to think differently about their future, changing expectations about whether they would have their own career or be a homemaker, and changing fertility preferences. Indeed, we observe a large delay in sexual debut and a marked reduction in childbearing for women affected by the reform. Interestingly, we observe a rightward shift in the distribution of age at sexual debut not just from 16 to 17, the age when most girls would have been in grade ten, but through age 20, suggesting that education did not just limit opportunities for sex among school-going youth, [23] but actually changed preferences towards later sexual debut (Table S8).

The effects of schooling on women's labor market participation and fertility have been observed in other settings without hyper-endemic HIV, [25-28,75,76] and schooling has been cited as a key policy lever in spurring the demographic transition from high to low fertility. [77-80] In Botswana, these changes may have had the added benefit of reducing HIV infection. Whether the observed reductions in HIV risk are unintentional consequences of increased utilization of condoms as birth control, or a result of explicit decisions by economically empowered and/or more forward-looking women to reduce their exposure to HIV, cannot be ascertained. However, the benefits of secondary school in reducing HIV risk appear to be closely related to effects of education on the economic and reproductive empowerment of women in Botswana.

Although the additional education induced by the reform had little economic impact for men, we find evidence that extra years of schooling may have had impacts on cognitive skills, as suggested by very large increases in literacy. The proportion of men reporting that they could not read and understand a letter or newspaper with ease dropped by nearly half per year of schooling induced by the reform. These skills may have improved men's ability to use information about HIV in making complex decisions in their lives. [12] Indeed we find evidence that men are more likely to discuss HIV with others and, similarly to women, to test for HIV, use condoms, and report that it is acceptable for women to carry condoms. Although knowledge about HIV prevention methods did not change, men may have acquired additional skills to utilize that knowledge in their lives.

In summary, we draw three conclusions from our discussion of mechanisms. First, education had no effect on HIV knowledge, but did affect norms and behaviors. Second, observed changes in behavior occurred on the margins of risk reduction within relationships and delayed sexual debut for women. We found no evidence of partner reduction, abstinence after debut, or selection of less risky partners as pathways. Third, education led to fundamental changes in market labor supply for women, which may have empowered women to reduce exposure to HIV due to increased bargaining power, future-orientation, or changed fertility preferences. The ability to chart out causal pathways in this degree of detail is rare and a testament to the data collected in the BAIS surveys. Finally, the fact that the 1996 reform affected a number of proximate HIV risk factors (e.g., condom use, sexual debut) and outcomes correlated with unprotected sex (e.g., fertility) generates added confidence in our main results.

Our study has some limitations. First, consent rates were imperfect, and migration or mortality could have influenced the composition of the study sample. However, neither consent rates nor birth cohort sizes varied systematically with exposure to the reform, suggesting that any bias from these sources is minimal. We also imputed HIV status for people without valid HIV test result and find similar results using the full sample. Second, our analysis relies on the assumption that HIV infection does not cause people to stay in secondary school or drop out. Infection rates are likely very low prior to grade ten, with only 10.1% of women and 14.9% of men having their sexual debut before age 16 (BAIS 2008). The vast majority of people infected with HIV would be asymptomatic during the period of their schooling making reverse causality unlikely. In addition, we focus on a cohort born between 1975 and 1989 whose childhood occurred prior to the advent of pediatric ART. Any children infected during birth or breastfeeding would almost certainly have died prior to secondary school and would not be in the sample. Third, we only observe people through age 32 years. We cannot know whether we are measuring HIV infections truly averted or delayed. However, this is a common limitation of prevention studies. In spite of this limitation, our analysis of cumulative incidence captures much longer follow-up than most RCTs, which observe incidence over a shorter, e.g., 3yr [46], horizon.

Fourth, by exploiting exogenous variation in schooling, we avoided issues of self-selection of high-risk individuals into more (or less) schooling, and thereby control for such unobserved confounders as: socioeconomic status, risk aversion, future-orientation, self-confidence, etc. However, our analysis nevertheless relies on the assumption that conditional on covariates – there were no other cohort-specific effects that would have affected HIV risk aside from exposure to the reform. There are many reasons why HIV risk might change across birth cohorts but the likely candidates – infection rates among sexual partners, access to HIV treatment, changes in prevention programming – are phenomena that are either gradual over time (changes in the epidemic context) or, if they are sudden, affect people of many different ages (e.g., a national prevention campaign): in both cases, these phenomena would result in gradual changes in HIV infection across birth cohorts. We control for such changes in risk across birth cohorts using a linear (or quadratic) trend, which picks up all observed and unobserved factors that change smoothly (linearly, quadratically) across birth cohorts. The validity of our natural experiment would be jeopardized only if there were unobserved factors that led to a discontinuous change in HIV risk for the specific birth cohorts affected by the policy reform. One example would be an HIV prevention program targeted to a specific grade in school. However, Botswana's HIV education programming was not in place in 1996. [81] In robustness checks, an alternative difference-in-differences identification strategy yielded similar results, lending support for this assumption. Finally, as with all infectious diseases, there could be spillover effects beyond the individuals directly affected by the reform, which would change the interpretation of our estimates; however, given that people have sexual relationships across birth cohorts, these spillovers would be expected to have a smooth impact on HIV infection across birth cohorts and would not bias our estimates, which are based on a discontinuous change in exposure to the reform.

Many studies have reported correlations between schooling and HIV infection risk. [4,6-9] This study is among the first to use a natural experiment to assess the causal relationship between schooling and HIV infection. [10,82] [†] A sharp policy change that increased access to secondary school enabled us to rule out unobserved confounders that may have biased previous estimates. Further research on the relationship between education and HIV risk could draw upon randomized trials or other natural experiments – including compulsory schooling laws, class size, and school admission lotteries – to tease out mechanisms and determine generalizability across other contexts. [49,83-85] One attractive feature of the policy change that we evaluate is the sheer size of the schooling gains that resulted. Botswana's 1996 education reform focused on supply-side factors – increasing the number of seats, reducing travel times to school, and making continuation to 10th grade the default option. These changes led to very large increases in educational attainment, particularly among those students most likely to drop out of school. To increase educational attainment in other settings, supply-side interventions that increase opportunities for secondary schooling should be considered alongside demand-side interventions such as conditional cash transfers.

Expanding access to secondary school had a large protective effect against HIV infection in Botswana. Our findings confirm what has been long suspected: that secondary schooling is an important structural determinant of HIV infection. In Botswana, and other settings with large, generalized HIV epidemics, estimates of the returns to secondary schooling may be grossly underestimated due to the exclusion of health benefits.

[†] The study by Agüero and Bharadwaj (2014) was underpowered to detect effects on HIV infection, while the study by Behrman (2014) did not adjust for age, which may confound their cohort-based identification strategy.

REFERENCES

1. Joint United Nations Programme on HIV/AIDS. Global report: UNAIDS report on the global AIDS epidemic 2013. Geneva, Switzerland; 2013.
2. Vandemoortele J, Delamonica E. The 'education vaccine' against HIV. *Current Issues in Comparative Education* 2000; **3**(1).
3. Jukes M, Simmons S, Bundy D. Education and vulnerability: the role of schools in protecting young women and girls from HIV in southern Africa. *Aids* 2008; **22 Suppl 4**: S41-56.
4. Dallabetta GA, Miotti PG, Chiphangwi JD, et al. High Socioeconomic-Status Is a Risk Factor for Human-Immunodeficiency-Virus Type-1 (Hiv-1) Infection but Not for Sexually-Transmitted Diseases in Women in Malawi - Implications for Hiv-1 Control. *J Infect Dis* 1993; **167**(1): 36-42.
5. Fortson JG. The gradient in sub-Saharan Africa: socioeconomic status and HIV/AIDS. *Demography* 2008; **45**(2): 303-22.
6. Kirunga CT, Ntozi JP. Socio-economic determinants of HIV serostatus: a study of Rakai District, Uganda. *Health transition review : the cultural, social, and behavioural determinants of health* 1997; **7 Suppl**: 175-88.
7. Barnighausen T, Hosegood V, Timaeus IM, Newell ML. The socioeconomic determinants of HIV incidence: evidence from a longitudinal, population-based study in rural South Africa. *Aids* 2007; **21 Suppl 7**: S29-38.
8. Fylkesnes K, Musonda RM, Sichone M, Ndhlovu Z, Tembo F, Monze M. Declining HIV prevalence and risk behaviours in Zambia: evidence from surveillance and population-based surveys. *Aids* 2001; **15**(7): 907-16.
9. Hargreaves JR, Bonell CP, Boler T, et al. Systematic review exploring time trends in the association between educational attainment and risk of HIV infection in sub-Saharan Africa. *Aids* 2008; **22**(3): 403-14.
10. Agüero JB, P. Do the More Educated Know More about Health? Evidence from Schooling and HIV Knowledge in Zimbabwe. *Econ Devel Cult Change* 2014; **62**(3).
11. de Walque D. Does Education Affect HIV Status? Evidence from five African Countries. *World Bank Econ Rev* 2009; **23**(2): 209-33.
12. Cutler DM, Lleras-Muney A. Understanding differences in health behaviors by education. *Journal of health economics* 2010; **29**(1): 1-28.
13. Psacharopoulos G. Returns to Investment in Education - a Global Update. *World Dev* 1994; **22**(9): 1325-43.
14. Schultz TW. Capital Formation by Education. *J Polit Econ* 1960; **68**(6): 571-83.
15. Siphambe HK. Rates of return to education in Botswana. *Econ Educ Rev* 2000; **19**(3): 291-300.
16. Willis RJ. A new approach to the economic theory of fertility behavior. *J Polit Econ* 1973; **81**(2).
17. Baird SJ, Garfein RS, McIntosh CT, Ozler B. Effect of a cash transfer programme for schooling on prevalence of HIV and herpes simplex type 2 in Malawi: a cluster randomised trial. *Lancet* 2012; **379**(9823): 1320-9.
18. Dunkle KL, Jewkes RK, Brown HC, Gray GE, McIntryre JA, Harlow SD. Transactional sex among women in Soweto, South Africa: prevalence, risk factors and association with HIV infection. *Social science & medicine* 2004; **59**(8): 1581-92.

19. Mabsout R, van Staveren I. Disentangling Bargaining Power from Individual and Household Level to Institutions: Evidence on Women's Position in Ethiopia. *World Dev* 2010; **38**(5): 783-96.
20. Epstein E, Guttman R. Mate selection in man: evidence, theory, and outcome. *Social biology* 1984; **31**(3-4): 243-78.
21. Pettifor AE, Rees HV, Kleinschmidt I, et al. Young people's sexual health in South Africa: HIV prevalence and sexual behaviors from a nationally representative household survey. *Aids* 2005; **19**(14): 1525-34.
22. Alsan MM, Cutler DM. Girls' education and HIV risk: evidence from Uganda. *Journal of health economics* 2013; **32**(5): 863-72.
23. Black SE, Devereux PJ, Salvanes KG. Staying in the Classroom and out of the maternity ward? The effect of compulsory schooling laws on teenage births*. *The Economic Journal* 2008; **118**(530): 1025-54.
24. Kravdal O. Education and Fertility in Sub-Saharan Africa. *Demography* 2002; **39**(2): 233-50.
25. Martin TC. Women's Education and Fertility: Results from 26 Demographic and Health Surveys. *Studies in Family Planning* 1995; **26**(4): 187-202.
26. McCrary J, Royer H. The Effect of Female Education on Fertility and Infant Health: Evidence from School Entry Policies Using Exact Date of Birth. *Am Econ Rev* 2011; **101**(1): 158-95.
27. Monstad K, Propper C, Salvanes KG. Education and Fertility: Evidence from a Natural Experiment*. *Scandinavian Journal of Economics* 2008; **110**(4): 827-52.
28. Osili UO, Long BT. Does female schooling reduce fertility? Evidence from Nigeria. *Journal of Development Economics* 2008; **87**(1): 57-75.
29. Case A, Paxson C. HIV Risk and Adolescent Behaviors in Africa. Allied Social Science Associations (ASSA) Annual Meeting. San Diego, California; 2013.
30. Kohler HP, Thornton RL. Conditional Cash Transfers and HIV/AIDS Prevention: Unconditionally Promising? *The World Bank Economic Review* 2011; **26**(2): 165-90.
31. Clark D, Royer H. The Effect of Education on Adult Mortality and Health: Evidence from Britain. *Am Econ Rev* 2013; **103**(6): 2087-120.
32. Cutler D, Lleras-Muney A. Education and Health: Evaluating Theories and Evidence. *National Bureau of Economic Research Working Paper Series* 2006; **No. 12352**.
33. de Walque D. Does education affect smoking behaviors? Evidence using the Vietnam draft as an instrument for college education. *Journal of health economics* 2007; **26**(5): 877-95.
34. Dupas P. Health Behavior in Developing Countries. *Annual Review of Economics* 2011; **3**(1): 425-49.
35. Lleras-Muney A. The Relationship Between Education and Adult Mortality in the United States. *Rev Econ Stud* 2005; **72**(1): 189-221.
36. Ross CE, Wu C-l. The Links Between Education and Health. *American Sociological Review* 1995; **60**(5): 719-45.
37. Berkman LF, Kawachi I. Social epidemiology. New York: Oxford University Press; 2000.
38. Cutler D, Deaton A, Lleras-Muney A. The Determinants of Mortality. *Journal of Economic Perspectives* 2006; **20**(3): 97-120.
39. Glied S, Lleras-Muney A. Technological innovation and inequality in health. *Demography* 2008; **45**(3): 741-61.

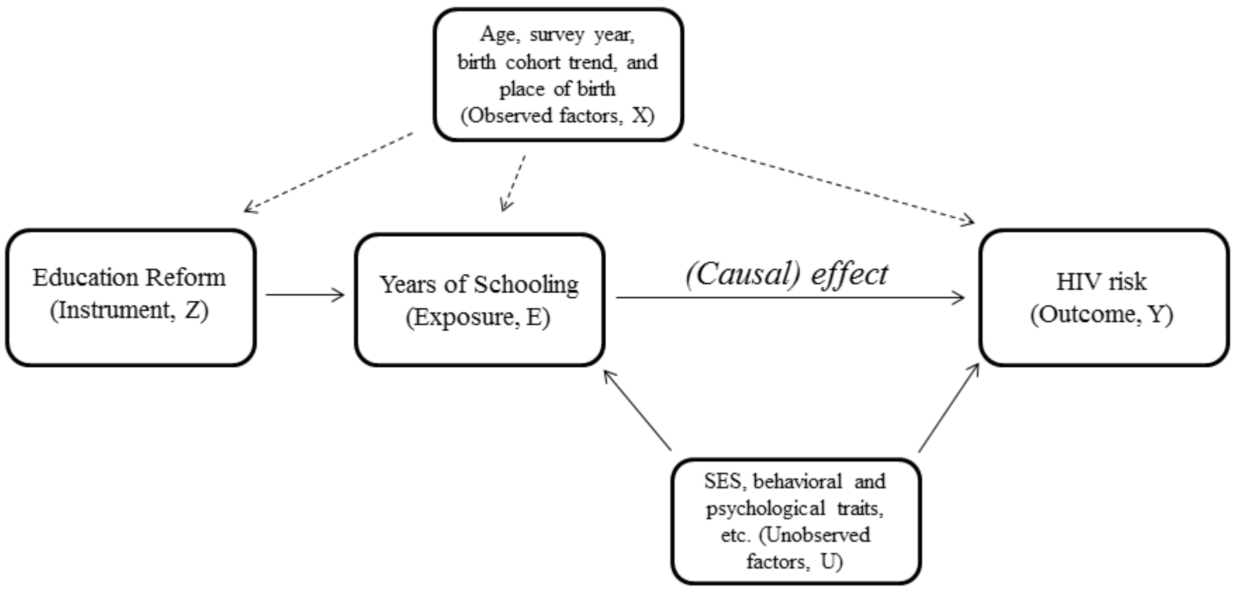
40. Kaufman JS, Cooper RS. Seeking causal explanations in social epidemiology. *American journal of epidemiology* 1999; **150**(2): 113-20.
41. Krieger N. Theories for social epidemiology in the 21st century: an ecosocial perspective. *International journal of epidemiology* 2001; **30**(4): 668-77.
42. Meara ER, Richards S, Cutler DM. The Gap Gets Bigger: Changes In Mortality And Life Expectancy, By Education, 1981-2000. *Health Affairs* 2008; **27**(2): 350-60.
43. Cho H, Hallfors DD, Mbai II, et al. Keeping Adolescent Orphans in School to Prevent Human Immunodeficiency Virus Infection: Evidence From a Randomized Controlled Trial in Kenya. *J Adolescent Health* 2011; **48**(5): 523-6.
44. Duflo E, Dupas P, Kremer M, Sinei S. Education and HIV/AIDS prevention : evidence from a randomized evaluation in Western Kenya. Washington, D.C.: World Bank, Development Economics, Office of the Senior Vice President and Chief Economist; 2006.
45. Hallfors D, Cho H, Rusakaniko S, Iritani B, Mapfumo J, Halpern C. Supporting adolescent orphan girls to stay in school as HIV risk prevention: evidence from a randomized controlled trial in Zimbabwe. *American journal of public health* 2011; **101**(6): 1082-8.
46. 068 HPTN. Effects of cash transfer for the prevention of HIV in young South African women. http://www.hptn.org/research_studies/hptn068.asp (accessed Mar 27 2014).
47. Abdool-Karim Q. A proof of concept cluster randomised controlled trial to evaluate the impact of a cash incentivised prevention intervention to reduce HIV infection in high school learners in rural KwaZulu-Natal, South Africa. http://clinicaltrials.gov/ct2/show/NCT01187979 (accessed 27 Mar 2014).
48. Borkum E. Grade structure, educational attainment and labor market outcomes: Evidence from Botswana. 2009.
49. Duflo E. Schooling and labor market consequences of school construction in Indonesia: Evidence from an unusual policy experiment. *Am Econ Rev* 2001; **91**(4): 795-813.
50. Chicoine L. Education and Fertility: Evidence from a Policy Change in Kenya. *Institute for the Study of Labor (IZA), Research Paper Series*; **6778**
51. Botswana Go. Education Report Drawn From 2009/10 Botswana Core Welfare Indicators Survey. Gaborone, Botswana: Central Statistics Office Botswana, 2013.
52. Botswana Ro. The revised national policy on education. Gaborone, Botswana: Botswana Government Printer; 1994.
53. Central Statistics Office GoB. 2004 Botswana AIDS Impact Survey II. 2004.
54. Central Statistics Office GoB. 2008 Botswana AIDS Impact Survey III. 2008.
55. Pettifor AE, van der Straten A, Dunbar MS, Shiboski SC, Padian NS. Early age of first sex: a risk factor for HIV infection among women in Zimbabwe. *Aids* 2004; **18**(10): 1435-42.
56. Wooldridge JM. Introductory Econometrics: A Modern Approach: Michigan State University; 2013.
57. Angrist JD, Krueger AB. Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments. *J Econ Perspect* 2001; **15**(4): 69–85.
58. Little RJ, Rubin DB. Causal Effects in Clinical and Epidemiological Studies Via Potential Outcomes: Concepts and Analytical Approaches. *Annual Review of Public Health* 2000; **21**(1): 121-45.
59. Botswana. Central Statistics Office., National AIDS Co-ordinating Agency (Botswana). Botswana AIDS impact survey III : statistical report. Gaborone: Central Statistics Office; 2009.
60. Imbens GW, Angrist JD. Identification and Estimation of Local Average Treatment Effects. *Econometrica* 1994; **62**(2): 467.

61. Angrist JD, Imbens GW, Rubin DB. Identification of causal effects using instrumental variables. *J Am Stat Assoc* 1996; **91**(434): 444-55.
62. Hernan MA, Robins JM. Instruments for causal inference: an epidemiologist's dream? *Epidemiology* 2006; **17**(4): 360-72.
63. Barnighausen T, Bloom DE, Humair S. Economics of antiretroviral treatment vs. circumcision for HIV prevention. *Proceedings of the National Academy of Sciences of the United States of America* 2012; **109**(52): 21271-6.
64. Hallett TB, Baeten JM, Heffron R, et al. Optimal uses of antiretrovirals for prevention in HIV-1 serodiscordant heterosexual couples in South Africa: a modelling study. *PLoS medicine* 2011; **8**(11): e1001123.
65. Kahn JG, Marseille E, Auvert B. Cost-effectiveness of male circumcision for HIV prevention in a South African setting. *PLoS medicine* 2006; **3**(12): e517.
66. Pretorius C, Stover J, Bollinger L, Bacaer N, Williams B. Evaluating the cost-effectiveness of pre-exposure prophylaxis (PrEP) and its impact on HIV-1 transmission in South Africa. *PloS one* 2010; **5**(11): e13646.
67. Kuh D. Life course epidemiology. *Journal of Epidemiology & Community Health* 2003; **57**(10): 778-83.
68. Harling G, Newell M-L, Tanser F, Kawachi I, Subramanian SV, Bärnighausen T. Do Age-Disparate Relationships Drive HIV Incidence in Young Women? Evidence from a Population Cohort in Rural KwaZulu-Natal, South Africa. *JAIDS Journal of Acquired Immune Deficiency Syndromes* 2014; **66**(4): 443-51.
69. Leclerc-Madlala S. Age-disparate and intergenerational sex in southern Africa: the dynamics of hypervulnerability. *Aids* 2008; **22 Suppl 4**: S17-25.
70. Longfield K, Glick A, Waithaka M, Berman J. Relationships Between Older Men and Younger Women: Implications for STIs/HIV in Kenya. *Studies in Family Planning* 2004; **35**(2): 125-34.
71. Luke N. Confronting the 'sugar daddy' stereotype: age and economic asymmetries and risky sexual behavior in urban Kenya. *International family planning perspectives* 2005; **31**(1): 6-14.
72. Levinsohn J, Dinkelman T, Majelantle R. When Knowledge is not Enough: HIV/AIDS Information and Risky Behavior in Botswana. *National Bureau of Economic Research Working Paper Series* 2006; **No. 12418**.
73. siphambe h. Rates of Return to Education in Botswana: Results from the 2002/2003 Household Income and Expenditure Survey Data Set. *South African Journal of Economics* 2008; **76**(4): 641-51.
74. Pollak R. Bargaining Power in Marriage: Earnings, Wage Rates and Household Production. *National Bureau of Economic Research Working Paper Series* 2005; **No. 11239**.
75. Ainsworth M, Beegle K, Nyamete A. The Impact of Women's Schooling on Fertility and Contraceptive Use: A Study of Fourteen Sub-Saharan African Countries. *The World Bank Economic Review* 1996; **10**(1): 85-122.
76. Sackey HA. Female labour force participation in Ghana : the effects of education. Nairobi: African Economic Research Consortium; 2005.
77. Becker SO, Cinnirella F, Woessmann L. The trade-off between fertility and education: evidence from before the demographic transition. *Journal of Economic Growth* 2010; **15**(3): 177-204.

78. Bloom D, Canning D. Global Demographic Change: Dimensions and Economic Significance. *National Bureau of Economic Research Working Paper Series* 2004; **No. 10817**.
79. Bloom DE, Canning D, Fink G, Finlay JE. Fertility, female labor force participation, and the demographic dividend. *Journal of Economic Growth* 2009; **14**(2): 79-101.
80. Caldwell JC. Mass Education as a Determinant of the Timing of Fertility Decline. *Population and Development Review* 1980; **6**(2): 225-55.
81. Tsheko G. HIV/AIDS and Education in Botswana. In: Alexander W. Wiseman RNG, ed. The Impact of HIV/AIDS on Education Worldwide (International Perspectives on Education and Society, Volume 18): Emerald Group Publishing Limited; 2012: 141-59.
82. Behrman JA. The effect of increased primary schooling on adult women's HIV status in Malawi and Uganda: Universal Primary Education as a natural experiment. *Social science & medicine* 2014.
83. Angrist JD, Krueger AB. Does compulsory school attendance affect schooling and earnings? *Q J Econ* 1991; **106**(4): 979-1014.
84. Angrist JD, Lavy V. Using Maimonides' rule to estimate the effect of class size on scholastic achievement. *Q J Econ* 1999; **114**(2): 533-75.
85. Angrist JD, Pathak PA, Walters CR. Explaining charter school effectiveness. *Am Econ J-Appl Econ* 2013; **5**(4): 1-27.

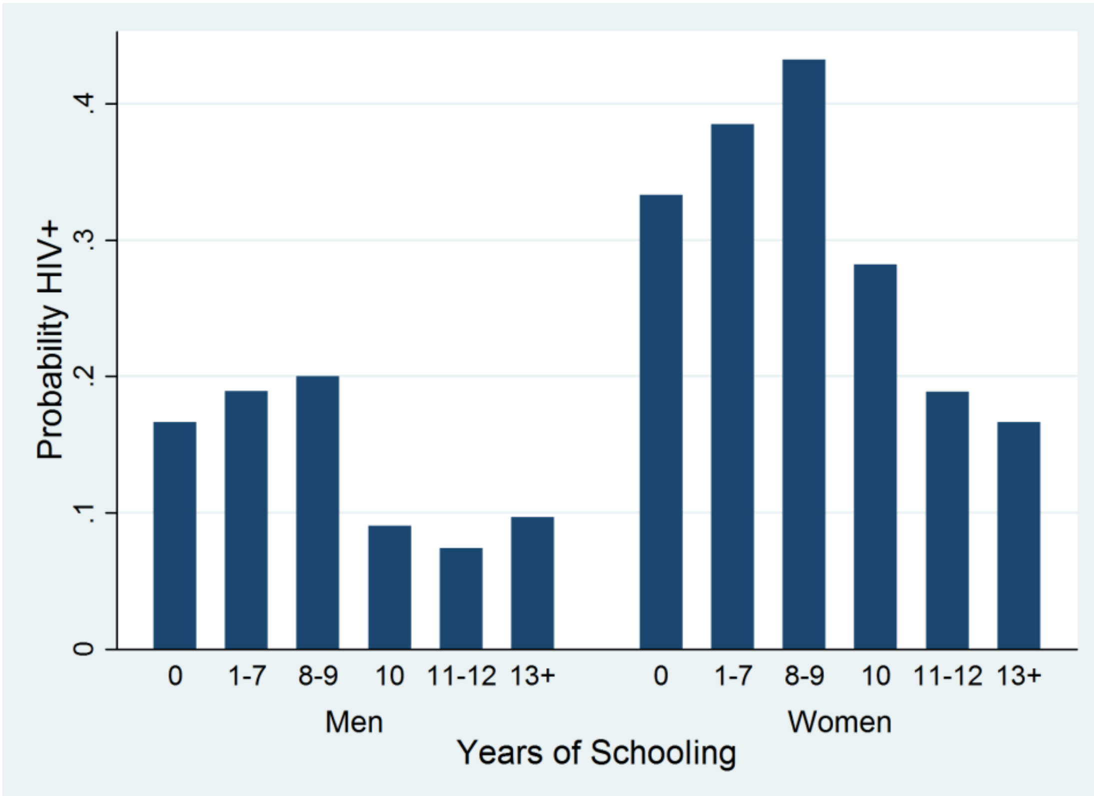
FIGURES

Figure 1: Causal Diagram.



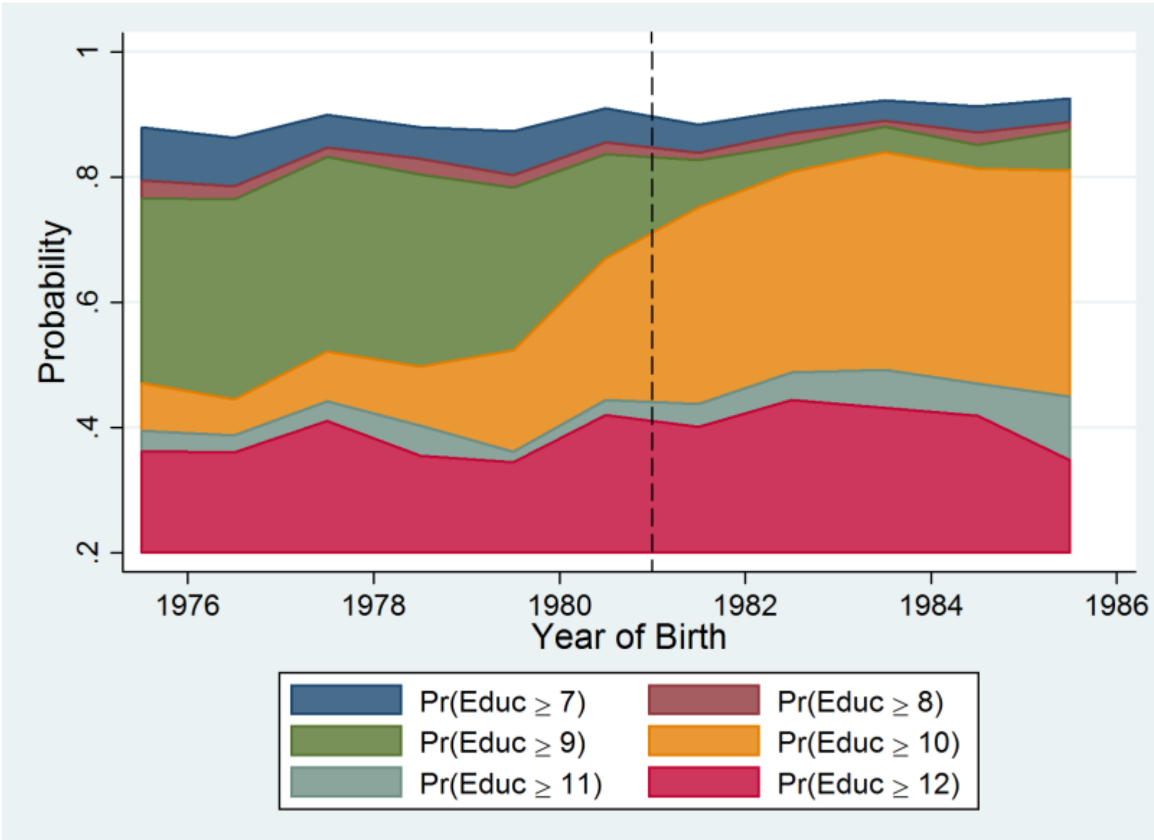
Directed acyclic graph showing the instrumental variable assumptions underpinning our study. Conditional on X, Z is a valid instrument if Z causally affects E, Z is uncorrelated with U, and Z affects Y only through E. Under the assumption that Z only affects E in one direction, 2SLS identifies a local average treatment effect (LATE).

Figure 2: HIV Prevalence by Years of Schooling in Botswana.



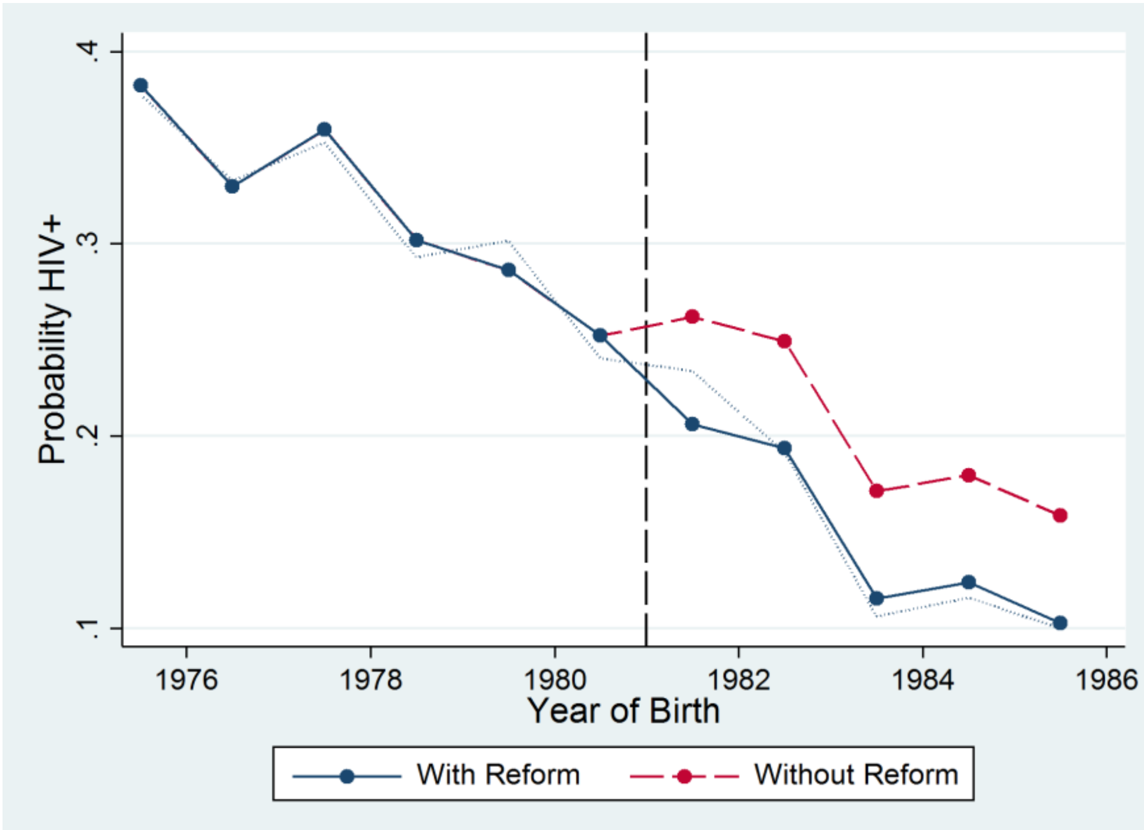
HIV prevalence by years of schooling completed and gender. Sample includes survey respondents who were citizens of Botswana, at least 18 years old at the time of the surveys, born in or after 1975, and had a valid HIV test result. Source: Botswana AIDS Impact Survey II (2004) and III (2008).

Figure 3: Educational Attainment by Birth Cohort in Botswana.



Pr(Educ ≥ X) is the probability that the respondent has attained at least X years of schooling. Sample includes survey respondents who were citizens of Botswana, at least 18 years old at the time of the surveys, born between 1975 and 1985, and had a valid HIV test result. Survey weights used as provided. Source: Botswana AIDS Impact Survey II (2004) and III (2008).

Figure 4: HIV Prevalence by Birth Cohort in Botswana.



HIV prevalence by birth cohort in Botswana with and without education reform. (Predicted HIV Infection with Reform: solid blue line; Predicted HIV Infection without Reform: broken red line; Observed HIV Infection: dotted blue line). Sample includes survey respondents who were citizens of Botswana, at least 18 years old at the time of the surveys, born between 1975 and 1985, and had a valid HIV test result. Survey weights used as provided. Source: Botswana AIDS Impact Survey II (2004) and III (2008).

TABLES

Table 1: Summary Statistics.

Variables	Percent / Mean (SD)			
Survey Year	BAIS II (2004)		BAIS III (2008)	
Subsample	Female	Male	Female	Male
HIV Positive (%)	28.3	11.1	27.3	12.4
Age	22.7 (3.1)	22.6 (3.2)	24.9 (4.2)	24.7 (4.3)
Years of Schooling	10.0 (3.0)	9.7 (4.0)	10.5 (3.2)	10.3 (3.8)
Has At Least Ten Years of Schooling (%)	62.4	65.2	72.6	73.0
Ever Had Sex (%)	88.2	77.9	92.7	83.1
Age at First Intercourse	18.0 (2.0)	17.8 (2.5)	18.2 (2.5)	18.5 (3.0)
Ever Married (%)	4.93	1.00	7.10	2.60
Literacy (%)	83.0	80.0	91.1	86.0
Total N with HIV Result	1,760	1,354	2,205	1,699

Sample includes survey respondents who were citizens of Botswana, at least 18 years old at the time of the surveys, born in or after 1975, and had a valid HIV test result. Total N with Age at First Intercourse was 1,520 for women and 1,012 for men in BAIS II (2004), and 1,987 for women and 1,348 for men in BAIS III (2008). Sample weights used as provided. Source: Botswana AIDS Impact Survey II (2004) (N: 14,802) and III (2008) (N: 13,479).

Table 2: Natural Experiment: First Stage, Intention-To-Treat, and 2SLS Results.

	(1)	(2)	(3)
Model	**First Stage**	**Intention-to-treat**	**2SLS**
Dependent Variable	**Years of Schooling**	**HIV-Positive**	**HIV-Positive**
Coefficient on Reform Indicator			
Female	0.635***	-0.074**	-0.116**
	(0.223)	(0.031)	(0.058)
Observations	3,965	3,965	3,965
R-squared	0.034	0.095	-
Probability HIV-positive, Pre-Reform	-	0.323	0.287
Male	1.005***	-0.050*	-0.050*
	(0.322)	(0.026)	(0.029)
Observations	3,053	3,053	3,053
R-squared	0.033	0.070	-
Probability HIV-positive, Pre-Reform	-	0.168	0.164
Both Sexes	0.792***	-0.064***	-0.081***
	(0.188)	(0.021)	(0.031)
Observations	7,018	7,018	7,018
R-squared	0.036	0.123	-
Probability HIV-positive, Pre-Reform	-	0.255	0.238

Regressions 1 to 2 are OLS models. Regression 3 is a 2SLS model, in which exposure to the reform was used as an instrument for years of schooling. All models included the following controls: single-year age indicators, a linear term for year of birth, an indicator for survey wave and indicators for district of birth. Regressions for the subsample with both sexes additionally control for age*sex, districtofbirth*sex, yearofbirth*sex and surveywave*sex interactions. The instrument was defined as = 1 if YOB > 1980. F-Statistics in the 2SLS models were 8.6 for women, 9.5 for men, and 18.0 in the sample with both sexes. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Sample includes survey respondents who were citizens of Botswana, at least 18 years old at the time of the surveys, born in or after 1975, and had a valid HIV test result. No weights were used. Source: Botswana AIDS Impact Survey II (2004) and III (2008).

' Positive	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)[¥]	(12)
s variable												

Table 3: Sensitivity Analyses: Assessing the Robustness of the 2SLS Results to: Additional Controls, Sample Restrictions, Sample Weights, Imputed HIV Estimates, and a 2SLS Difference-in-Differences Strategy.

	.	.	.	.	.	.	.	.	.	.	.	.
)istrictofBirth,)istrictofBirth#i.Sex	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	-
_Year##i.Sex, ex, ex	-	-	-	-	-	-	-	-	-	-	-	✓
i.Sex	-	✓	-	-	-	-	-	-	-	-	-	-
x): c.Age#i.Sex, #i.Sex, c.Agek20##i.Sex, k30##i.Sex	-	-	✓	-	-	-	-	-	-	-	-	-
	-	-	-	✓	-	-	-	-	-	-	-	-
#i.Sex	✓	✓	✓	-	-	✓	✓	✓	✓	-	-	-
#i.Sex, r##i.Sex	-	-	-	-	✓	-	-	-	-	-	-	-
tor, tor#i.Sex	-	-	-	-	-	✓	-	-	-	-	-	-
ndows												
	✓	✓	✓	✓	✓	✓	-	-	✓	✓	✓	✓
	-	-	-	-	-	-	✓	-	-	-	-	-
	-	-	-	-	-	-	-	✓	-	-	-	-
	-	-	-	-	-	-	-	-	✓	-	-	-
	-	-	-	-	-	-	-	-	-	✓	✓	-
ations	7,018	7,018	7,018	7,018	7,018	7,018	5,712	4,987	7,018	10,039	8,281	7,018
tistic	17.1	12.5	14.8	17.9	0.9	13.3	5.9	1.9	13.4	n/a	n/a	0.6

odels, in which exposure to the reform was used as an instrument for years of schooling. Model 3 controls for a third order spline in age (with knots at 20, 25, and urrleman et al. 1989). The instrument was defined as = 1 if YOB > 1980. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. ¥ Uses additional tercourse and Ever Married, to impute HIV status. Sample includes survey respondents who were citizens of Botswana, at least 18 years old at the time of the surveys, sult. Source: Botswana AIDS Impact Survey II (2004) and III (2008).

Table 4: 2SLS Results: The Effect of Education on Sexual Intercourse, Contraceptive Use, Fertility, HIV Testing Behavior, and HIV Knowledge. (Panel A)

Coefficient on Years of Schooling	**(1)** **Female**	**(2)** **Male**
Dependent Variable		
Ever Had Intercourse (1=yes, 0=no)	-0.007	0.056*
	(0.027)	(0.030)
Observations	3,965	3,050
F-Statistic	8.6	9.6
Proportion Ever Had Intercourse, Pre-Reform	0.962	0.932
Age at First Intercourse	0.761***	0.065
	(0.261)	(0.209)
Observations	3,507	2,360
F-Statistic	11.9	10.4
Mean Age At First Intercourse, Pre-Reform	18.6	18.6
First sex ever: Did you use anything to protect yourself (eg, condom)? (1=yes, 0=no)	0.127***	0.055*
	(0.047)	(0.028)
Observations	3,582	2,458
F-Statistic	9.9	11.0
Proportion Protected, Pre-Reform	0.828	0.863
Ever Given Birth (1=yes, 0=no)	-0.158***	-
	(0.057)	-
Observations	3,644	-
F-Statistic	10.0	-
Proportion Ever Given Birth, Pre-Reform	0.728	-
Have you ever been tested for HIV, the virus that causes AIDS? (1=yes, 0=no)	0.110*	0.120**
	(0.061)	(0.052)
Observations	3,793	2,922
F-Statistic	7.7	7.9
Proportion Ever Tested, Pre-Reform	0.720	0.573
Indicator for knowledge of at least one HIV prevention strategy (1=yes, 0=no)	-0.021	0.007
	(0.020)	(0.020)
Observations	3,791	2,919
F-Statistic	7.9	7.9
Proportion Yes on Indicator, Pre-Reform	0.949	0.952

All regressions are 2SLS models, in which exposure to the reform was used as an instrument for years of schooling. All models included the following controls: single-year age indicators, a linear term for year of birth, an indicator for survey wave and indicators for district of birth. The instrument was defined as = 1 if YOB > 1980. The indicator for knowledge of HIV prevention strategies was defined as 1 if respondent could name at least one out of the six following HIV prevention strategies: condoms, fewer partners, mutually faithful relationship, abstinence, avoid injections with contaminated needles, and avoid blood transfusions. Those responding "don't know" to an HIV knowledge question were accounted for as incorrect. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Sample includes survey respondents who were citizens of Botswana, at least 18 years old at the time of the surveys, born in or after 1975, and had a valid HIV test result. No weights were used. Source: Botswana AIDS Impact Survey II (2004) and III (2008).

Table 4: 2SLS Results: The Effect of Education on HIV Misperceptions, Number of Partners, HIV Discussion, HIV Attitudes, Literacy, and Labor Force Participation. (Panel B)

Coefficient on Years of Schooling	(1) Female	(2) Male
Dependent Variable		
Indicator for any misperceptions about HIV (1=yes, 0=no)	-0.062	-0.038
	(0.059)	(0.043)
Observations	3,782	2,915
F-Statistic	7.2	7.6
Proportion Yes on Indicator, Pre-Reform	0.564	0.603
Indicator for 2 or more sexual partners in the last 12 months (1=two or more, 0=one or zero)	0.044	0.111**
	(0.045)	(0.051)
Observations	3,658	2,877
F-Statistic	6.3	8.5
Proportion One on Indicator, Pre-Reform	0.115	0.232
During the past 4 weeks, have you discussed HIV/AIDS with anyone? (1=yes, 0=no or not sure)	0.003	0.087*
	(0.060)	(0.046)
Observations	3,791	2,918
F-Statistic	7.7	8.2
Proportion Which Discussed HIV/AIDS, Pre-Reform	0.484	0.471
Do you think it should be acceptable for a woman to obtain male condoms? (1=yes, 0=no or not sure)	0.080**	0.096**
	(0.038)	(0.039)
Observations	3,832	2,957
F-Statistic	9.6	9.0
Proportion Yes on Outcome, Pre-Reform	0.933	0.875
Can you read and understand a letter / newspaper / bible? (1=easily, 0=no or with difficulty)	0.003	0.084***
	(0.036)	(0.023)
Observations	3,962	3,051
F-Statistic	8.7	9.7
Proportion Easily on Outcome, Pre-Reform	0.866	0.831
Labor Force Participation (1=yes, 0=no)	0.172**	0.048
	(0.076)	(0.039)
Observations	3,942	3,037
F-Statistic	8.4	9.3
Proportion Labor Force Participation, Pre-Reform	0.706	0.846

ressions are 2SLS models, in which exposure to the reform was used as an instrument for years of schooling. All models included the following controls: single-year age indicators, a linear term for year of birth, an indicator for survey wave and indicators for district of birth. The instrument was defined as = 1 if YOB > 1980. The indicator for any misperceptions was defined as 1 if respondent incorrectly answered any of the following four questions: whether HIV spreads via mosquitos, sharing a meal with an HIV+ person, due to witchcraft, and whether a healthy looking person can be HIV+. Those responding "don't know" to an HIV misconception question were accounted for as incorrect. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Sample includes survey respondents who were citizens of Botswana, at least 18 years old at the time of the surveys, born in or after 1975, and had a valid HIV test result. No weights were used. Source: Botswana AIDS Impact Survey II (2004) and III (2008).

Table 5: Cost-effectiveness Ratio of Secondary School and Known HIV Prevention Interventions.

Intervention	**Medical Male Circumcision**	**Treatment as Prevention (CD4 ≥350/µL)**	**Pre-exposure Prophylaxis**	**Antiretroviral treatment (CD4 <350/µL)**	**Secondary School**
Cost-effectivness Ratio ($ / infection averted)	551; 1,096	8,375	12,500 - 20,000; -6000 - 26,000	6,790	27,753
Study (year)	Kahn et al. (2006), Barnighausen et al. (2012)	Barnighausen et al. (2012)	Pretorius et al. (2010), Hallett et al. (2011)	Barnighausen et al. (2012)	Authors (2014)

APPENDIX

A. Description of Cost Effectiveness Calculations

The per-pupil-per-year cost of secondary education was $2,248 using the average of the 2005 and 2007 UNESCO Institute for Statistics estimates. Since Batswana who stayed in school for an additional year had a 8.1 percentage point lower risk of HIV infection, the cost per HIV infection averted was $27,753 USD. Standard cost-effectiveness benchmarks indicate that an intervention is very cost-effective if it costs less than 1 x per capita GDP for each disability-adjusted life year (DALY) averted.

We used a simple model to estimate the number of DALYs resulting from an HIV infection under two scenarios: the person initiates ART when eligible; the person does not initiate antiretroviral therapy (ART). We based all estimates on an infection occurring at age 20. Under the "no ART scenario," we assumed that HIV infection was followed by 7 years of asymptomatic HIV (disability weight 0), one year of symptomatic HIV (disability weight 0.221), two years of AIDS (disability weight 0.547), and then death, occurring ten years after initial infection. Under the "ART scenario," we assumed 7 years of asymptomatic HIV, one year of symptomatic HIV, followed by 27 years of ART (disability weight 0.053). Disability weights were obtained from the 2010 Global Burden of Disease (Salomon et al. 2012). To calculate DALYs, we need an estimate of the number of life years that a person would live if he or she did not contract HIV at age 20; we used WHO life table estimates for Botswana for 2009, indicating a life expectancy of 45 years at age 20 (WHO 2012). Survival on ART was estimated as 75% of "normal" life expectancy at age of ART initiation (36 years at age 30), based on published estimates from neighboring South Africa (Johnson et al. 2013). We used per patient per year cost estimates of $202 for pre-ART and $880 for ART published in Menzies et al. (2011).

Based on these inputs, and discounting both DALYs and costs at 3%, we estimated that an HIV infection at age 20 would lead to: 16.3 DALYs for a person who did not initiate ART; and 3.5 DALYs for a person who initiated ART, with a present-discounted lifetime treatment cost of $12,400. These estimates of lifetime treatment costs are of similar magnitude to an independent estimate of the present-discounted fiscal costs of an HIV infection in Botswana, projected at twice GDP per capita (Lule & Haacker, 2012).

For the "no ART scenario", we divided the $27,753 per infection averted by the 16.3 DALYs per HIV infection in the absence of ART, yielding a cost-effectiveness ratio (CER) of $1,703 /DALY. For the "ART scenario", we divided the net cost of secondary school ($27,753 - $12,400 = $15,353) by the 3.5 DALYs per HIV infection with ART initiation when eligible, yielding a CER of $4,387 /DALY. Each of these CERs is less than Botswana's $5,178 per capita GDP, implying that secondary school is very cost-effective as an HIV prevention intervention.

B. Additional Robustness Checks

Alternative Statistical Analyses

Table S7 in the Appendix displays ITT results using either logistic regression or standard errors clustered at the year of birth level. The clustered standard errors were smaller than the non-clustered standard errors.

Attrition and Consent Rates

One concern is that results could be biased by differential consent rates by birth cohort or by selection bias from mortality risk associated with being born after 1980. Increased education may have improved access to antiretroviral therapy, which became available in 2003 in Botswana; however, this would lead to a higher HIV prevalence among those with more education, and hence a bias against the direction of our results. Figure S2 shows the proportion of respondents without an HIV test and the size of the surviving birth cohorts in the study sample. There was no evidence of a post-1980 cohort effect in either of these variables that might bias our estimates. We also assessed whether differential non-consent might have biased our results, by imputing HIV status for respondents with missing HIV biomarker data among Batswana. Results using imputed HIV estimates were similar to our main model. Further, we show in a regression context that consent rates do not change much with exposure to the reform. To do so, we ran the main intention-to-treat regression model, but using "consent" as the outcome. Table S6 in the displays ITT results for consent rates by sex and in the pooled sample. The small size of the effect is unlikely to explain the large effects of schooling on HIV risk we observe. Lastly, we note that the datasets we used do not contain information on interviewer identity, which would have allowed us to use Heckman-type selection models to correct for selection on unobserved variables (Bärnighausen et al. 2011).

Alternative Explanations

In our main results, we controlled non-parametrically for age, district of birth, and survey year and included a linear trend for year of birth. Our identifying assumption is that there are no other cohort-specific exposures that influence HIV risk for persons born after 1980, conditional on long run trends in HIV risk across birth cohorts. This assumption could be violated if long run trends are non-linear, or if some other intervention affected specific birth cohorts (or equivalent, targeted specific age groups in specific years). First, to assess the robustness of our results to the presence of non-linearities in long-run cohort trends in HIV risk, we controlled for a quadratic in year of birth, quadratic term in age, and cubic spline in age in addition to the linear term included in the main analysis. Second, we also estimated our main model for a narrower window of birth cohorts – 1981 +/- 4, 5 years. The assumption that underlying trends are approximately linear is more plausible the narrower the window of cohorts included. Third, to rule out the possibility that other national policy changes might have affected HIV risk for the same cohorts, we used an alternative identification strategy. The education policy reform would be expected to have the greatest impact on years of schooling in those districts where a large fraction of students completed exactly nine years of schooling in the pre-reform period. Figure S3 in the Appendix displays educational attainment by districts with either a high or low proportion of people with exactly nine years of schooling. We created an indicator for whether a subject's district of birth had high vs. low grade-nine completion, and used as our instrument the interaction of this variable with the indicator for reform cohort, while controlling for the main effects of each variable.

Weighting

In analytical inference, the use of sample weights is subject to controversy. We added sample weights to our main model as an additional robustness check. In all descriptive results, we used sample weights.

Results of Robustness Checks

Tables 3 and S5 shows the robustness of our ITT and 2SLS results to additional controls in the pooled sample and by sex, including a quadratic term in age, cubic spline in age, quadratic term in year of birth, year of birth and survey year interactions, using sample weights, year of birth and reform indicator interactions, narrower birth cohort windows, imputed HIV estimates and a 2SLS difference-in-difference estimator using an indicator for whether a subject's district of birth had high vs. low grade-nine completion. 2SLS estimates in robustness checks were similar in direction and magnitude. In the pooled sample, using either a quadratic term in age, cubic spline in age, quadratic term in year of birth, or a quadratic in year of birth with survey year interactions, Batswana who stayed in school for an additional year had a 7 percentage point lower risk of HIV infection ($p = 0.048$, $p = 0.041$, $p = 0.014$, and $p = 0.010$, respectively), using year of birth and survey year interactions or sample weights they had a 8 percentage point lower risk of HIV infection ($p = 0.008$ and $p = 0.025$, respectively). Using imputed HIV estimates, they had a 9 - 11 percentage point lower risk of HIV infection depending on the use of additional covariates, such as age at first intercourse, to impute HIV estimates ($p = 0.045$ and $p = 0.036$, respectively). Using a narrower birth cohort of 1981 +/- 5 years and a 2SLS difference-in-difference estimator, schooling appeared similarly protective but did not reach conventional benchmarks for statistical significance. In women, however, using a 2SLS difference-in-difference estimator, they had a 13 percentage point lower risk of HIV infection ($p = 0.048$). Results in robustness checks by sex were otherwise similar to our main results above.

C. References for Appendix

Bärnighausen T, Bor J, Wandira-Kazibwe S, Canning D. (2011) Correcting HIV prevalence estimates for survey nonparticipation using Heckman-type selection models**.** *Epidemiology,* **22**(1):27-35.

Deaton A. The analysis of household surveys : a microeconometric approach to development policy. Baltimore, MD: Published for the World Bank by Johns Hopkins University Press; 1997

Durrleman S, Simon R. (1989) Flexible regression models with cubic splines. *Stat Med*, **8**(5):551-61

Johnson LF, Mossong J, Dorrington RE, Schomaker M, Hoffmann CJ, et al. (2013) Life Expectancies of South African Adults Starting Antiretroviral Treatment: Collaborative Analysis of Cohort Studies. *PLoS Med* **10**(4): e1001418. doi:10.1371/journal.pmed.1001418

Lule E, Haacker M. (2012) The Fiscal Dimension of HIV/AIDS in Botswana, South Africa, Swaziland, and Uganda. Washington, DC: The World Bank.

Menzies NA, Berruti AA, Berzon R, Filler S, Ferris R. (2011). The cost of providing comprehensive HIV treatment in PEPFAR-supported programs. *AIDS,* **25**(14):1753-60.

Salomon JA, Vos T, Hogan DR, Gagnon M, Naghavi M, et al. (2012) Common values in assessing health outcomes from disease and injury: disability weights measurement study for the Global Burden of Disease Study 2010. *The Lancet*, **380**(9859): 2129-43.

WHO. (2012). World Health Statistics. Life tables for WHO Member States. Geneva, World Health Organization, 2012.

UNESCO. (2014). Institute for Statistics. Education Indicators.

D. Supplementary Figures

Figure S1: Participant Flow Diagram.

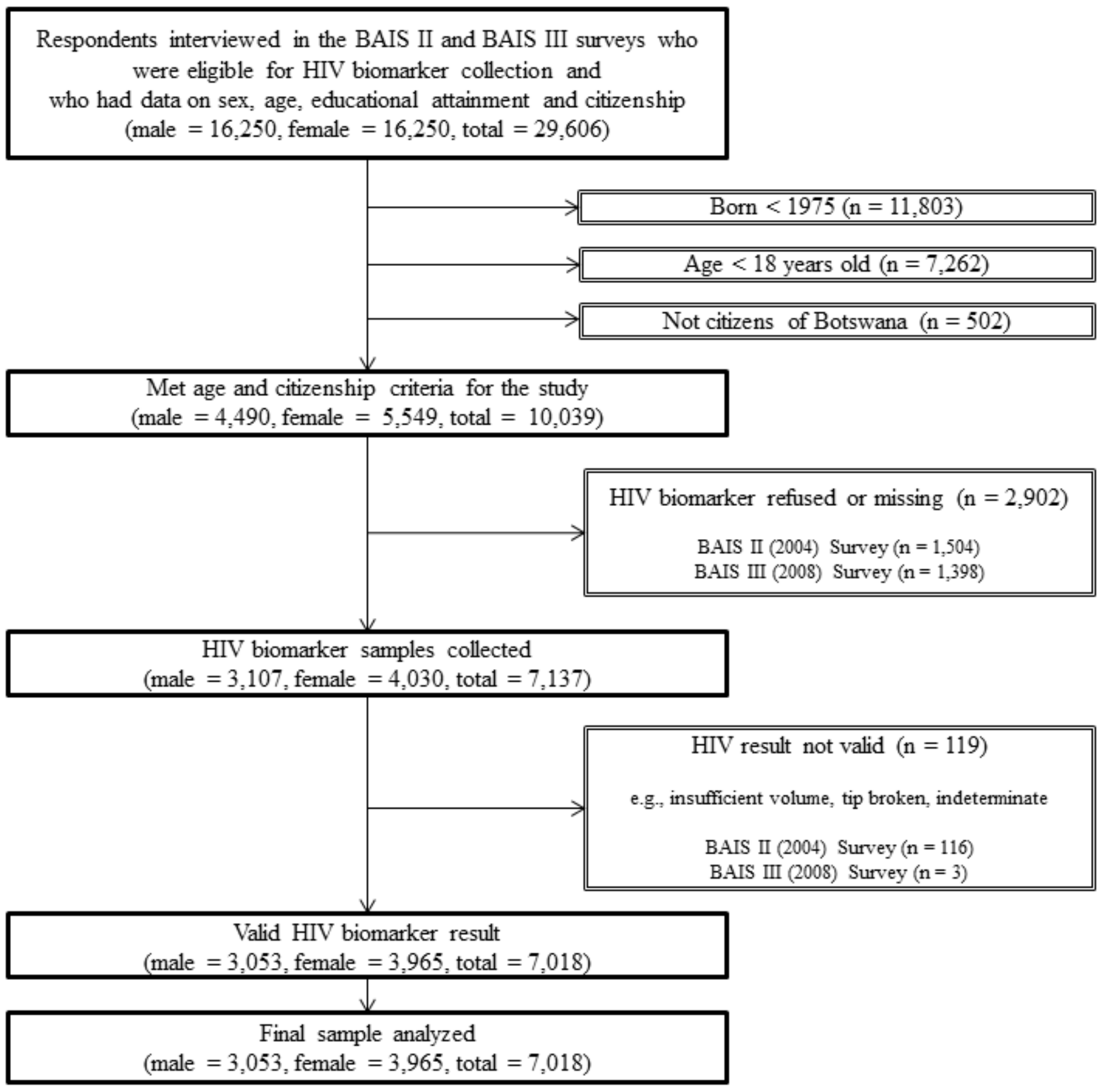
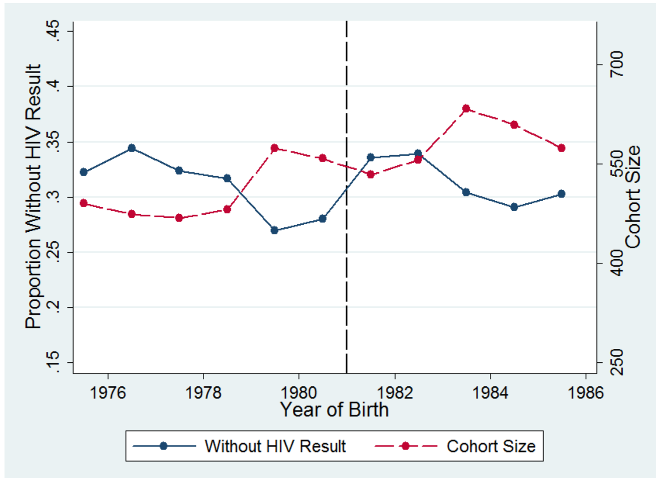


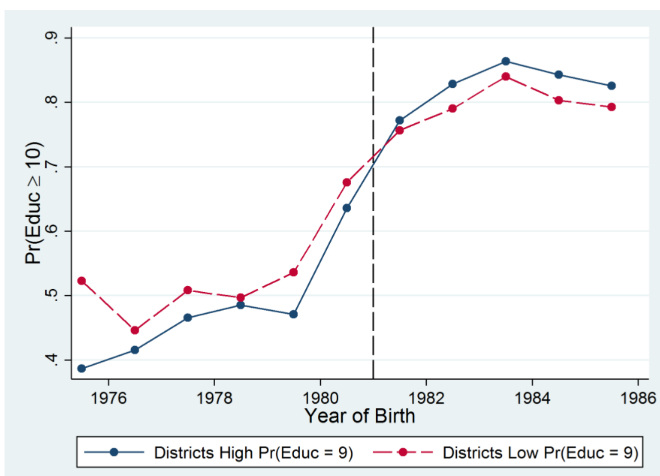
Diagram showing the flow of participants through the study's inclusion and exclusion criteria. Source: Botswana AIDS Impact Survey II (2004) and III (2008).

Figure S2: Proportion without HIV Test Result, and Cohort Size.



Sample includes survey respondents who were citizens of Botswana, at least 18 years old at the time of the surveys, born between 1975 and 1985, and had a valid HIV test result. (Proportion Without HIV Test Result: solid blue line; Cohort Size: broken red line). Survey weights used as provided. Source: Botswana AIDS Impact Survey II (2004) and III (2008).

Figure S3: Educational Attainment by Birth Cohort and District in Botswana.



Pr(Educ $\geq$ 10) is the probability that the respondent has attained at least 10 years of schooling. Sample includes survey respondents who were citizens of Botswana, at least 18 years old at the time of the surveys, born between 1975 and 1985, and had a valid HIV test result. Survey weights used as provided. Source: Botswana AIDS Impact Survey II (2004) and III (2008).

E. Supplementary Tables

Table S1: OLS Regressions: Association Between Years of Schooling and HIV Status.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Coefficients on Schooling Measures	**Female**	**Female**	**Male**	**Male**	**Both Sexes**	**Both Sexes**	**Both Sexes**
Predictor							
Years of Schooling	-0.019***		-0.007***		-0.013***		
	(0.002)		(0.002)		(0.001)		
Years of Schooling (0-9)		0.004		0.001		0.003	
		(0.004)		(0.003)		(0.002)	
Years of Schooling (10-13)		-0.045***		-0.022***		-0.036***	
		(0.006)		(0.005)		(0.004)	
At Least 7 Years							-0.003
							(0.029)
At Least 8 Years							0.024
							(0.045)
At Least 9 Years							-0.007
							(0.041)
At Least 10 Years							-0.035*
							(0.018)
At Least 11 Years							-0.038*
							(0.020)
At Least 12 Years							-0.042**
							(0.019)
At Least 13 Years							-0.048***
							(0.013)
Observations	3,965	3,541	3,053	2,658	7,018	6,199	7,018
R-squared	0.111	0.126	0.075	0.089	0.132	0.151	0.141

Sample includes survey respondents who were citizens of Botswana, at least 18 years old at the time of the surveys, born in or after 1975, and had a valid HIV test result. All regressions include age dummies, a linear term for year of birth, an indicator for survey wave and dummies for district of birth. Models 5 to 7 additionally control for age*sex, districtofbirth*sex, yearofbirth*sex and surveywave*sex interactions. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. No weights used. Source: Botswana AIDS Impact Survey II (2004) and III (2008).

Table S2: First Stage Regressions: Effect of the Education Reform on Years of Schooling.

Coefficient on Reform Indicator	(1) Female	(2) Male	(3) Both Sexes	(4) Both Sexes
Dependent Variable				
Years of Schooling	0·635***	1·005***	0·792***	
	(0·223)	(0·322)	(0·188)	
At Least 7 Years of Schooling				0·026
				(0·016)
At Least 8 Years of Schooling				0·043**
				(0·019)
At Least 9 Years of Schooling				0·042**
				(0·020)
At Least 10 Years of Schooling				0·249***
				(0·024)
At Least 11 Years of Schooling				0·069***
				(0·026)
At Least 12 Years of Schooling				0·082***
				(0·026)
At Least 13 Years of Schooling				0·031
				(0·020)
Observations	3,965	3,053	7,018	7,018
R-squared	0·034	0·033	0·036	-

All regressions include age dummies, a linear term for year of birth, an indicator for survey wave and dummies for district of birth. Columns 3 and 4 additionally control for age*sex, districtofbirth*sex, yearofbirth*sex and surveywave*sex interactions. Standard errors in parentheses. *** p<0·01, ** p<0·05, * p<0·1. No weights used. Sample includes survey respondents who were citizens of Botswana, at least 18 years old at the time of the surveys, born in or after 1975, and had a valid HIV test result. Source: Botswana AIDS Impact Survey II (2004) and III (2008).

Table S3: Intention-To-Treat Results: The Effect of the Education Reform on Sexual Intercourse, Marriage, Current Sexual and HIV Testing Behavior, HIV Knowledge and Attitudes, Labor Market Outcomes, and Literacy. (Panel A)

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Ever Had Intercourse (1=yes, 0=no)	Age at First Intercourse	First sex ever: Did you use anything to protect yourself (eg, condom)? (1=yes, 0=no)	Ever Married (1=yes, 0=no)	Years of Pre-marital Intercourse	Ever Given Birth (1=yes, 0=no)	How many times have you given birth? (never given birth=0)	Did you have sexual intercourse in the last 12 months? (1=yes, 0=no or never had intercourse)	Last sex: Did you or your partner use a condom, conditional on intercourse in the last 12 months? (1=yes, 0=no)	How old is your most recent partner, conditional on intercourse in the last 12 months? (years)
-0.004	0.592***	0.092***	0.013	-0.688***	-0.114***	-0.181**	-0.008	0.094***	-0.207
(0.017)	(0.155)	(0.027)	(0.015)	(0.167)	(0.033)	(0.079)	(0.023)	(0.035)	(0.361)
3,965	3,507	3,582	3,963	3,497	3,644	3,656	3,954	3,041	3,041
0.178	0.089	0.078	0.072	0.621	0.198	0.228	0.162	0.024	0.440
0.964	18.4	0.792	0.064	6.327	0.785	1.378	0.894	0.712	30.3
0.057**	0.073	0.064**	0.010	-0.076	-	-	0.058*	-0.019	-0.510*
(0.026)	(0.237)	(0.030)	(0.008)	(0.241)	-	-	(0.030)	(0.034)	(0.288)
3,050	2,360	2,458	3,052	2,357	-	-	3,046	2,144	2,131
0.279	0.108	0.047	0.047	0.548	-	-	0.239	0.043	0.44
0.928	18.6	0.853	0.026	6.683	-	-	0.873	0.822	22.9
0.022	0.386***	0.081***	0.012	-0.444***	-	-	0.020	0.048*	-0.332
(0.015)	(0.132)	(0.020)	(0.009)	(0.139)	-	-	(0.018)	(0.025)	(0.244)
7,015	5,867	6,040	7,015	5,854	-	-	7,000	5,185	5,172
0.253	0.099	0.069	0.078	0.592	-	-	0.211	0.046	0.610
0.948	18.4	0.818	0.048	6.481	-	-	0.885	0.761	27.0

ch include age dummies, a linear term for year of birth, an indicator for survey wave and dummies for district of birth. Models for both sexes additionally control urofbirth*sex and surveywave*sex interactions. The instrument was defined as = 1 if YOB > 1980. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. defined as Age at First Marriage minus Age at First Intercourse (if married) or Age at Time of Survey minus Age at First Intercourse (if never married). Sample re citizens of Botswana, at least 18 years old at the time of the surveys, born in or after 1975, and had a valid HIV test result. No weights were used. Source: 004) and III (2008).

Table S3: Intention-To-Treat Results: The Effect of the Education Reform on Sexual Intercourse, Marriage, Current Sexual and HIV Testing Behavior, HIV Knowledge and Attitudes, Labor Market Outcomes, and Literacy. (Panel B)

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Ever had a sexual partner who was 10 years older / younger, conditional on intercourse in the last 12 months? (1=yes, 0=no)	Indicator for 2 or more sexual partners in the last 12 months (1=two or more, 0=one or zero)	Time since first sexual intercourse with most recent partner, conditional on intercourse in the last 12 months (days)	Relationship to most recent partner, conditional on intercourse in the last 12 months? (1=spouse, live-in partner, 0=not living in, other)	Have you ever heard of the virus HIV or an illness called AIDS? (1=yes, 0=no)	Have you ever been tested for HIV, the virus that causes AIDS? (1=yes, 0=no)	In the past 12 months have you been tested for HIV, the virus that causes AIDS? (1=yes, 0=no)	Indicator for knowledge of at least one HIV prevention strategy (1=yes, 0=no)	Indicator for knowledge of at least two HIV prevention strategies (1=yes, 0=no)	Mentioned condom use as an HIV prevention strategy (1=yes, 0=no)
0.002	0.025	5.8	-0.057	-0.006	0.066**	0.067*	-0.013	-0.014	-0.026
(0.031)	(0.023)	(84.6)	(0.038)	(0.011)	(0.031)	(0.035)	(0.014)	(0.033)	(0.022)
3,364	3,658	2,953	3,057	3,874	3,793	3,787	3,791	3,791	3,795
0.035	0.016	0.171	0.089	0.011	0.234	0.098	0.020	0.020	0.025
0.232	0.099	1,285	0.470	0.990	0.683	0.493	0.956	0.706	0.896
-0.027	0.109***	-36.7	0.048	0.005	0.110***	0.069*	0.006	0.004	0.001
(0.027)	(0.037)	(76.6)	(0.039)	(0.012)	(0.038)	(0.037)	(0.019)	(0.040)	(0.029)
2,328	2,877	2,086	2,155	2,995	2,922	2,918	2,919	2,919	2,922
0.043	0.041	0.147	0.146	0.013	0.249	0.137	0.016	0.018	0.022
0.114	0.220	894	0.266	0.980	0.551	0.342	0.951	0.678	0.839
-0.010	0.061***	-11.7	-0.014	-0.001	0.085***	0.067***	-0.005	-0.006	-0.015
(0.022)	(0.021)	(59.0)	(0.028)	(0.008)	(0.024)	(0.025)	(0.011)	(0.025)	(0.018)
5,692	6,535	5,039	5,212	6,869	6,715	6,705	6,710	6,710	6,717
0.059	0.056	0.195	0.132	0.012	0.28	0.137	0.019	0.022	0.025
0.180	0.153	1,108	0.379	0.985	0.626	0.427	0.954	0.693	0.871

ich include age dummies, a linear term for year of birth, an indicator for survey wave and dummies for district of birth. Models for both sexes additionally control for
fbirth*sex and surveywave*sex interactions. The instrument was defined as = 1 if YOB > 1980. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The
:vention strategies was defined as 1 if respondent could name either at least one (model 8) or two (model 9) out of the six following HIV prevention strategies: condoms,
:lationship, abstinence, avoid injections with contaminated needles, and avoid blood transfusions. Those responding "don't know" to an HIV knowledge question were
includes survey respondents who were citizens of Botswana, at least 18 years old at the time of the surveys, born in or after 1975, and had a valid HIV test result. No weights
Impact Survey II (2004) and III (2008).

Table S3: Intention-To-Treat Results: The Effect of the Education Reform on Sexual Intercourse, Marriage, Current Sexual and HIV Testing Behavior, HIV Knowledge and Attitudes, Labor Market Outcomes, and Literacy. (Panel C)

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Mentioned mutually faithful relationship as an HIV prevention strategy (1=yes, 0=no)	Mentioned abstinence as an HIV prevention strategy (1=yes, 0=no)	Can people reduce their chances of getting HIV/AIDS by using a condom correctly every time they have sex? (1=yes, 0=no, don't know)	Can people reduce their chances of getting HIV/AIDS by having only one uninfected sex partner who has no other partners? (1=yes, 0=no,	Indicator for any misperceptions about HIV (1=yes, 0=no)	Do you think it should be acceptable for a woman to obtain male condoms? (1=yes, 0=no or not sure)	During the past 4 weeks, have you discussed HIV/AIDS with anyone? (1=yes, 0=no or not sure)	During the past 4 weeks, have you discussed HIV/AIDS with your spouse? (1=yes, 0=no)	During the past 4 weeks, have you discussed HIV/AIDS with your sexual partner? (1=yes, 0=no)	During the past 4 weeks, have you discussed HIV/AIDS with your family member(s)? (1=yes, 0=no)
0.044	0.021	0.007	-0.003	-0.036	0.060***	0.002	-0.002	0.058***	-0.014
(0.031)	(0.035)	(0.020)	(0.024)	(0.036)	(0.023)	(0.036)	(0.014)	(0.022)	(0.020)
3,794	3,793	3,791	3,780	3,782	3,873	3,791	3,791	3,791	3,791
0.038	0.017	0.015	0.012	0.023	0.037	0.031	0.023	0.020	0.019
0.230	0.616	0.924	0.876	0.585	0.900	0.483	0.045	0.110	0.091
-0.036	-0.002	0.023	0.033	-0.034	0.074***	0.081**	0.005	0.003	-0.032**
(0.035)	(0.042)	(0.025)	(0.029)	(0.042)	(0.029)	(0.041)	(0.011)	(0.024)	(0.016)
2,921	2,921	2,917	2,912	2,915	2,992	2,918	2,918	2,918	2,918
0.042	0.018	0.017	0.015	0.019	0.036	0.039	0.02	0.039	0.01
0.253	0.635	0.898	0.863	0.609	0.868	0.455	0.022	0.075	0.040
0.010	0.011	0.014	0.012	-0.035	0.066***	0.036	0.001	0.034**	-0.022
(0.023)	(0.027)	(0.015)	(0.018)	(0.027)	(0.018)	(0.027)	(0.009)	(0.016)	(0.013)
6,715	6,714	6,708	6,692	6,697	6,865	6,709	6,709	6,709	6,709
0.04	0.018	0.016	0.013	0.024	0.04	0.035	0.025	0.028	0.028
0.240	0.624	0.913	0.870	0.595	0.886	0.471	0.035	0.095	0.069

ich include age dummies, a linear term for year of birth, an indicator for survey wave and dummies for district of birth. Models for both sexes additionally control for birth*sex and surveywave*sex interactions. The instrument was defined as = 1 if YOB > 1980. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The as defined as 1 if respondent incorrectly answered any of the following four questions: whether HIV spreads via mosquitos, sharing a meal with an HIV+ person, due to ooking person can be HIV+. Those responding "don't know" to an HIV knowledge or HIV misconception question were accounted for as incorrect. Sample includes survey otswana, at least 18 years old at the time of the surveys, born in or after 1975, and had a valid HIV test result. No weights were used. Source: Botswana AIDS Impact Survey

Table S3: Intention-To-Treat Results: The Effect of the Education Reform on Sexual Intercourse, Marriage, Current Sexual and HIV Testing Behavior, HIV Knowledge and Attitudes, Labor Market Outcomes, and Literacy. (Panel D)

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
During the ast 4 weeks, have you discussed V/AIDS with other elative(s)? (1=yes, 0=no)	During the past 4 weeks, have you discussed HIV/AIDS with a friend? (1=yes, 0=no)	During the past 4 weeks, have you discussed HIV/AIDS with a co-worker? (1=yes, 0=no)	During the past 4 weeks, have you discussed HIV/AIDS with a healthcare worker? (1=yes, 0=no)	In the past 4 weeks, have you heard or seen any information about HIV/AIDS? (1=yes, 0=no)	Labor Force Participation (1=yes, 0=no)	Unemployed (1= yes, 0=no, missing if does not participate in labor force)	Indicator for housework (1=worked at home/househol d/homemaker, 0=other)	Can you read and understand a letter / newspaper / bible? (1=easily, 0=no or with difficulty)	Mentioned newspaper as a source of information about HIV/AIDS in past 4 weeks (1=yes, 0=no)
-0.013	0.009	-0.013	0.014	-0.011	0.108***	0.017	-0.043	0.002	0.041
(0.016)	(0.028)	(0.017)	(0.016)	(0.035)	(0.033)	(0.043)	(0.030)	(0.023)	(0.030)
3,791	3,791	3,791	3,791	3,774	3,942	2,458	3,925	3,962	3,778
0.020	0.013	0.04	0.019	0.018	0.078	0.128	0.021	0.035	0.015
0.062	0.157	0.084	0.057	0.649	0.654	0.416	0.261	0.865	0.199
0.014	0.025	0.013	-0.001	0.049	0.048	0.061	-0.033	0.085***	0.060*
(0.015)	(0.035)	(0.022)	(0.012)	(0.040)	(0.034)	(0.043)	(0.026)	(0.032)	(0.037)
2,918	2,918	2,918	2,918	2,909	3,037	2,239	3,031	3,051	2,911
0.025	0.018	0.058	0.012	0.026	0.156	0.158	0.034	0.029	0.012
0.023	0.244	0.111	0.031	0.662	0.842	0.219	0.095	0.825	0.214
-0.001	0.015	-0.002	0.007	0.014	0.082***	0.038	-0.039*	0.037*	0.049**
(0.011)	(0.022)	(0.013)	(0.010)	(0.026)	(0.024)	(0.030)	(0.021)	(0.019)	(0.023)
6,709	6,709	6,709	6,709	6,683	6,979	4,697	6,956	7,013	6,689
0.023	0.018	0.051	0.023	0.022	0.121	0.154	0.053	0.035	0.016
0.045	0.195	0.096	0.046	0.655	0.736	0.320	0.189	0.848	0.205

ıclude age dummies, a linear term for year of birth, an indicator for survey wave and dummies for district of birth. Models for both sexes additionally control irth*sex and surveywave*sex interactions. The instrument was defined as = 1 if YOB > 1980. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. were citizens of Botswana, at least 18 years old at the time of the surveys, born in or after 1975, and had a valid HIV test result. No weights were used. Source:) and III (2008).

Table S4: 2SLS Results: The Effect of Education on Sexual Intercourse, Marriage, Current Sexual and HIV Testing Behavior, HIV Knowledge and Attitudes, Labor Market Outcomes, and Literacy. (Panel A)

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Ever Had ntercourse =yes, 0=no)	Age at First Intercourse	First sex ever: Did you use anything to protect yourself (eg, condom)? (1=yes, 0=no)	Ever Married (1=yes, 0=no)	Years of Pre-marital Intercourse	Ever Given Birth (1=yes, 0=no)	How many times have you given birth? (never given birth=0)	Did you have sexual intercourse in the last 12 months? (1=yes, 0=no or never had intercourse)	Last sex: Did you or your partner use a condom, conditional on intercourse in the last 12 months? (1=yes, 0=no)	How old is your most recent partner, conditional on intercourse in the last 12 months? (years)
-0.007	0.761***	0.127***	0.021	-0.893***	-0.158***	-0.246**	-0.013	0.142**	-0.311
(0.027)	(0.261)	(0.047)	(0.025)	(0.304)	(0.057)	(0.107)	(0.035)	(0.068)	(0.538)
3,965	3,507	3,582	3,963	3,498	3,644	3,656	3,954	3,041	3,041
8.6	11.9	9.9	8.3	11.6	10.0	10.4	9.1	7.0	7.0
0.962	18.6	0.828	0.071	6.1	0.728	1.286	0.890	0.759	30.1
0.056*	0.065	0.055*	0.01	-0.084	-	-	0.058*	-0.019	-0.561
(0.030)	(0.209)	(0.028)	(0.009)	(0.212)	-	-	(0.033)	(0.036)	(0.385)
3,050	2,360	2,458	3,052	2,357	-	-	3,046	2,144	2,131
9.6	10.4	11.0	9.6	10.5	-	-	9.2	7.1	6.0
0.932	18.6	0.863	0.027	6.7	-	-	0.878	0.819	22.8
0.027	0.421***	0.090***	0.015	-0.493***	-	-	0.025	0.060*	-0.433
(0.019)	(0.161)	(0.026)	(0.012)	(0.175)	-	-	(0.023)	(0.032)	(0.329)
7,015	5,867	6,040	7,015	5,855	-	-	7,000	5,185	5,172
18.1	22.3	20.9	17.8	22.1	-	-	18.3	14.0	13.0
0.954	18.5	0.840	0.051	6.4	-	-	0.890	0.776	26.9

h exposure to the reform was used as an instrument for years of schooling. All models included the following controls: single-year age indicators, a linear term wave and indicators for district of birth. Models for both sexes additionally control for age*sex, districtofbirth*sex, yearofbirth*sex and surveywave*sex d as = 1 if YOB > 1980. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Years of Premarital Intercourse was defined as Age at First Marriage d) or Age at Time of Survey minus Age at First Intercourse (if never married). Sample includes survey respondents who were citizens of Botswana, at least 18 n in or after 1975, and had a valid HIV test result. No weights were used. Source: Botswana AIDS Impact Survey II (2004) and III (2008).

Table S4: 2SLS Results: The Effect of Education on Sexual Intercourse, Marriage, Current Sexual and HIV Testing Behavior, HIV Knowledge and Attitudes, Labor Market Outcomes, and Literacy: (Panel B)

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Ever had a exual partner who was 10 years older / younger, onditional on intercourse in the last 12 months? 1=yes, 0=no)	Indicator for 2 or more sexual partners in the last 12 months (1=two or more, 0=one or zero)	Time since first sexual intercourse with most recent partner, conditional on intercourse in the last 12 months (days)	Relationship to your most recent partner, conditional on intercourse in the last 12 months? (1=spouse, live-in partner, 0=not living in, other)	Have you ever heard of the virus HIV or an illness called AIDS? (1=yes, 0=no)	Have you ever been tested for HIV, the virus that causes AIDS? (1=yes, 0=no)	In the past 12 months have you been tested for HIV, the virus that causes AIDS? (1=yes, 0=no)	Indicator for knowledge of at least one HIV prevention strategy (1=yes, 0=no)	Indicator for knowledge of at least two HIV prevention strategies (1=yes, 0=no)	Mentioned condom use as an HIV prevention strategy (1=yes, 0=no)
0.003	0.044	8.1	-0.087	-0.009	0.110*	0.111	-0.021	-0.023	-0.044
(0.042)	(0.045)	(118.5)	(0.063)	(0.017)	(0.061)	(0.068)	(0.020)	(0.056)	(0.041)
3,364	3,658	2,953	3,057	3,874	3,793	3,787	3,791	3,791	3,795
9.7	6.3	8.0	6.8	8.3	7.7	7.6	7.9	7.9	7.8
0.232	0.115	1,289	0.441	0.987	0.720	0.531	0.949	0.698	0.881
-0.026	0.111**	-39.5	0.049	0.005	0.120**	0.075*	0.007	0.005	0.002
(0.026)	(0.051)	(82.0)	(0.045)	(0.012)	(0.052)	(0.044)	(0.020)	(0.043)	(0.031)
2,328	2,877	2,086	2,155	2,995	2,922	2,918	2,919	2,919	2,922
8.3	8.5	6.0	6.7	9.4	7.9	7.8	7.9	7.9	8.1
0.112	0.232	888	0.273	0.980	0.573	0.355	0.952	0.678	0.840
-0.011	0.082**	-14.6	-0.018	-0.001	0.115***	0.092**	-0.006	-0.009	-0.02
(0.025)	(0.034)	(72.8)	(0.035)	(0.010)	(0.040)	(0.039)	(0.016)	(0.035)	(0.025)
5,692	6,535	5,039	5,212	6,869	6,715	6,705	6,710	6,710	6,717
18.2	14.9	13.9	13.5	17.8	15.6	15.4	15.8	15.8	15.8
0.178	0.173	1,104	0.375	0.985	0.657	0.452	0.952	0.691	0.866

:h exposure to the reform was used as an instrument for years of schooling. All models included the following controls: single-year age indicators, a linear term for year nd indicators for district of birth. Models for both sexes additionally control for age*sex, districtofbirth*sex, yearofbirth*sex and surveywave*sex interactions. The 1980. The indicator for knowledge of HIV prevention strategies was defined as 1 if respondent could name either at least one (model 8) or two (model 9) out of the six ndoms, fewer partners, mutually faithful relationship, abstinence, avoid injections with contaminated needles, and avoid blood transfusions. Those responding "don't ere accounted for as incorrect. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Sample includes survey respondents who were citizens of Botswana, at veys, born in or after 1975, and had a valid HIV test result. No weights were used. Source: Botswana AIDS Impact Survey II (2004) and III (2008).

Table S4: 2SLS Results: The Effect of Education on Sexual Intercourse, Marriage, Current Sexual and HIV Testing Behavior, HIV Knowledge and Attitudes, Labor Market Outcomes, and Literacy. (Panel C)

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Mentioned mutually faithful elationship as an HIV prevention strategy (1=yes, 0=no)	Mentioned abstinence as an HIV prevention strategy (1=yes, 0=no)	Can people reduce their chances of getting HIV/AIDS by using a condom correctly every time they have sex? (1=yes, 0=no, don't know)	Can people reduce their chances of getting HIV/AIDS by having only one uninfected sex partner who has no other partners? (1=yes, 0=no,	Indicator for any misperceptions about HIV (1=yes; 0=no)	Do you think it should be acceptable for a woman to obtain male condoms? (1=yes, 0=no or not sure)	During the past 4 weeks, have you discussed HIV/AIDS with anyone? (1=yes, 0=no or not sure)	During the past 4 weeks, have you discussed HIV/AIDS with your spouse? (1=yes, 0=no)	During the past 4 weeks, have you discussed HIV/AIDS with your sexual partner? (1=yes, 0=no)	During the past 4 weeks, have you discussed HIV/AIDS with your family member(s)? (1=yes, 0=no)
0.073	0.034	0.012	-0.005	-0.062	0.092**	0.003	-0.003	0.094**	-0.022
(0.054)	(0.057)	(0.032)	(0.040)	(0.059)	(0.040)	(0.060)	(0.023)	(0.047)	(0.035)
3,794	3,793	3,791	3,780	3,782	3,873	3,791	3,791	3,792	3,792
7.7	7.7	7.8	7.7	7.2	8.7	7.7	7.7	7.7	7.7
0.254	0.628	0.928	0.874	0.564	0.933	0.484	0.044	0.142	0.083
-0.038	-0.002	0.026	0.038	-0.038	0.072**	0.087*	0.006	0.003	-0.035
(0.043)	(0.045)	(0.027)	(0.034)	(0.043)	(0.030)	(0.046)	(0.012)	(0.026)	(0.022)
2,921	2,921	2,917	2,912	2,915	2,992	2,918	2,918	2,918	2,918
8.0	8.1	7.7	7.2	7.6	9.6	8.2	8.2	8.2	8.2
0.246	0.634	0.902	0.869	0.603	0.875	0.471	0.023	0.075	0.034
0.014	0.015	0.019	0.017	-0.049	0.082***	0.048	0.002	0.045*	-0.029
(0.031)	(0.035)	(0.021)	(0.025)	(0.036)	(0.025)	(0.035)	(0.012)	(0.024)	(0.020)
6,715	6,714	6,708	6,692	6,697	6,865	6,709	6,709	6,710	6,710
15.8	15.7	15.4	14.8	14.7	18.3	15.8	15.8	15.8	15.8
0.244	0.628	0.918	0.875	0.582	0.905	0.484	0.036	0.107	0.061

ch exposure to the reform was used as an instrument for years of schooling. All models included the following controls: single-year age indicators, a linear term for year nd indicators for district of birth. Models for both sexes additionally control for age*sex, districtofbirth*sex, yearofbirth*sex and surveywave*sex interactions. The 1980. The indicator for any misperceptions was defined as 1 if respondent incorrectly answered any of the following four questions: whether HIV spreads via mosquitos, e to witchcraft, and whether a healthy looking person can be HIV+. Those responding "don't know" to an HIV knowledge or HIV misconception question were ors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Sample includes survey respondents who were citizens of Botswana, at least 18 years old at the time of the surveys, IV test result. No weights were used. Source: Botswana AIDS Impact Survey II (2004) and III (2008).

Table S4: 2SLS Results: The Effect of Education on Sexual Intercourse, Marriage, Current Sexual and HIV Testing Behavior, HIV Knowledge and Attitudes, Labor Market Outcomes, and Literacy. (Panel D)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	During the past 4 weeks, have you discussed HIV/AIDS with other relative(s)? (1=yes, 0=no)	During the past 4 weeks, have you discussed HIV/AIDS with a friend? (1=yes, 0=no)	During the past 4 weeks, have you discussed HIV/AIDS with a co-worker? (1=yes, 0=no)	During the past 4 weeks, have you discussed HIV/AIDS with a healthcare worker? (1=yes, 0=no)	In the past 4 weeks, have you heard or seen any information about HIV/AIDS? (1=yes, 0=no)	Labor Force Participation (1=yes, 0=no)	Unemployed (1= yes, 0=no, missing if does not participate in labor force)	Indicator for housework (1=worked at home/household/homemaker, 0=other)	Can you read and understand a letter / newspaper / bible? (1=easily, 0=no or with difficulty)	Mentioned newspaper as a source of information about HIV/AIDS in past 4 weeks (1=yes, 0=no)
	-0.022	0.015	-0.022	0.023	-0.019	0.172**	0.038	-0.070	0.003	0.069
	(0.028)	(0.046)	(0.030)	(0.027)	(0.060)	(0.076)	(0.100)	(0.049)	(0.036)	(0.049)
	3,791	3,793	3,791	3,791	3,774	3,942	2,458	3,925	3,962	3,778
	7.7	7.7	7.7	7.7	7.6	8.4	2.8	8.2	8.7	7.6
	0.055	0.163	0.077	0.065	0.643	0.706	0.433	0.239	0.866	0.222
	0.015	0.026	0.014	-0.001	0.052	0.048	0.069	-0.033	0.084***	0.064*
	(0.017)	(0.037)	(0.023)	(0.012)	(0.041)	(0.039)	(0.056)	(0.027)	(0.023)	(0.038)
	2,918	2,920	2,918	2,918	2,909	3,037	2,239	3,031	3,051	2,911
	8.2	8.0	8.2	8.2	8.2	9.3	5.4	9.3	9.7	8.3
	0.026	0.249	0.114	0.031	0.671	0.846	0.235	0.092	0.831	0.224
	-0.002	0.021	-0.003	0.01	0.019	0.105***	0.058	-0.050*	0.046**	0.066**
	(0.015)	(0.029)	(0.018)	(0.014)	(0.034)	(0.039)	(0.051)	(0.027)	(0.019)	(0.031)
	6,709	6,713	6,709	6,709	6,683	6979	4,697	6,956	7,013	6,689
	15.8	15.7	15.8	15.8	15.7	17.7	8.2	17.5	18.4	15.9
	0.045	0.201	0.095	0.049	0.660	0.757	0.340	0.178	0.857	0.223

ch exposure to the reform was used as an instrument for years of schooling. All models included the following controls: single-year age indicators, a linear term y wave and indicators for district of birth. Models for both sexes additionally control for age*sex, districtofbirth*sex, yearofbirth*sex and surveywave*sex ed as = 1 if YOB > 1980. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Sample includes survey respondents who were citizens of Botswana, at rveys, born in or after 1975, and had a valid HIV test result. No weights were used. Source: Botswana AIDS Impact Survey II (2004) and III (2008).

Table S5: Sensitivity Analyses: Assessing the Robustness of the 2SLS Results to: Additional Controls, Sample Restrictions, Sample Weights, Imputed HIV Estimates, and a 2SLS Difference-in-Differences Strategy (Panel A)

sitive	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)[¥]	(12)
riable												
	-0.116**	-0.093	-0.115	-0.140*	-0.123*	-0.112	-0.279	-0.339	-0.113*	-0.141	-0.130*	-0.125**
	(0.058)	(0.065)	(0.081)	(0.079)	(0.068)	(0.084)	(0.303)	(0.675)	(0.060)	(0.080)	(0.070)	(0.063)
	✓	-	[illegible]	✓	[illegible]	✓	✓	[illegible]	✓	✓	✓	✓
	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
irth	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	-
ghedu9,	-	-	-	-	-	-	-	-	-	-	-	✓
	-	✓	-	-	-	-	-	-	-	-	-	-
, c.Age2, 5, c.Agek30	-	-	✓	-	-	-	-	-	-	-	-	-
	-	-	-	✓	-	-	-	-	-	-	-	-
	✓	✓	✓	-	-	✓	✓	✓	✓	-	-	-
	-	-	-	-	✓	-	-	-	-	-	-	-
	-	-	-	-	-	✓	-	-	-	-	-	-
ws												
	✓	✓	✓	✓	✓	✓	-	-	✓	✓	✓	✓
	-	-	-	-	-	-	✓	-	-	-	-	-
	-	-	-	-	-	-	-	✓	-	-	-	-
	-	-	-	-	-	-	-	-	✓	-	-	-
	-	-	-	-	-	-	-	-	-	✓	✓	-
	3,965	3,965	3,965	3,965	3,965	3,965	3,240	2,846	3,965	5,549	4,866	3,965
	8.3	6.0	4.6	5.3	0.2	3.9	0.9	0.2	7.7	n/a	n/a	0.1

ls, in which exposure to the reform was used as an instrument for years of schooling. Model 3 controls for a third order spline in age (with knots at 20, 25, 9). The instrument was defined as = 1 if YOB > 1980. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. ¥ Uses additional covariates, Age at rried, to impute HIV status. Sample includes survey respondents who were citizens of Botswana, at least 18 years old at the time of the surveys, and had a 3otswana AIDS Impact Survey II (2004) and III (2008).

Table S5: Sensitivity Analyses: Assessing the Robustness of the 2SLS Results to: Additional Controls, Sample Restrictions, Sample Weights, Imputed HIV Estimates, and a 2SLS Difference-in-Differences Strategy (Panel B)

itive	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)¥	(12)
riable												
	-0.050*	-0.050	-0.038	-0.035	-0.033	-0.031	-0.029	-0.034	-0.043	-0.056	-0.053	-0.094
	(0.029)	(0.035)	(0.026)	(0.024)	(0.023)	(0.027)	(0.033)	(0.052)	(0.036)	-0.052	-0.047	(0.173)
	✓	-	-	✓	✓	✓	✓	✓	✓	✓	✓	✓
	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
irth	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	-
ghedu9,	-	-	-	-	-	-	-	-	-	-	-	✓
	-	✓	-	-	-	-	-	-	-	-	-	-
, c.Age2, ;, c.Agek30	-	-	✓	-	-	-	-	-	-	-	-	-
	-	-	-	✓	-	-	-	-	-	-	-	-
	✓	✓	✓	-	-	✓	✓	✓	✓	-	-	-
	-	-	-	-	✓	-	-	-	-	-	-	-
	-	-	-	-	-	✓	-	-	-	-	-	-
ws												
	✓	✓	✓	✓	✓	✓	-	-	✓	✓	✓	✓
	-	-	-	-	-	-	✓	-	-	-	-	-
	-	-	-	-	-	-	-	✓	-	-	-	-
	-	-	-	-	-	-	-	-	✓	-	-	-
	-	-	-	-	-	-	-	-	-	✓	✓	-
	3,053	3,053	3,053	3,053	3,053	3,053	2,472	2,141	3,053	4,490	3,415	3,053
	8.8	6.6	10.6	13.1	0.7	10.0	6.0	2.1	5.8	n/a	n/a	0.5

ls, in which exposure to the reform was used as an instrument for years of schooling. Model 3 controls for a third order spline in age (with knots at 20, 25, 9). The instrument was defined as = 1 if YOB > 1980. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. ¥ Uses additional covariates, Age at rried, to impute HIV status. Sample includes survey respondents who were citizens of Botswana, at least 18 years old at the time of the surveys, and had a 3otswana AIDS Impact Survey II (2004) and III (2008).

Table S6: Intention-To-Treat Results: The Effect of The Reform on No Consent to HIV Test.

	(1)	(2)	(3)
Subsample	**Female**	**Male**	**Both Sexes**
Dependent Variable	**NoHIVConsent**	**NoHIVConsent**	**NoHIVConsent**
Predictor			
Reform Indicator	0.030	0.044	0.036*
	(0.027)	(0.032)	(0.021)
Observations	5,442	4,395	9,837
R-squared	0.017	0.015	0.017
Mean Dependent Variable, Pre-Reform	0.279	0.290	0.284

All regressions are OLS models, which include age dummies, a linear term for year of birth, an indicator for survey wave and dummies for district of birth. Model 3 additionally controls for age*sex, districtofbirth*sex, yearofbirth*sex and surveywave*sex interactions. The indicator for no HIV consent was defined as 1 if no HIV test result was available, excluding invalid results (eg, indeterminate result, insufficient volume of blood withdrawn). Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. No weights used. Sample includes survey respondents who were citizens of Botswana, at least 18 years old at the time of the surveys, and born in or after 1975. Source: Botswana AIDS Impact Survey II (2004) and III (2008).

Table S7: Intention-To-Treat Results: Using Logistic Regression and SE's Clustered at the YOB Level.

	(1)	(2)	(3)	(4)	(5)	(6)
Coefficient on Reform Indicator	**Female**	**Male**	**Both Sexes**	**Female**	**Male**	**Both Sexes**
Dependent Variable						
HIV-Positive	-0.050*	-0.049*	-0.049**	-0.074***	-0.050***	-0.064***
	(0.030)	(0.027)	(0.021)	(0.023)	(0.014)	(0.017)
Observations	3,965	3,053	7,018	3,965	3,053	7,018
R-squared	-	-	-	0.095	0.069	0.122
Proportion HIV positive, Pre-Reform	0.316	0.171	0.254	0.323	0.168	0.255

Models 1 to 3 are logistic regression models (marginal effects reported). Models 4-6 are OLS models clustering standard errors at the YOB level. All models include age dummies, a linear term for year of birth, an indicator for survey wave and dummies for district of birth. Models 3 and 6 additionally control for age*sex, districtofbirth*sex, yearofbirth*sex and surveywave*sex interactions. The instrument was defined as = 1 if YOB > 1980. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Sample includes survey respondents who were citizens of Botswana, at least 18 years old at the time of the surveys, born in or after 1975, and had a valid HIV test result. No weights were used. Source: Botswana AIDS Impact Survey II (2004) and III (2008).

Table S8: Intention-To-Treat Results: The Effect of the Reform on an Indicator for Age at First Intercourse by Age 16, 18, 20, or 22.

Dependent Variable	**(1)** Indicator for Age at First Intercourse ≤ 16 (1=yes, 0=no)	**(2)** Indicator for Age at First Intercourse ≤ 18 (1=yes, 0=no)	**(3)** Indicator for Age at First Intercourse ≤ 20 (1=yes, 0=no)	**(4)** Indicator for Age at First Intercourse ≤ 22 (1=yes, 0=no)
Coefficient on Reform Indicator				
Female	-0.091***	-0.116***	-0.083***	-0.018
	(0.029)	(0.036)	(0.025)	(0.012)
Observations	3,507	3,507	3,507	3,507
R-squared	0.041	0.104	0.073	0.053
Mean Dependent Variable, Pre-Reform	0.187	0.544	0.850	0.963
Male	0.027	-0.039	-0.046	0.001
	(0.041)	(0.044)	(0.032)	(0.020)
Observations	2,360	2,360	2,360	2,360
R-squared	0.046	0.105	0.108	0.080
Mean Dependent Variable, Pre-Reform	0.228	0.491	0.789	0.931
Both Sexes	-0.044*	-0.085***	-0.068***	-0.011
	(0.024)	(0.028)	(0.020)	(0.011)
Observations	5,867	5,867	5,867	5,867
R-squared	0.047	0.105	0.092	0.071
Mean Dependent Variable, Pre-Reform	0.205	0.522	0.824	0.949

All regressions are OLS models, which include age dummies, a linear term for year of birth, an indicator for survey wave and dummies for district of birth. Models for both sexes additionally control for age*sex, districtofbirth*sex, yearofbirth*sex and surveywave*sex interactions. The instrument was defined as = 1 if YOB > 1980. Standard errors in parentheses. *** p<0·01, ** p<0·05, * p<0·1. Sample includes survey respondents who were citizens of Botswana, at least 18 years old at the time of the surveys, born in or after 1975, and had a valid HIV test result. No weights were used. Source: Botswana AIDS Impact Survey II (2004) and III (2008).