The effect of increased primary schooling on adult women's HIV status in Malawi and Uganda: Universal Primary Education as a natural experiment

Working Paper: Please do not cite without author's permission

Julia Andrea Behrman

Doctoral Candidate
Department of Sociology
New York University
295 Lafayette Ave, 4th floor
New York, NY 10012
Jab965@nyu.edu
+001 610.389.4207

Acknowledgement: I am grateful to Jere Behrman, Amber Peterman, Florencia Torche, Paula England and Jennifer Hill for helpful commentary on earlier versions of this paper. Support for this study was provided by the grant *Team 1000+ Saving Brains: Economic Impact of Poverty-Related Risk Factors for Cognitive Development and Human Capital* "0072-03" provided to the Grantee, The Trustees of the University of Pennsylvania by Grand Challenges Canada.

Abstract: This paper explores the causal relationship between primary schooling and adult HIV status in two East African countries with some of the highest HIV infection rates in the world. Using data from the most recent Demographic Health Surveys in Malawi (2010) and Uganda (2011), the paper takes advantage of a natural experiment, the implementation of Universal Primary Education policies in the mid 1990s. An instrumented fuzzy regression discontinuity approach is used to model the relationship between increased primary schooling and adult HIV status. The results indicate that in Malawi a one year increase in schooling for a girl leads to a 6-7 percent reduction in probability of testing positive for HIV as an adult and in Uganda a one year increase in schooling leads to a 2-4 percent reduction in probability of testing positive for HIV as an adult. These results are robust to a variety of model specifications. In a series of supplementary analyses a number of potential pathways through which such effects may occur are explored. Findings indicate increased exposure to primary school affects overall schooling attainment and effects adolescent sexual behavior to some extent. However primary schooling has no effect on recent (adult) sexual behavior.

Keywords: HIV/AIDS, sub-Saharan Africa, schooling, Universal Primary Education, risky sexual behavior, natural experiment, regression discontinuity

Introduction

Over 35 million people infected with the human immunodeficiency virus (HIV) currently live in sub-Saharan Africa (UNAIDS 2013). Given the magnitude of the epidemic there has been extensive focus on how to reduce infection rates for adolescents and young adults between the ages of 15 and 24, a group identified as particularly at risk for new infection (UNAIDS 2008). The literature suggests that adolescent African females face a special set of vulnerabilities. In part this is attributed to social norms that put women, particularly young women, at greater risk, for example the prevalence of relationships between older males and much younger females. In part the gender difference in HIV rates also reflects the fact that women are biologically more susceptible to contracting the virus through sexual intercourse (Padian et al. 1991). In Africa today women comprise over 60 percent of HIV infections, a dramatic shift from the early days of the epidemic in Africa when infection rates were higher in men (UNAIDS 2013).

Young females face heightened risk for infection just as the point when they leave school, thus increasing girls schooling has been put forward as a mechanism to reduce HIV transmission. It is argued that more schooling will enable young women to lessen the possibility of transmission through greater knowledge of risks and more capacity for reducing risks through condom use or other preventative behavior. The focus on schooling in the HIV literature corresponds with a more general trend in the global development community on expanding access to schooling for children throughout the developing world. Perhaps most visibly, the second Millennium Development goal calls for universal completion of primary school and gender equity in school access (UN 2001).

In this paper I explore the causal relationship between girls' primary schooling and their adult HIV status in Malawi and Uganda using the implementation of Universal Primary Education (UPE) policies in the mid-nineties as natural experiments. Using data from the most recent Demographic Health Survey (DHS) in Uganda and Malawi, I model the relationship between primary schooling and adult HIV status using a fuzzy regression discontinuity approach.

While there has been substantial interest in the potential relationship between schooling and HIV status, very few studies from developing countries have been able to infer causality. I further contribute to the literature by focusing on the relationship between primary, as opposed to secondary, schooling and later outcomes. Studies focusing exclusively on the effect of secondary school on HIV status likely suffer from selection bias since it is a different, potentially less vulnerable population, who is able to attend secondary school in the first place. Furthermore, I am able to extend existing studies, most of which have relatively short follow-up periods, by exploring the effect of schooling on adult HIV status. Comparing the experiences of Uganda and Malawi is substantively interesting and is important for establishing the external validity of findings given that the trajectory of the epidemic took different forms in each country.

Literature review

In a comprehensive review of the literature Jukes et al. (2008) distinguish between two distinct pathways through which formal schooling can affect HIV infection. The first pathway concerns the consequences of cognitive and social learning processes while the second pathway is related to the consequences of school attendance regardless of learning. With respect to the first pathway, evidence on whether changes in knowledge and attitudes about HIV risk lead to behavioral change is mixed. In Tanzania, an adolescent sexual health program in 20 communities was found to significantly affect knowledge and attitudes about HIV but did not reduce HIV in either the short or long term (Ross et al 2007; Doyle et al. 2010). On the other hand, a randomized experiment in Kenya finds the provision of information on relative risk by age led to a reduction in teen pregnancy (a proxy for unprotected sex) and a shift from older to younger partners (Dupas 2011). However, the same study found the government's HIV risk campaign, which focused on abstinence as a strategy for risk reduction, had no effect on behavior, thus indicating that the type of information or messaging may matter for behavior change.

Alternatively, as Jukes et al. (2008) also suggest, schooling may not affect learning but may nonetheless play a protective role simply by delaying the onset of sexual activity, marriage and/or childbearing, all of which put women at risk for infection. In a randomized experiment in Western Kenya girls who received an educational subsidy for the last three years of primary school were significantly less likely to drop out of primary school, marry, or start childbearing even following the end of the subsidy (Duflo, Dupas, Kremer 2012). A number of emerging studies use educational policy change in Africa to look at schooling and early fertility outcomes. Ozier (2011) uses performance on the national Kenyan secondary school entrance exam to conduct a regression discontinuity analysis, finding that secondary schooling leads to a reduction in adolescent pregnancy for females. Also in Kenya, Ferre (2007) uses a 1995 education reform as an instrument for years of schooling and finds a one-year increase in schooling leads to an 8 percent decrease in the probability of adolescent pregnancy. In nearby Ethiopia, Pradhan and Canning (2013) use elimination of school fees as a natural experiment. Using a regression discontinuity approach, they find each additional year of schooling lowers the probability of adolescent pregnancy by 6 percent. All of these findings are particularly interesting given that in the early years of the epidemic higher socio-economic status, potentially including schooling attainment, was associated with riskier sexual behavior for men and also to some extent for women (Lopman et al. 2007; Weltz et al. 2007; Mishra et al. 2007).

The association between schooling and HIV in the African context has been documented in a number of cross-sectional studies (Glynn et al. 2004; Pettifor et al. 2005; de Walque 2007; Hargeeves et al. 2008). However, provision of credible evidence of a causal relationship between schooling and HIV status is limited to two recent experimental studies, to the best of my knowledge. In a cluster randomized trial of a cash transfer program designed to keep adolescent girls (ages 13-22) in school Baird and colleagues (2012) find HIV prevalence in the combined treatment group, which included conditional and unconditional cash transfer beneficiaries, to be 1-2 percent as compared to 3 percent in the control group at the end of an 18-month exposure period. Baird and colleagues attribute their findings to changes in respondent (self-reported) sexual behavior. They find the transfers decrease frequency of sexual activity and age of sexual partner with program beneficiaries choosing younger partners. Interestingly, the program does not affect other aspects of risky sexual behavior, including age of sexual debut or unprotected sex.

In another recent study involving causal inference, Duflo and colleagues (2012) evaluate the effects of a randomized HIV intervention amongst 6th graders (approximately age 13) from 328 schools in Western Kenya over a seven-year period. The intervention compared three approaches to behavior changes; (1) educational subsidies; (2) an HIV information campaign focused on abstinence; or (3) a combination of both approaches. The second approach was found to be least effective, failing to reduce pregnancy or sexually transmitted infections, including HIV and HSV-2. The first approach reduced adolescent girls' rates of school exit, pregnancy and marriage but not the rates of sexually transmitted infection. The third approach had larger effects on reductions in sexually transmitted infections, but less pronounced effects on schooling and self-reported sexual behavior than the first approach. Additional exploration of the type of partnerships girls pursue led Duflo and colleagues to conclude that pregnancy and STI infection should not be viewed as determined solely by unprotected sex, but are also related to the level of commitment of adolescent relationships.

Universal Primary Education in Malawi & Uganda

The primary school systems in both Malawi and Uganda follow the legacy of the British colonial system. In both countries primary school is officially supposed to run from ages 6 to 13, though as in much of Africa grade repetition is common and many children start late and end early. Both countries were early implementers of Universal Primary Education Policies aimed at increasing primary school enrolment principally through the elimination of school fees. In 1991 the primary school Gross Enrollment Ratio (GER), the number of children who are actually in school divided by the number of children who are of school age, was 71 percent for Uganda and 75 percent for Malawi (UNESCO 2011a; UNESCO 2011b). By 2011 these figures had risen to 110 percent for Uganda and 140 percent for Malawi (ibid). The percent for this ratio can be greater than 100 because students outside of the primary school age range may still be in school.

Starting in 1991 Malawi adopted a sequential approach to elimination of school fees which entailed provision of free waivers for grade one in the first year of implementation and for grade two in the second year (Kattan 2006; World Bank 2009). In 1994 this plan was replaced by a "big bang" approach in which primary school fees were eliminated across all grades effective starting September 1994. The government also eliminated "indirect" fees, such as contributions to school development funds, and the requirement to wear uniforms to school. The Ministry of Education, the agency charged with carrying out implementation, launched a mass media campaign in July of 1994 to ensure that the public was aware of the policy. In addition to elimination of fees a number of other changes were put into place to deal with the influx of new students; local languages become the dominant language of instruction for early grades, 20,000 new teachers were recruited and received expedited training and the budgetary allocation to the education sector significantly increased. In the year following the 1994 "big bang" implementation 1 million new children entered primary school, an increase in enrollment of over 50 percent (ibid). Because entry was allowed at any grade increases were actually largest in the older grades. In the final grade of schooling (standard 8) enrollment increased by 79 percent (ibid).

In contrast to Malawi's initial sequential approach, Uganda adopted a "big bang" approach to elimination of school fees from the start (Deininger 2003; Grogan 2006; Kattan 2006). All school fees were eliminated at all levels for all primary school children effective January 1st 1997. The policy eliminated both direct school fees and indirect school fees such as donations to parent's funds. School uniforms, another impediment to school access for the poor, were also eliminated. The implementation of this policy was a highly decentralized process with individual districts across the country taking charge of carrying out implementation. A large nation-wide information dissemination campaign was launched and included target messages specifically focused on educating girls. In an evaluation of the program Deininger (2003) finds evidence that UPE led to dramatic increases in primary school attendance and reduced gender and income inequities in primary school attendance across the country.

In both countries an important critique of UPE policies is that improving the access to education came at the expense of deteriorating quality of primary schooling. In both contexts the student-to-teacher ratio rose fairly significantly and critics lamented that new teachers were poorly trained or unprepared (Deininger 2003; Kadzamira and Rose 2003). In some regions the government did not follow through on promise to provide materials and parents were forced to shoulder this cost (ibid). All of this has important implications for what transpires in the classroom. Grogan (2006) finds that UPE policies in Uganda increased the probability of on-time entry, but decreased the probability of a child being able to complete simple reading tests.

Data

Data for this paper comes from the 2010 round of the Malawi Standard Demographic Health Survey (DHS) and the 2011 round of the Uganda AIDS Indicator Survey (AIS). The DHS and AIS are nationally representative cross-sectional household-based surveys collected by ICF International in collaboration with host country governments. Standardized questionnaires facilitate cross-national comparisons.

In the 2010 Malawi survey a nationally representative cross-sectional sample of 27,345 households was interviewed from urban and rural areas. In one third of these households, HIV tests were also administered to all men and women in the household between the ages of 15 and 49. Thus the full female HIV sample for Malawi consists of 19,363 women, of whom nine percent (n=1,834) are HIV positive. Five respondents were excluded due to indeterminate test results. My analysis focuses on a sub-sample of women born between 1974 and 1984 (n=2,512) of whom sixteen percent (n=421) are HIV positive (Table 1).

The Uganda AIS collects much of the same demographic information as in the standard DHS survey, but with an additional focus on sexually transmitted infections (STIs). All women and men ages 15-49 in the household are eligible to be interviewed and tested as part of the AIS. The full Uganda sample includes 11,967 women of whom eight percent (n=944) are HIV positive. One respondent was excluded due to indeterminate test results. My analysis focuses on a sub-sample of women born between 1977 and 1987, of whom nine percent (n=362) are HIV positive (Table 1). In both countries test refusal rates were very low (less than one percent) in the full sample of men and women. Thus there is minimal concern about non-response bias.

I focus on women in this analysis due to their increased biological and social vulnerability to HIV. Indeed, in the full sample of HIV positive respondents in Malawi 63 percent (n=890) are female and 37 percent (n=530) are male. Similarly in the full Uganda sample 63 percent (n=944) of HIV positive respondents are female and 37 percent (n=551) are male. These patterns hold for the sub-samples of interest as well. In addition to increased biological vulnerability, HIV infection in women also poses a risk for the transmission of infection to the next generation via mother to child transmission. Furthermore female children have historically faced disadvantages in access to school and stood to gain the most from elimination of school fees. Thus we would expect the jump in schooling to be higher for girls than boys.

[Insert Table 1]

Empirical Strategy

In this paper I use a regression discontinuity design (RDD) to assess the causal effect of primary schooling on adult HIV status. The RDD is a special case of an observational study design in which assignment to *treatment (exposure to UPE)* is decided solely based on values of one measured variable, *Z (birth cohort)*, with all values of *Z* on one side of a specified cutoff, *c (13* *and younger at policy implementation)*, given the treatment and all observations with values on the other side denied the treatment. The RDD takes advantage of the fact that girls just above primary school age and just below primary school age at the year of policy implementation will be comparable on both observed and unobserved characteristics and differ only in their exposure to the UPE policy. Thus, girls 13 and younger may be able to continue or extend their schooling thanks to elimination of fees, but girls over 13 will not have this opportunity. This empirical strategy offers a number of advantages over Ordinary Least Squares (OLS) or logistic regression. It is likely that there are characteristics such as socio-economic status, cognitive ability or personal preferences that predict both schooling attainment and likelihood of avoiding HIV infection. If so, then OLS or logistic regression estimates overstate the schooling impact because schooling is partially proxying for these omitted characteristics (Table A1).

However, in East Africa grade repetition is common and it is not unusual for children to start school late or end school early. Some girls who are beyond primary school age will nonetheless be exposed to UPE because they are still in primary school when UPE is introduced. Likewise, some girls who are primary school age will not attend primary school even with the elimination of school fees because truancy is not enforced, child labor is often needed for household livelihood activities, or education for girls is not valued by households. Graphically it is possible to see some evidence of non-compliance, particularly in Malawi (Figures 1 & 2). This non-compliance with treatment assignment means this is a fuzzy, as opposed to a clean, RDD: thus I introduce the use of an instrumental variable in the RDD analysis. I use two stage least squares regression where the outcome is a dichotomous indicator of HIV status at survey and the treatment is a continuous variable that measures number of years of school completed. The treatment is instrumented using a dichotomous indicator of exposure to UPE while age 13 or younger. In Malawi exposure is assigned to girls born in 1981 and later and in Uganda exposure is assigned to girls born in 1984 and later.

In equation (1), the first stage, I regress *D*, the treatment, on *Z*, the randomly assigned instrument. In equation (2), the second stage, I regress *Y*, the outcome, on the predicted value of *D* from the first stage. Due to concerns about collapsibility I use linear models. Given that both Malawi and Uganda are religious and ethnically heterogeneous countries, I control for religion and ethnicity. Religion is represented by four indicator variables for (1) Christian (non-Catholic); (2) Catholic; (3) Muslim; and (4) Other religion. I use Christian (non-Catholic) as the reference category. Uganda and Malawi have over 20 different ethnicities recorded in the DHS and AIS. To capture ethnicity I create an indicator to represent whether the respondent is a member of the "dominant" ethnic group of each country, or the group with the largest response rate. In Uganda this is the Baganda tribe and in Malawi this is the Chewa tribe. I do not control for current levels of household wealth or any other post-treatment variables as they may be endogenous to the model.

(1) $D_i = \alpha_0 + \alpha_1 Z_i + \ldots \alpha_k X_k + v_i$

(2) $Y_i = \beta_0 + \beta_1 D_i + \cdots \beta_k X_k + \mathcal{E}_i$

The estimand of interest for this analysis is the complier average causal effect (CACE), a special case of the Local Average Treatment Effect. In equation (3) *Z* denotes the randomly assigned instrument, *Y(0*) and *Y(1)* stand for potential outcomes, *Y* denotes the observed outcome, *D(0)= D(Z=0) and D(1)=D(Z=1)* denote potential "program exposure" under each assignment of the instrument, *D* denotes observed participation behavior (in other words the treatment)..

(3) $$\frac{E[Y(Z=1)-Y(Z=0)]}{E[D(Z=1)-D(Z=0)]}$$

Inference is made over the interval of birth cohorts specified for compliers, who are girls who would have adhered to their treatment assignment no matter which side of the age threshold they fell on. I must assume that there are no additional pathways through which exposure to UPE could affect HIV status. That is to say, I must assume that Z_i is uncorrelated with both $\mathcal{E}_i$ and v_i.

If this is not the case, then it will be impossible to know whether schooling, as opposed to these alternative pathways, affects HIV status. This approach relies heavily on the assumption of ignorability of the instrument, in other words that the instrument, exposure to UPE, is randomly assigned. The validity of this assumption depends considerably on the width of the interval of birth cohorts I consider around the specified cutoffs of age at exposure. Correctly specifying that this interval is adequately narrow is important to ensure treatment and control groups are comparable and differ only in their extent of exposure to UPE. I start with an interval of approximately 5 years above and below the cutoff to provide a relatively large sample with a range of differences in exposure that is nonetheless fairly comparable. I then consider two narrower intervals of birth cohorts: approximately 3 years above and below the cutoff and approximately 2 years above or below the cutoff.

In order to use UPE exposure as an instrument for years of schooling UPE must actually increase primary school attendance, in other words $cov(Z_i, D_i) \neq 0$. Since I have a single instrument and single endogenous regressor the instrument can be considered relevant if the t-value for the instrument is larger than 3.2 or the corresponding p-value below 0.0016 or the F for the excluded instrument is greater than 10 (Stock et al. 2002). I test this empirically in the first stage and find results to be highly significant at the $p< 0.001$ level (Tables 2 & 3).

To account for the fact that increasing attendance may not be tantamount to increasing schooling attainment if grade repetition is common, I re-run all estimates using a dichotomous indicator of primary school completion as the outcome (Tables A2 & A3). I also explore the possibility that UPE may most effect girls at the start, as opposed to the end, of their schooling career. I run a separate second set of analyses for a younger sub-sample of girls using a different instrument that indicates exposure to UPE for the entire schooling career (ages 7 or younger) as opposed to exposure for only part of the schooling career (ages 8 and older). First stage results indicate this is a weak instrument thus I do not pursue this strategy (results available upon request from the author).

One potential concern with this approach is that children may be HIV positive prior to starting school and may leave school early because they are ill or because families are less interested in investing in schooling for sick children. Given the trajectory of the epidemic, rates of mother-to-child transmission should be low for children born in the mid 1970s to the mid 1980s, though it is difficult to know precisely due the paucity of data for this period. Nonetheless, the poor quality of treatment in both countries well into the 21st century means that it would have been nearly impossible for an HIV positive child born in the 1980s to survive to be interviewed as an adult in 2010/2011. Thus for the women in this sample the prime pathway of HIV transmission should be through sexual intercourse which, barring instances of childhood sexual abuse, would occur in adolescence or adulthood. In the Malawi sample the average age of first intercourse is 16.65 (SD 2.6) and in the Uganda sample the average age of first intercourse is 16.83 (2.7). Amongst the treated sample, I find very small numbers of girls who report sexual initiation prior to exposure to the policy (n=12 for Malawi and =15 for Uganda). I run analyses including and excluding these respondents and find no difference in results.

First-stage results

In the first stage I regress years of schooling completed on exposure to UPE. In Malawi, I start with the interval of birth cohorts 1974-1984 and find girls exposed to UPE have an average of one more year of schooling than girls not exposed (Table 2 column 2). When I limit this to the interval 1978-1982 exposed girls have an average of .76 more years of school than non-exposed girls (Table 2 column 4). Both of these results are highly significant at the $p<0.001$ level and represent fairly large jumps in a country where overall levels of schooling are low. A graphical representation of this discontinuity can be seen in Figure 1. In Uganda, I start with the interval of birth cohorts 1977-1987 and find girls exposed to UPE have an average of 1.5 more years of schooling than girls not exposed (Table 3 column 2). For interval 1982-1986, exposed girls have an average of .96 more years of school then girls who are not exposed. Both results are highly significant at the p<.001 level. A graphical representation of this discontinuity can be seen in Figure 2.

[Insert Table 2]

[Insert Table 3]

Second-stage results

In the second stage, I regress adult HIV positive status on the predictions from the first stage. In Malawi, the effect (-0.06) is negative and highly significant at the p<0.001 level for the birth cohort interval 1974-1984 (Table 4 column 2). When I re-estimate this relationship using a smaller interval of birth cohorts (1976-1983) the effect size remains unchanged with significance at the p<0.01 level (Table 4 column 3). I limit the birth cohort intervals even further to 1978-1982 and find the effect (-0.07) is significant at the p<0.05 level (Table 4 column 4). This indicates that a one year increase in schooling for a girl in Malawi leads to a 6-7 percent reduction in probability of testing positive for HIV as an adult as compared to if she had not attended for that year.

In Uganda the effect is also negative (-0.02) and is highly significant at the p<0.001 level for the birth cohort intervals 1977-1987 (Table 5 column 2). When I limit this to a smaller interval of birth cohorts (1980-1986) the magnitude of the effect remains unchanged with statistical significance at the p<0.01 level (Table 5 column 3). When I further limit the interval to 1982-1986 the effect becomes -0.04 and is significant at the p<0.01 level (Table 5 column 4). These results indicate that a one year increase in schooling for a girl in Uganda leads to a 2-4 percent reduction in probability of testing positive for HIV as an adult as compared to if she had not attended the extra year.

[Insert Table 4]

[Insert Table 5]

Extensions: Exploration of impact pathways

I explore a number of pathways through which schooling could affect HIV status using the same strategy to make causal inference. I start by looking to see if girls who are exposed to more primary school are actually achieving at higher levels as opposed to just attending longer and potentially repeating grades. I look at the effect of increased primary schooling on: (1) completion of primary school, which in both contexts entails passing a national examination; and (2) attendance of secondary school.

In Malawi there is a positive effect (0.07) of primary schooling on primary school completion that is significant at the $p<.001$ level (Table A2 column 1). Additionally, there is a positive effect (.06) of primary school on secondary school attendance that is significant at the $p<.001$ level (Table A2 column 2). These results indicate that a one year increase in primary schooling for a girl in Malawi leads to 7 percent increasing in probability of completing primary school and a 6 percent increase in the probability of attending secondary as compared to if she had not attended the extra year. In Uganda, there is also a positive effect (0.1) of primary schooling on completion of primary school that is significant at the $p<.001$ level (Table A3 column 1). Likewise, there is a positive effect (.08) of primary schooling on completion of secondary school that is significant at the $p<.001$ level (Table A3 column 2). These results indicate that a one year increase in primary schooling for a girl in Uganda leads to 10 percent increase in the probability of completing primary school and a 8 percent increase in the probability of attending secondary school as compared to if she had not attended the extra year. The significance of these results is robust to the specification of tighter intervals of birth cohorts in both countries (results available upon request from the author).

Then I explore the effect of schooling on three indicators of adolescent/early adult sexual behavior related to risk including: (1) age at sexual debut; (2) age at marriage; and (3) age difference between respondent and current spouse. In Malawi I find a negative effect (-.38) of schooling on age at marriage that is significant at the $p<0.01$ level (Table A2 column 4).

Interestingly, I find the opposite to be true in Uganda. Schooling has a positive effect (.29) on age at first marriage that is significant at the $p<.001$ level (Table A3 column 4). Likewise, in Uganda there is also a positive effect (.22) of schooling on age at first sex that is significant at the $p<.001$ level (Table A3 column 3). However, none of these results are robust to the specification of a tighter interval of birth cohorts, raising some concerns about the population I can make inference over (results available upon request from the author). Finally I explore the effect of schooling on recent (adult) sexual activity including: (1) whether the respondent used a condom during last sexual encounter; and (2) number of sexual partners (including spouse) in the last 12 months. Results are not significant in either country, suggesting a limited effect of schooling on later adult behavior (Tables A2 & A3).

Discussion

Building on the earlier work of Pettifor et al. (2005), de Walque (2007), Hargeeves et al. (2008) and others on the negative association between girls schooling and HIV status I use instrumented fuzzy RDD to causally demonstrate the negative effect of schooling on adult women's HIV status. I extend the findings of Baird and colleagues (2012) by demonstrating the effect of primary school, as opposed to secondary school, on adult HIV status. This is an important distinction because many students, particularly female students, do not reach secondary school due to financial or family constraints. In addition, primary school ends around ages 13 to 14 in both Uganda and Malawi, just at the point that many young adolescent females are engaging in sexual activity for the first time.

My investigation of the pathways through which schooling may affect the behavior associated with HIV transmission provides mixed results. On the one hand, increasing primary schooling for girls leads to higher probability of completion of primary school and attending secondary school in both countries. However, I find mixed effects of schooling on adolescent (self-reported) sexual behavior. In Malawi there is a negative effect of schooling on age at first marriage and in Uganda there is a positive effect of schooling on age at marriage and sexual debut. However, these results are not significant if I narrow the interval of birth cohorts of inference. In addition, I find no effect on adult (self-reported) sexual behavior. Interestingly, OLS and logistic regression estimates of the relationship between schooling and (self-reported) sexual behavior provide quite different results (Tables A4 & A5). In both countries, there is a highly significant positive association between schooling and age at sexual debut, age at marriage and likelihood of using a condom and a highly significant negative relationship between schooling and age difference between spouses. These differences in findings demonstrate how past explorations of the relationship between schooling, sexual behavior and HIV status may be biased due to unobserved heterogeneity.

The largest limitation in this analysis is that it is difficult to control for the possibility that older women may have higher rates of HIV infection simply because they have had longer time to be exposed to sexual activity. Furthermore, if the rate of infection changed dramatically over time then younger and older girls may have actually been exposed to different infection environments. Some of this is accounted for by narrowing the interval around the cutoff age to ensure that respondents are not so far apart in age and quite comparable on a variety of dimensions. The robustness of results across intervals lends some credibility to findings. Results from the logistic regression analysis of the association between schooling and HIV status provide further insight (Table A1). In Uganda there is a significant negative relationship (-.06) between schooling and HIV positive status and in Malawi there is a significant positive relationship (.04) between schooling and HIV positive status. Likely in these estimates increased schooling is proxying for time. Thus, results reflect the fact that the epidemic peaked earlier in Uganda then in Malawi, signaling that risk of infection for young people declined over time in Uganda and increased over time in Malawi (for a graphical representation of HIV positive status over time see Figures 3 & 4). If anything, the fact that I find robust negative results in both contexts using the RDD approach in spite of these differing time trends provides further support to the RDD findings.

This study has demonstrated that increased exposure to primary school has important long-term implications that extend beyond classroom learning. Importantly, I have demonstrated that endogeneity makes a difference: OLS estimates differ considerably from those provided by the fuzzy regression discontinuity approach. More causal work is needed to expand the evidence base on the effect of schooling on adult HIV status and other STIs. In addition, further research should better explore the complicated pathways through which schooling affects HIV status. Thus far most of the research has focused on the effect of schooling on sexual behavior. Considerably less attention has been paid to the role of schools as sites of cognitive and non-cognitive learning, perhaps due to concerns about poor quality of African schools. Future research should take up how improvements in school quality—such as student-to-teacher ratios, curriculum development and teacher evaluation—could contribute to both knowledge and behavior change in ways that improve the lives and livelihoods of young Africans

References

Baird, S., R. S. Garfein, C. McIntosh, & B. Ozler. (2012). Effect of a cash transfer programme for schooling on prevalence of HIV and herpes simplex type 2 in Malawi: a cluster randomised trial . *The Lancet,* 379, 1320–29

Deininger, K. (2003). Does cost of schooling affect enrollment by the poor? Universal primary education in Uganda. *Economics of Education Review*, 22, 291–305.

de Walque, D. (2007). How does the impact of an HIV/AIDS information campaign vary with educational attainment? Evidence from rural Uganda. *Journal of Development Economics,* 84, 686–714

Doyle A.M., Ross D.A., et al. (2010). Long-Term Biological and Behavioural Impact of an Adolescent Sexual Health Intervention in Tanzania: Follow-up Survey of the Community-Based MEMA kwa Vijana Trial. *PLoS Medicine*, 7:6, 1-14.

Duflo, E., Dupas P. & M. Kremer. (2012). Education, HIV, and Early Fertility: Experimental Evidence from Kenya. Unpublished results.

Dupas, P. (2011). Do Teenagers respond to HIV Risk Information? Evidence from a field experiment in Kenya. AEJ: *Applied Economics*, 3(1), 1-36.

Ferre, C. (2007). Age at first child: does education delay fertility timing? The case of Kenya. Unpublished results

Glynn, J.R., M. Carae, A. Buve, S. Anagonou, L. Zekeng, M. Kahindo, & R. Musonda. (2004). Does increased general schooling protect against HIV infection? A study in four African cities. *Tropical Medicine and International Health.* 9(1): 4-14.

Grogan, L. (2006). Who benefits from Universal Primary Education in Uganda? Unpublished results.

Hargreaves, J.R., Bonell, C.P., Boler, T., et al. (2008). Systematic review exploring time trends in the association between educational attainment and risk of HIV infection in sub-Saharan Africa. *AIDS,* 22**,** 403–414.

Kadzamira, E. & P. Rose. (2003). Can free primary education meet the needs of the poor? Evidence from Malawi. *International Journal of Educational Development,* 23, 501–516

Kattan, R. B. (2006). Implementation of Free Basic Education Policy. The World Bank Education Working Paper Series, 7. World Bank: Washington DC.

Lopman B, Lewis J, Nyamukapa C, Mushati P, Chandiwana S, Gregson S. (2007). HIV incidence and poverty in Manicaland, Zimbabwe: is HIV becoming a disease of the poor? *AIDS*, 2, S57–S66.

Mishra V, Assche SB, Greener R, Vaessen M, Hong R, Ghys PD, Boerma JT, Van Assche A, Khan S, Rutstein S. (2007). HIV infection does not disproportionately affect the poorer in sub-Saharan Africa. *Aids,* 21, S17–S28.

Ozier, O. (2011). The impact of secondary schooling in Kenya: a regression discontinuity analysis. Unpublished results

Padian N.S., Shiboski, S.C., & N.P. Jewell. (1991). Female-to-male transmission of human immunodeficiency virus. *Journal of the American Medical Association,* 266, 1664–1667.

Pettifor, A.E., Rees, H.V., Kleinschmidt, I., et al. (2005). Young people's sexual health in South Africa: HIV prevalence and sexual behaviors from a nationally representative household survey. *AIDS,* 19**,** 1525–34.

Pradhan, E. & D. Canning. (2013). The effect of education on teen fertility: evidence from the abolition of school fees in Ethiopia. Unpublished results.

Ross, D.A., Changalucha, J., et al. (2007). Biological and behavioural impact of an adolescent sexual health intervention in Tanzania: a community-randomized trial. *AIDS,* 21, 1943–1955.

Stock, J. H., J. H. Wright and M. Yogo. (2002). A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments. *Journal of Business and Economic Statistics*, 20(4), 518-29.

United Nations (UN). (2001). Road map towards the implementation of the United Nations Millennium Declaration: Report of the Secretary-General. New York: United Nations Department of Public Information.

United Nations Programme on HIV/AIDS (UNAIDS). (2008). Report on the global AIDS epidemic. Geneva: UNAIDS.

United Nations Educational Scientific and Cultural Organization (UNESCO). (2011a). Malawi country statistics. Accessed December 4th 2013. http://stats.uis.unesco.org/unesco/TableViewer/document.aspx?ReportId=121&IF_Language=en&BR_Country=4540

United Nations Educational Scientific and Cultural Organization (UNESCO). (2011b). Uganda country statistics. Accessed December 4th 2013. http://stats.uis.unesco.org/unesco/TableViewer/document.aspx?ReportId=121&IF_Language=en&BR_Country=8000

United Nations Programme on HIV/AIDS (UNAIDS). (2013). UNAIDS global report on the AIDS epidemic. New York: United Nations. Accessed December 12 2013. http://www.unaids.org/en/media/unaids/contentassets/documents/epidemiology/2013/gr2013/UNAIDS_Global_Report_2013_en.pdf

Welz T, Hosegood V, Jaffar S, Batzing-Feigenbaum J, Herbst K, Newell ML. (2007). Continued very high prevalence of HIV infection in rural KwaZulu-Natal, South Africa: a population-based longitudinal study. *AIDS*, 21(1), 467–1472.

World Bank. (2009). Abolishing School Fees in Africa: Lessons from Ethiopia, Ghana, Kenya, Malawi and Mozambique. Research Report. Washington D.C.: The World Bank.

APPENDIX

Table A1. Logistic regression estimates predicting HIV status in Malawi & Uganda

VARIABLES	(1) Malawi full sample	(2) Malawi 1974-1984	(3) Malawi 1976-1983	(4) Uganda full sample	(5) Uganda 1977-1987	(6) Uganda 1980-1986
Years schooling	0.000	0.019	0.039*	-0.039***	-0.046**	-0.058***
	(0.011)	(0.016)	(0.018)	(0.010)	(0.015)	(0.017)
Catholic	-0.099	-0.063	-0.118	-0.008	-0.152	-0.073
	(0.094)	(0.145)	(0.169)	(0.088)	(0.134)	(0.161)
Muslim	0.065	0.133	0.124	-0.208	-0.306	0.245
	(0.116)	(0.178)	(0.212)	(0.334)	(0.595)	(0.609)
Other religion	0.091	-0.272	0.043	-0.070	-0.186	-0.254
	(0.359)	(0.571)	(0.606)	(0.098)	(0.151)	(0.183)
Chewa ethnicity	-0.580***	-0.653***	-0.570***			
	(0.093)	(0.145)	(0.163)			
Baganda ethnicity				0.527***	0.423**	0.487*
				(0.097)	(0.164)	(0.198)
Constant	-1.830***	-1.540***	-1.710***	-2.330***	-1.968***	-1.956***
	(0.083)	(0.118)	(0.139)	(0.088)	(0.122)	(0.147)
Observations	7,396	2,512	1,824	11,967	3,829	2,692

Robust standard errors in parentheses (clustered at enumeration area level)
*** p<0.001, ** p<0.01, * p<0.05

APPENDIX

Table A2. Second stage estimates predicting additional pathways in Malawi (birth cohorts 1974-1984)

VARIABLES	(1) Complete primary	(2) Attend secondary	(3) Age first sex	(4) Age marriage	(5) Spouse age difference	(6) Condom used last sex	(7) Number partners last 12 months
Years school	0.074***	0.057***	0.063	-0.380**	-0.353	-0.001	0.002
	(0.013)	(0.010)	(0.102)	(0.144)	(0.200)	(0.010)	(0.018)
Catholic	-0.016	0.000	0.247	0.744**	-0.060	0.019	0.011
	(0.021)	(0.018)	(0.161)	(0.230)	(0.339)	(0.017)	(0.056)
Muslim	-0.010	-0.041+	-0.362	-1.084***	0.316	-0.029	-0.029
	(0.029)	(0.022)	(0.232)	(0.310)	(0.492)	(0.023)	(0.037)
Other religion	-0.164**	-0.108**	0.320	-0.894	-1.775+	-0.030	-0.032
	(0.050)	(0.039)	(0.750)	(0.909)	(0.907)	(0.049)	(0.073)
Chewa ethnicity	0.002	-0.004	0.695***	-0.033	-0.557*	-0.050***	0.025
	(0.017)	(0.015)	(0.154)	(0.206)	(0.277)	(0.014)	(0.029)
Constant	-0.112	-0.102+	16.136***	19.595***	7.699***	0.099+	0.889***
	(0.070)	(0.058)	(0.568)	(0.790)	(1.098)	(0.055)	(0.100)
Observations	2,512	2,512	2,291	2,468	2,036	2,237	2,508
R-squared	0.454	0.494	0.069			0.005	

Robust standard errors in parentheses (clustered at the enumeration area level)
*** p<0.001, ** p<0.01, * p<0.05

APPENDIX

Table A3. Second stage estimates predicting additional pathways in Uganda (birth cohorts 1977-1987)

VARIABLES	(1) Completion primary	(2) Attend secondary	(3) Age first sex	(4) Age first marriage	(5) Spouse age difference	(6) Condom used last sex	(7) Number partners last 12 months
Years school	0.100***	0.084***	0.219***	0.294**	-0.165	0.004	0.012
	(0.006)	(0.006)	(0.051)	(0.103)	(0.187)	(0.006)	(0.012)
Catholic	0.011	0.009	0.347**	0.176	-0.217	0.005	-0.027
	(0.012)	(0.011)	(0.117)	(0.164)	(0.295)	(0.012)	(0.020)
Muslim	-0.005	-0.023	0.169	0.465	-0.736	-0.001	-0.060
	(0.043)	(0.051)	(0.316)	(0.542)	(0.790)	(0.047)	(0.058)
Other religion	-0.003	0.004	-0.350**	-0.290+	0.271	-0.020+	0.003
	(0.013)	(0.012)	(0.115)	(0.175)	(0.321)	(0.011)	(0.026)
Baganda ethnicity	0.048+	0.040+	-0.473*	-0.050	0.761	0.070**	-0.070
	(0.028)	(0.023)	(0.194)	(0.341)	(0.631)	(0.025)	(0.044)
Constant	-0.203***	-0.226***	15.625***	16.079***	7.680***	0.045	0.894***
	(0.035)	(0.033)	(0.304)	(0.562)	(1.013)	(0.036)	(0.056)
Observations	3,829	3,829	3,755	3,553	2,984	3,442	3,826
R-squared	0.628	0.633	0.121	0.129	0.011	0.028	

Robust standard errors in parentheses (clustered at enumeration area level)
*** p<0.001, ** p<0.01, * p<0.05

APPENDIX

Table A4. OLS & logistic regression estimates predicting additional pathways in Malawi (birth cohorts 1978-1982)

VARIABLES	3 Age first sex	4 Age marriage	5 Spouse age difference	6 Condom used last sex	7 Number partners last 12 months
Years school	0.240***	0.330***	-0.201***	0.083**	-0.001
	(0.023)	(0.026)	(0.041)	(0.032)	(0.003)
Catholic	0.216	0.045	0.020	0.306	-0.051+
	(0.202)	(0.206)	(0.359)	(0.282)	(0.027)
Muslim	-0.378	-0.168	1.011+	-2.093*	-0.017
	(0.247)	(0.294)	(0.538)	(1.011)	(0.034)
Other religion	1.189	1.342	-1.306	0.358	0.006
	(1.043)	(1.220)	(1.361)	(1.038)	(0.076)
Chewa ethnicity	0.756***	0.393*	-0.418	-0.843**	0.020
	(0.180)	(0.178)	(0.326)	(0.323)	(0.021)
Constant	(0.167)	(0.190)	(0.334)	-2.884***	0.020
	15.184***	15.775***	6.698***	(0.293)	0.907***
Observations	1,074	1,145	948	1,037	1,166
R-squared	0.143	0.163	0.036		0.005

Robust standard errors in parentheses (clustered at the enumeration area level)
*** p<0.001, ** p<0.01, * p<0.05

APPENDIX

Table A5. OLS & logistic regression estimates predicting additional pathways in Uganda (birth cohorts 1982-1986)

VARIABLES	1 Age first sex	2 Age marriage	3 Spouse age difference	4 Condom used last sex	5 Number partners last 12 months
Years school	0.223***	0.359***	-0.133**	0.129***	-0.001
	(0.018)	(0.023)	(0.041)	(0.023)	(0.002)
Catholic	0.342*	0.356+	-0.257	0.031	-0.064+
	(0.142)	(0.197)	(0.358)	(0.197)	(0.036)
Muslim	0.588	0.455	0.317	0.289	-0.059
	(0.399)	(0.701)	(1.188)	(0.724)	(0.077)
Other religion	-0.335*	-0.282	0.239	-0.427+	-0.028
	(0.143)	(0.205)	(0.431)	(0.234)	(0.038)
Chewa ethnicity	-0.509**	-0.541*	0.738+	0.751***	-0.032
	(0.184)	(0.220)	(0.445)	(0.216)	(0.028)
Constant	(0.151)	(0.202)	(0.404)	-3.443***	-0.032
	15.663***	15.748***	7.459***	(0.237)	1.005***
Observations	1,991	1,874	1,609	1,844	2,028
R-squared	0.118	0.139	0.008		0.003

Robust standard errors in parentheses (clustered at the enumeration area level)
*** p<0.001, ** p<0.01, * p<0.05

FIGURE CAPTIONS

Figure 1. Average years of completed schooling by birth cohort in Malawi
Source: Created by the author

Figure 2. Average years of completed schooling by birth cohort in Uganda
Source: Created by the author

Figure 3. Overlaid lowess curve of birth cohort and HIV positive status in Malawi
Source: Created by the author

Figure 4. Overlaid lowess curve of birth cohort and HIV positive status in Uganda
Source: Created by the author

Table 1. Descriptive statistics (unweighted)

VARIABLES	Country	N	Mean	SD	Min	Max
Current age in 2010	Malawi	2512	30.1	3.23	25	36
Current age in 2011	Uganda	3829	27.89	3.06	23	34
Years schooling completed	Malawi	2512	5.24	3.78	0	17
Years schooling completed	Uganda	3829	5.75	4.18	0	19
Completed primary school	Malawi	2512	0.27	0.44	0	1
Completed primary school	Uganda	3829	0.38	0.48	0	1
Attended secondary school	Malawi	2512	0.19	0.39	0	1
Attended secondary school	Uganda	3829	0.26	0.44	0	1
Age at sexual debut [1]	Malawi	2303	16.65	2.66	8	31
Age at sexual debut [2]	Uganda	3770	16.83	2.7	4	30
Age at first marriage [3]	Malawi	2468	17.63	3.13	8	33
Age at first marriage [4]	Uganda	3553	17.66	3.91	4	33
Age difference between current spouse [5]	Malawi	2036	5.69	4.9	-12	40
Age difference between current spouses [6]	Uganda	2984	6.85	6.2	-21	47
Number of sexual partners last 12 months [7]	Malawi	2508	0.91	0.5	0	20
Number of sexual partners last 12 months [8]	Uganda	3826	0.94	0.49	0	20
Condom used last intercourse [9]	Malawi	2237	0.08	0.26	0	1
Condom used last intercourse [10]	Uganda	3442	0.08	0.27	0	1
HIV positive in 2010	Malawi	2512	0.16	0.37	0	1
HIV positive in 2011	Uganda	3829	0.09	0.29	0	1
Chewa ethnicity	Malawi	2512	0.29	0.45	0	1
Baganda ethnicity	Uganda	3829	0.16	0.37	0	1
Christian non-Catholic	Malawi	2512	0.69	0.46	0	1
Christian non-Catholic	Uganda	3829	0.32	0.47	0	1
Catholic	Malawi	2512	0.2	0.39	0	1
Catholic	Uganda	3829	0.41	0.49	0	1
Muslim	Malawi	2512	0.1	0.31	0	1
Muslim	Uganda	3829	0.01	0.12	0	1
Other religion	Malawi	2512	0.01	0.1	0	1
Other religion	Uganda	3829	0.26	0.43	0	1

(1) 209 respondents excluded because never had sex (n=6); or do not recall when had first sex (n=3); or provided inconsistent response (n=200)

(2) 59 respondents excluded because never had sex (n=42); or provided inconsistent response (n=17)

(3) 44 respondents excluded because not yet married

(4) 276 respondents excluded because not yet married

(5) 476 respondents excluded because provided inconsistent response (n=50); or not currently married (n=125); or missing information (n=301)

(6) 845 respondents excluded because provided inconsistent response (n=71); or not currently married (n=300); or missing information (n=474)

(7) 4 respondents excluded because refused to answer

(8) 3 respondents excluded because refused to answer

(9) 275 respondents excluded because missing information

(10) 387 respondents excluded because missing information

Table 2. First stage results predicting women's total years of schooling using 2010 Malawi Demographic and Health Survey

VARIABLES	(1) Full sample	(2) 1974-1984	(3) 1976-1983	(4) 1978-1982
Exposure age 13	2.127***	1.017***	0.721***	0.776***
	(0.051)	(0.083)	(0.099)	(0.221)
Catholic	0.547***	0.656***	0.708***	0.96***
	(0.083)	(0.132)	(0.144)	(0.285)
Muslim	-1.464***	-1.573***	-1.538***	-1.659***
	(0.137)	(0.157)	(0.178)	(0.362)
Other religion	-2.243***	-1.812***	-1.962***	-0.601
	(0.231)	(0.408)	(0.457)	(1.097)
Chewa ethnicity	-0.942***	-0.905***	-0.871***	-0.645**
	(0.095)	(0.128)	(0.137)	(0.244)
Constant	4.454***	5.150***	5.362***	5.259***
	(0.090)	(0.109)	(0.116)	(0.183)
Observations	23,018	7,990	5,891	1169
R-squared	0.119	0.053	0.043	0.04

Robust standard errors in parentheses (clustered at the enumeration area level)
*** p<0.001, ** p<0.01, * p<0.05

Table 3. First stage results predicting women's total years of schooling using 2011 Uganda AIDS Indicator Survey

	(1)	(2)	(3)	(4)
VARIABLES	Full sample	1977-1987	1980-1986	1982-1986
Exposure at 13	2.586***	1.533***	1.423***	0.960***
	(0.067)	(0.130)	(0.153)	(0.183)
Catholic	-0.816***	-0.717***	-0.676***	-0.652**
	(0.107)	(0.167)	(0.195)	(0.221)
Muslim	0.102	0.690	0.610	0.457
	(0.353)	(0.644)	(0.773)	(0.771)
Other religion	-0.122	0.118	0.042	-0.035
	(0.122)	(0.177)	(0.210)	(0.239)
Baganda ethnicity	2.889***	3.101***	3.258***	3.287***
	(0.149)	(0.223)	(0.257)	(0.278)
Constant	4.131***	4.857***	4.855***	5.325***
	(0.112)	(0.150)	(0.177)	(0.214)
Observations	11,994	3,840	2,699	2,037
R-squared	0.196	0.124	0.123	0.108

Robust standard errors in parentheses (clustered at the enumeration area level)
*** p<0.001, ** p<0.01, * p<0.05

Table 4. Second stage results predicting women's HIV positive status using 2010 Malawi Demographic and Health Survey

VARIABLES	(1) Full sample	(2) 1974-1984	(3) 1976-1983	(4) 1978-1982
Years schooling	-0.058***	-0.064***	-0.064**	-0.069*
	(0.005)	(0.016)	(0.023)	(0.033)
Catholic	0.030*	0.058*	0.055	0.048
	(0.012)	(0.029)	(0.035)	(0.045)
Muslim	-0.076***	-0.078*	-0.097+	-0.120+
	(0.020)	(0.040)	(0.050)	(0.066)
Other religion	-0.149***	-0.128	-0.069	-0.038
	(0.045)	(0.086)	(0.113)	(0.150)
Chewa	-0.103***	-0.134***	-0.124***	-0.119***
	(0.012)	(0.025)	(0.030)	(0.036)
Constant	0.465***	0.541***	0.543***	0.572**
	(0.031)	(0.093)	(0.130)	(0.190)
Observations	7,396	2,512	1,824	1,169

Robust standard errors in parentheses (clustered at the enumeration area level)
*** p<0.001, ** p<0.01, * p<0.05

Table 5. Second stage results predicting women's HIV positive status using 2011 Uganda AIDS Indicator Survey

VARIABLES	(1) Full sample	(2) 1977-1987	(3) 1980-1986	(4) 1982-1986
Years schooling	-0.020***	-0.021***	-0.024**	-0.039**
	-0.002	-0.006	-0.008	(0.015)
Catholic	-0.016*	-0.026+	-0.02	-0.028
	-0.007	-0.014	-0.017	(0.022)
Muslim	-0.008	-0.012	0.039	0.030
	-0.022	-0.046	-0.067	(0.087)
Other	-0.006	-0.014	-0.02	-0.024
	-0.007	-0.014	-0.016	(0.020)
Baganda	0.093***	0.092***	0.108**	0.146**
	-0.01	-0.024	-0.033	(0.054)
Constant	0.180***	0.214***	0.229***	0.325***
	-0.013	-0.035	-0.046	(0.089)
Observations	11967	3829	2692	2,031

Robust standard errors in parentheses (clustered at the enumeration area level)
*** p<0.001, ** p<0.01, * p<0.05

Figure 1.

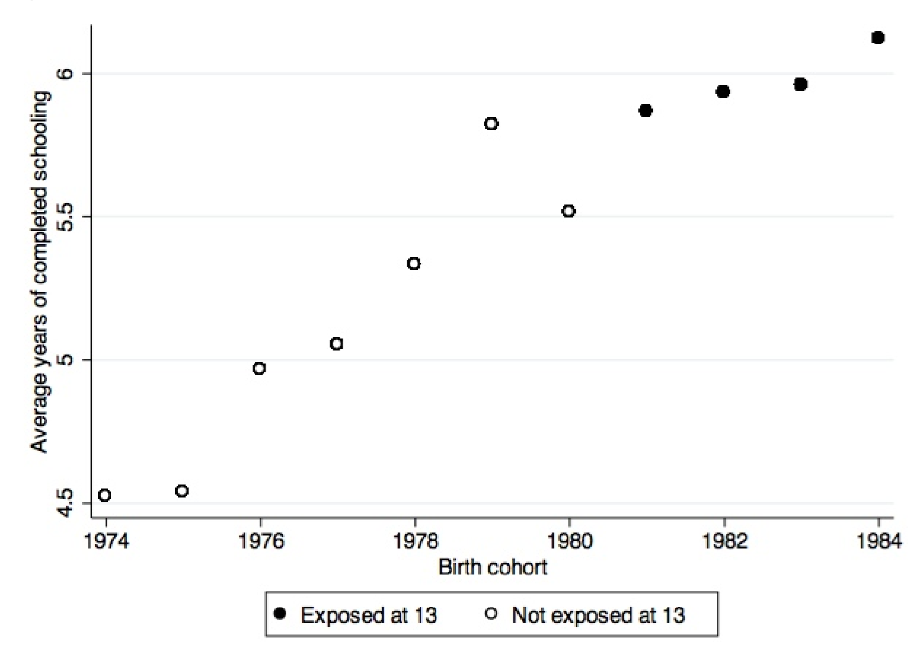

Figure 2.

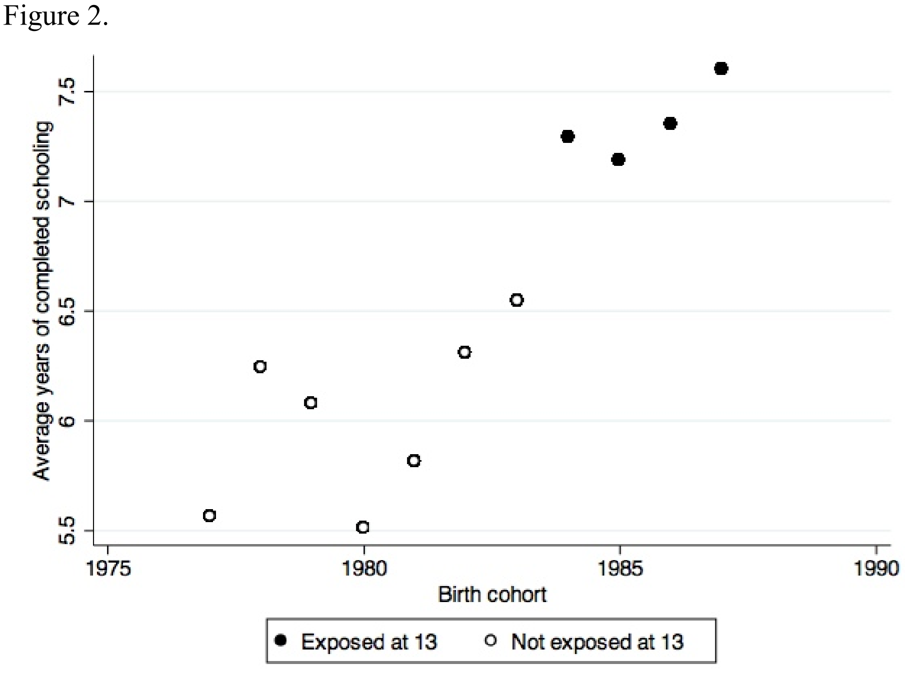

Figure 3.

Figure 4.