Does schooling protect sexual health? The influence of school attendance on herpes simplex type 2 infection among adolescents in Malawi

Barbara S. Mensch¹, Monica J. Grant², Christine A. Kelly², Erica Soler-Hampejsek¹, Satvika Chalasani⁴, Christopher Sudfeld⁵ and Paul C. Hewett⁶

¹Population Council, New York, ²University of Wisconsin, ³London School of Hygiene and Tropical Medicine, ⁴UNFPA, ⁵Harvard University, ⁶Population Council, Lusaka
Introduction
Over the last several decades educational participation among adolescents, particularly adolescent girls, has increased considerably in sub-Saharan Africa1 (AFDB 2011,2014; UNESCO 2011) with the consequence that young people are much more likely to reach sexual maturity while enrolled in school (Biddlecom et al., 2008). Moreover, while the age of first sex has either remained the same or risen, at least for young women, in many countries a shift has taken place in the context of sexual initiation from within to before marriage (Mensch, Grant and Blanc 2006). An analysis of DHS data from 27 sub-Saharan African countries revealed that among those who were unmarried, adolescent girls enrolled in school were less likely to initiate sex than their out of school counterparts. Moreover, among girls who were sexually active, those enrolled in school were more likely to report using contraception (Lloyd, 2009). While comparable data for young men are not as widely available, enrollment in school appears to be less strongly associated with sexual initiation among boys (Blanc et al 2005; Clark and Mathur 2012).2

This inverse association between school attendance and premarital sex has led some to assert that sexual activity and schooling appear to be incompatible for young women (Clark and Mathur et al 2012) and that schooling offers some measure of “protection” for unmarried adolescent girls in Africa who often have limited access to condoms and other methods of contraception (Lloyd 2007).

This paper investigates the association between school attendance and HSV-2 infection in the southern region of Malawi with data from a longitudinal survey of adolescents first interviewed in 2007. Beginning in the fourth round of data collection in 2010, respondents were tested for HIV and herpes simplex virus type 2 (HSV-2), commonly referred to as genital herpes. Because HIV prevalence among the respondents, who were aged 14–16 at the beginning of 2007, is too low to provide sufficient power for estimating associations reliably, our analysis is focused on HSV-2, which is almost always sexually transmitted and is thus considered a reliable marker of sexual behavior among adolescents (Obasi et al. 1999; Smith and Robinson 2002; Wagner et al. 1994). Genital herpes, which often produces lesions particularly during initial infection, is strongly associated with HIV in sub-Saharan Africa (Obasi et al. 1999). While HIV acquisition is known to be more likely in the presence of genital tract inflammation, the risk of acquiring HIV may be elevated even among those who are asymptomatic due to viral shedding (Celum et al. 2010; Glynn et al. 2008; O'Farrell 1999; Todd et al. 2006). HSV-2 co-infection also increases the risk of HIV transmission to uninfected partners (Reynolds 2009, and is a cause of neonatal morbidity and mortality due to exposure in the genital tract during childbirth (Sudfeld et al. 2013). While many of those seropositive for HSV-2 never have a clinical outbreak, prevention of infection has important consequences for public health particularly in low income countries where suppressive therapy is not readily available3.

Schooling and Sexual Behavior in sub-Saharan Africa
The observation that school enrollment delays sexual debut for adolescent females invites the

---

1 This assertion is based on attainment data and secondary school gross enrollment ratios as trend data on current participation by age are not easily accessible.
2 There are exceptions; one study using longitudinal data from rural South Africa observed no gender difference in the significant association between school attendance and the timing of sexual debut (McGrath et al. 2009).
3 While a clinical trial failed to demonstrate that suppressive therapy with acyclovir among those co-infected with HIV/HSV-2 reduces transmission of HIV to uninfected partners (Celum et al. 2008), it has been shown to have an effect on disease progression in those co-infected (Buvé 2010).
question of what might make school protective. Scholars have speculated as to what might explain the association, but few empirical studies have investigated the pathways. Some research has suggested that students may have smaller, less risky sexual networks than youth who are out of school, or that behavioral norms within students’ social networks may reinforce safer behaviors (Hargreaves et al. 2008; Jukes et al 2008). Others have suggested that being a “school girl” and wearing a school uniform signals or confers a special status even for girls who are post-puberty such that students are not considered appropriate sexual or marital partners (Bledsoe 1990). In contrast, microeconomic theory posits that individuals who invest in their education have more incentive to protect their health because of greater expected returns in the future (Becker 1993; de Walque 2007). Students who are attached to school and have higher aspirations for the future may be less inclined to engage in sexual activity (Kirby 2002; Frye 2012) in order to avoid either a poorly timed pregnancy or infection with HIV. School participation is also thought to challenge students’ traditional notions of status hierarchies, with education and occupation replacing age and gender as factors affecting societal ranking (LeVine et al. 2001). Where such a shift in world view takes place, it may affect the willingness of students to engage in sexual activity, particularly female students for whom the short-term consequences of unprotected sex are greater. The messages conveyed by teachers regarding the importance of staying in school and passing school-leaving exams may also affect student attitudes regarding the desirability of early sex and marriage. Students are also more likely to be exposed to—and to understand—safe sex messages through the formal school curriculum.

While arguments exist for why school attendance may be protective, there is also evidence that school can be a risky environment for girls. Numerous reports have been published about the widespread school related gender-based violence in developing countries (Wellesley Centers for Research on Women 2003; MSI 2008). Sexual harassment is often said to take place on the way to and from school as well as in school. Abuse of girls has been reported to be perpetrated by male classmates as well by male teachers who may exchange sex for better grades. This school related gender based violence is said to have serious consequences for emotional and physical health including increasing risk for sexually transmitted infections (MSI 2008).

In addition there is evidence, often anecdotal, that girls exchange sexual favors for school fees or other financial support in order to remain in school (Luke 2003). While primary education is free in most sub-Saharan African countries, tuition is generally charged for secondary school (UNESC0 2014). Even if tuition is free, students often need to purchase uniforms and other supplies.

Whether school is protective or potentially dangerous may also vary by student performance. A study in Cape Town found that while girls who attended school were less likely to be sexually active than their out-of-school counterparts, students who were attending the appropriate grade for age were more likely to have initiated sex than their female peers who had repeated grades. The explanation given for the increased likelihood of sexual debut for the girls who were in the appropriate grade for age is that they were exposed to male students 2-3 years older who may be more interested in engaging in sexual activity (Lam, Marteleto and Ranchhod 2013). Schooling discontinuities, including grade repetition and temporary withdrawals from school, have also been linked to a higher risk of schoolgirl pregnancy (Grant and Hallman 2008).

While explanations for the protective effect of schooling are theoretically plausible, an inverse association between school enrollment and STIs may be observed not because school attendance,

---

4 Note that out-of-school girls may also face unsafe environments and may be pressured to have sex; the question is whether school attendance puts girls at increased risk.
in and of itself, reduces vulnerability to infection, but because of selection effects. Those who are enrolled in school may be different from their less educated peers in ways that lead to lower risk. For example, they may be more ambitious and less willing to engage in risky behavior not because of messages conveyed in school but because of different underlying characteristics.

Another factor confounding the reported association between school enrollment and sexual behavior is the reliability of self-reported data, particularly if it varies by student status (Gregson et al. 2009; Langhaug et al. 2011; Lindstrom et al. 2010; Luke et al. 2011; Mensch et al. 2003; Mensch et al. 2008; Minnis et al. 2009; Nkpo et al. 2004). A recent analysis of MSAS data found that school-going females were more likely than females who were no longer enrolled to provide inconsistent reports about sexual behavior (Soler-Hampejsek et al. 2013) reflecting an apparent ambivalence about reporting sex; female students may be concerned that if family or school authorities were to discover that they were sexually active, their commitment to school may be questioned (Grant 2012).

**The Association between School Status and STIs**

While a considerable body of research exists on the association between educational attainment and HIV among adults, few studies of adolescents have investigated the association of school attendance and HIV or HSV-2. One exception is a study conducted in 2001 by Hargreaves et al. (2008) among a random population-based cross-sectional sample of rural South Africans aged 14-25, which found that school attendance was associated with lower HIV prevalence among young men but not young women. A recent randomized control trial in Malawi examined the effect of a cash transfer program for schooling on HIV and HSV-2 prevalence among young unmarried women. The cash transfer group was divided between a conditional group that received cash payment if school was attended for 80% of the days that it was in session during the prior month, and an unconditional group for whom payment was provided regardless of attendance. Significant differences were found between the cash transfer groups and a control group that did not receive any payment but not between the conditional and unconditional intervention groups. In observing lower infection rates in the cash transfer group, the researchers conclude that “financially empowering school-age girls and their families can have substantial effects on their sexual and reproductive health” (Baird et al. 2012:1328). The findings are consistent with the assertion that poverty makes girls more vulnerable to risk behavior. The study investigators do not comment on the absence of a difference in prevalence between the conditional and unconditional intervention groups. However, the results do not provide evidence for a protective effect of school participation on infection, although an earlier analysis suggested that attendance may delay the onset and reduce the frequency of sexual activity (Baird et al. 2010)

To the best of our knowledge, no panel studies of adolescents have examined the timing of HIV/HSV-2 acquisition relative to school attendance, although one study (de Walque 2007) has investigated the association between HIV and educational attainment.

**Study Aims**

Using data from a longitudinal survey of adolescents in rural Malawi first interviewed in 2007, and HSV-2 results from the fourth, fifth and sixth rounds of fieldwork in 2010, 2011 and 2013, the goal of this paper is to estimate the association between school attendance and HSV-2, and to examine whether the effects vary for males and females, reflecting different underlying mechanisms. We hypothesize that current school attendance is inversely related to HSV-2 among adolescent girls who may be disinclined to engage in sex while in school for fear of becoming
pregnant and having to drop out, but unrelated to HSV-2 among adolescent boys for whom the consequences of unprotected sex are considerably less severe in terms of school attendance. We also hypothesize that students with better school performance, measured as better reading and numeracy skills and higher grade attainment, will be less likely to become infected with HSV-2.

In developing these analyses we will investigate whether those who were not interviewed and/or were unwilling to be tested in Rounds 5 and 6 are selective, since sample attrition and testing refusal potentially bias estimates of disease incidence and prevalence (Obare 2010; Reniers and Eaton 2009; Clark and Houle 2012).

**Study Context**

*HIV and HSV-2 Prevalence in Malawi*

Malawi is among the 10 poorest countries in the world with a gross national income per capita of $860 in 2010 (http://www.prb.org/pdf12/2012-population-data-sheet_eng.pdf). Eighty-five percent of the population live in rural areas without access to proper sanitation, potable water, electricity, and all-weather roads. Age at marriage among women is early, with nearly half of 20-24 year olds in 2010 marrying by age 18 (National Statistical Office (NSO) and ICF Macro 2011). Moreover marriage in Malawi “fragile,” with high rates of separation and divorce even at young ages (Reniers, 2008), which may be both a cause and consequence of HIV.

According to the 2010 DHS, HIV prevalence in Malawi is currently estimated to be 10.6% with substantial variability across the country; in the South, the region where our longitudinal survey is situated, prevalence is estimated to be 14.5% (National Statistical Office (NSO) and ICF Macro 2011). As is the case elsewhere in sub-Saharan Africa, young Malawian women are disproportionately affected by the epidemic (Kumwenda et al. 2008). Moreover, the gender difference in prevalence is much greater among young people, with women under age 30 considerably more likely to be infected than their male peers. Indeed, among those aged 15-21 residing in rural areas in the Southern region, prevalence was reported to be about 11% for women and two percent for men.5 Notably, 2010 DHS data indicate that years of schooling and HIV prevalence have a u-shape association with the highest levels among those with no schooling and those with some secondary education (National Statistical Office (NSO) and ICF Macro 2011). On the other hand, incidence data from urban Malawi indicate a positive association between educational attainment and HIV among women (Kumwenda et al. 2006). DHS data indicate that prevalence is associated with wealth quintile for both men and women; those in a position of relative affluence are more likely to be infected (National Statistical Office (NSO) and ICF Macro 2011). Considerable variability in prevalence is also observed among ethnic groups. The Yao live primarily in the South, are predominantly Muslim and practice male circumcision, which is demonstrated to have a protective effect (Bailey et al. 2007; Gray et al. 2007), yet have one of the highest rates of HIV. Researchers have speculated that the Yao are more likely to engage in sexual practices that elevate risk (Poulin and Muula 2007).

Less is known about the prevalence of HSV-2 in Malawi than about the prevalence of HIV because testing is not common. However, some general patterns regarding the epidemiology of HSV-2 in sub-Saharan Africa are worth noting. Like HIV, HSV-2 is more common among young women than among young men both because women’s partners are typically older and because adolescent females are more susceptible to infection due to cervical ectopy (Smith and Robinson 2002). Prevalence increases steeply with age and then typically plateaus later in

---

5 Investigators’ tabulations from 2004 Malawi DHS.
An HSV-2 study in rural northern Malawi, based on stored sera collected between 1988 and 2005 from a random sample of the general population who were case-controls for mycobacterial disease studies and from women attending antenatal clinics, found high prevalence (Glynn et al. 2008). Age-standardized rates for the period from 1988-90 were 30% for men and 48% for women aged 15-44, and increased to 36% for men and 58% for women in the period between 1998 and 2001, and were then statistically similar for the following four year period. The sample of 15-19 year olds in the most recent period was too small to allow confident assertions about prevalence for adolescents. However, the prevalence estimate for those under age 25 was between 22% and 24% for the entire study period, which the authors did not separate by sex because of small sample size.

The 2008-9 randomized control trial assessing the effect of a conditional cash transfer program on the risk of HIV and HSV-2 infections among never-married young women aged 13-22 in Zomba district in the Southern region of the country, discussed earlier, observed relatively low prevalence of HIV and HSV-2 18 months after baseline. For the segment of the sample attending school at baseline, HSV-2 prevalence was 3% among the control group, 1% among the conditional cash transfer group and less than 1% among the unconditional cash transfer group. For the segment of the sample no longer attending school at baseline, HSV-2 prevalence was 8% among both the intervention and control groups. That the trial was limited to the unmarried is undoubtedly a factor in the low prevalence since all the married are likely sexually experienced. Malawi is characterized by early marriage, with median age at marriage among women estimated to be 17.6 in the Southern region (National Statistical Office (NSO) and ICF Macro 2011), suggesting that the trial participants, while drawn from a representative sample of households, were likely to be selective of the population in this age group.

**Education in Malawi**

While Malawi still ranks on the low end among African countries in grade attainment, there has been substantial progress in the last couple of decades. Since the elimination of all primary school fees in 1994, Malawi has achieved nearly universal access to primary school. In 2010, 4% of males and 8% of females aged 20-24 had never attended school compared with 11% of males and 34% of females aged 40-44. Moreover, 51% of young men and 38% of young women aged 20–24 had completed primary school (which extends through eight grades known as standards), compared with 29% of males aged 40-44 and 16% of females.

Although access to school has increased, it is thought to have had little positive impact on other critical schooling outcomes, namely retention, grade repetition, attainment, and skill acquisition. In fact, according to the results of numeracy tests administered in Standard 6 in Malawi, no students scored in the “competent” range or above (UNESCO 2005). Indeed, Malawi had the lowest numeracy scores of 14 countries in the Southern and Eastern African Consortium for Monitoring Educational Quality (SACMEQ). Furthermore, although policies and interventions to encourage girls’ school enrollment have contributed to near gender parity in primary school entry (Anzar et al. 2004; Chimombo et al. 2000), girls are still more likely to leave school than are boys and do so at a younger age (Baird et al. 2012). Moreover, due to late entry, repetition, and temporary withdrawal, adolescents of the same age are distributed across a wide range of

---

6 The differences between cohorts are likely larger due to selective mortality; those aged 40-44 who did not attend school or completed few years are less likely to have survived until 2010.

7 For a discussion of the effect of free primary education on school quality and community-state relations, see Kendall (2007).
standards.

**Data**
The Malawi Schooling and Adolescent Study (MSAS) is a six-round longitudinal study of 2,649 adolescents resident in two contiguous rural districts in the southern region of the country, and reported to be aged 14–16 in January 2007. The initial 2007 sample consisted of 1,764 students (875 girls and 889 boys) who were randomly selected from the enrollment rosters at 59 randomly selected primary schools in Machinga and Balaka districts. The probability of a particular school being included was proportional to its enrollment in 2006. At each school, approximately 30 students stratified by gender and age who were enrolled in standards 4–8, the last four years of primary school, were interviewed. An additional sample of 885 adolescents (462 girls and 423 boys) who were not enrolled in school was drawn from the communities surrounding the selected primary schools. These respondents, referred to as the “out-of-school sample” because of their status when first interviewed, were identified through key informants located at the school or in the randomly selected school catchment villages. The study’s ratio of 14–16 year olds attending standards 4–8 relative to those out of school was dictated by the proportion observed in the 2004 DHS for Malawi. Survey weights were constructed to correct for the differences in the probability of being included in the sample by school, gender, and enrollment status at Round 1. Follow-up interviews were conducted between 2007 and 2013. The study successfully re-interviewed 91%, 90%, 88%, 88% and 83% of the original sample in 2008, 2009, 2010, 2011, and 2013 respectively.

The MSAS adolescent instrument includes an extensive set of questions on household and family characteristics, educational attainment, schooling history and experiences, household labor and employment, health, marriage, and sexual behavior. Numeracy and literacy assessments were also conducted in each round.

**STI Testing Procedures**
Beginning in 2010 (Round 4), respondents were tested for HIV and HSV-2. HIV testing was conducted by enumerators trained in counseling and testing, following the Ministry of Health’s guidelines. Testing was conducted at the household for the majority of respondents, while others were tested at private external locations. After completing the main survey, the interviewer obtained consent for testing. For those respondents giving consent, HIV status was determined via whole blood obtained from finger pricks using EDTA capillary tubes. A serial algorithm of HIV testing was used: if respondents tested positive for Determine, they were retested using Uni-Gold™ Recombigen®. Both tests have a very high sensitivity (100%) and specificity (>99%) in clinic evaluations, including a controlled laboratory setting in rural Kenya (Foglia et al. 2004). In Rounds 4 and 5 BIOLINE HIV 1/2 3.0, was used as a tiebreaker in cases in which Determine and Unigold tests gave contradictory results.

The HSV-2 specimen was also collected at the household or private location via finger prick.

---

8 The 2004 Malawi DHS indicated that 77 percent of 14-16 year olds attending school were in standards 4-8; 6 percent were in standards 1-3 and 17 percent were in secondary school. Approximately 25 percent of 14-16 year olds were not in school. Thus the ratio of 14-16 year olds in standards 4-8 relative to those out of school based on the DHS was 70:30, whereas the MSAS ratio is 67:33.

9 Due to a reduction in funding, data were not collected in 2012.

10 BIOLINE was removed from the HIV testing algorithm in Round 6 (2013), in accordance with revised Ministry of Health guidelines. In cases where Determine and Uni-Gold produced discordant results, both tests were repeated in tandem.
A sample of whole blood was collected and stored in microtainers, which were transported to the College of Medicine-Johns Hopkins Research Project Laboratory at Queen Elizabeth Central Hospital in Blantyre for testing. The Kalon ELISA HSV-2 antibody test was used. In an evaluation of several assays using African sera, Kalon was found to have a sensitivity of 92.3% and specificity of 97.7% (van Dyck et al. 2004) although other validation studies have reported lower values for both measures (Delany-Moretwe 2010). In most evaluations, the sensitivity with Kalon is lower than with HerpeSelect, although the specificity is higher (Delany-Moretwe 2010). An external validation of the testing procedure for this study was successfully carried out in 2010 and 2013 by Contract Laboratory Services in Johannesburg.

With few exceptions, respondents who tested positive in Round 4 were not retested in Rounds 5 and those positive in Round 4 or 5 not retested in Round 6. In view of the small number of samples for which Kalon did not produce a definitive result — 1.6% of those tested in Round 4 — indeterminate results were not retested with HerpeSelect, as had originally been proposed. In Rounds 5 and 6, however, indeterminate specimens were retested twice with Kalon in an effort to resolve the indeterminacy. After specimen collection, all participants were provided information about HSV-2 detection, symptoms, safe sex practices, and treatment options. HSV-2 test results were made available to study participants at centralized health centers proximate to Balaka and Machinga districts. Respondents were provided vouchers with identification numbers to receive their test results and reimbursement for travel to the site. In Round 4, 23.2% of the 1855 participants tested received their result at a clinic. In Round 5, 26.3% of the 1763 respondents tested returned to the clinic to obtain their results. In Round 6, 28.4% of the 1575 participants tested received their result at a clinic and, after an adjustment of the dissemination protocol, an additional 5.0% received their results by phone.

Analysis

*Explanatory variables*

The focus of the analysis is on the respondent’s school status, based on enrollment in the school terms since the prior round, and in the interval preceding that prior round. Respondents who were not enrolled in school for any of the school terms were coded as being out of school. Respondents who were enrolled in all school terms were coded as being in school. Individuals who were enrolled in at least one, but not all, terms were labelled as having “left school between the prior round and the current round.” Note that the majority of these respondents actually dropped out of school during the inter-survey period. Three additional schooling variables are also included: highest grade completed, English literacy, and numeracy skills. English literacy was determined by the ability to read two sentences in English without difficulty. Numeracy skills were measured by the total number of correct answers to 12 math problems.

In addition to these schooling variables, the models included a set of individual characteristics, family background measures, and community HSV-2 prevalence. Individual characteristics

---

11 In Round 4 we had anticipated using dried blood spots to test for HSV-2. However, a validation exercise we conducted for a household-based survey in Uganda a month prior to the Malawi fieldwork indicated that dried blood spots were not optimal for HSV-2 testing and that serum collection was preferable. Additional IRB clearance was necessary for this change to testing procedures, which was not obtained until after the start of fieldwork. Although we attempted to retrace respondents who had completed the survey but had not been offered HSV-2 testing before the IRB approvals, 168 respondents could not be relocated during the time available and therefore did not complete HSV-2 testing in Round 4.

12 Note that because of the change in the testing protocol described above, we were unable to give respondents a fixed date for the start of results dissemination. In the desire to reach as many respondents as possible, including those who might not have otherwise travelled to the clinics, we provided results over the phone for some respondents. An additional 22.4 percent of Round 4 respondents were reported to have obtained their results via phone.
included the respondent’s age, ethnic group, marital status, paid or unpaid work, whether the respondent had ever moved, and residence in an urban area. Family background measures included parents’ education, parents’ survival and co-residence with the respondent, and an index of household asset ownership. Community HSV-2 prevalence was measured as the proportion of respondents from each original school sampling area who were HSV-2 positive. All explanatory variables other than age and ethnic group are time varying.

Analytic Methods
Our key interest is in the association between school enrollment and HSV-2 seroconversion from Round 4 to Round 6. Our initial measurement of HSV-2 infection, however, is left censored; we only know the current infection status at Round 4, but not the time interval when the infection first occurred. Therefore, we divide our analysis into two stages. The first stage of the analysis examines HSV-2 prevalence in Round 4, the first survey round in which blood samples were collected and tested. The second stage of the analysis will use discrete time event history models to examine transitions to HSV-2 infection over the subsequent two survey rounds, controlling for left and right censoring.

The first stage of the analysis initially uses a probit regression to examine the covariates associated with being HSV-2 positive for those respondents who were tested. The second model corrects these estimates for testing refusal by using a biprobit regression with selection. We use a set of four variables to predict testing refusal that are not associated with HSV-2 infection status: exam status, long distance migration, interviewed during Ramadan, and current marital status. Qualitative questions about testing refusal indicated that students who were preparing for the national exams at the end of primary school, the second year of secondary school, and the last year of secondary school were reluctant to be tested for fear that concerns about the test result would distract them from their studies. These responses also indicated that many Muslim respondents who were interviewed during Ramadan (indicated by the date of the interview) believed that drawing blood would force them to break the fast. Travel logistics meant that it was not possible to collect blood samples from respondents who had moved a great distance outside of the original sample area. Finally, currently married respondents were often discouraged by their spouse from consenting to be tested. The testing selection equation and the test result equations are estimated simultaneously. All variables in the prevalence regression were measured at the same survey round at which the blood sample was collected. Results from these regressions are presented in Table 2 and are discussed briefly in the next section.

The second stage of the analysis will examine the association between school enrollment and HSV-2 seroconversion from Round 4 to Round 6. One approach is to model HSV-2 seroconversion, conditional on being HSV-2 negative at Round 4. This method, however, will yield biased results as a non-trivial proportion of respondents were already HSV-2 positive at the time of the initial test. Therefore, we will explore methods that will allow us to control for both left and right censoring of discrete time data. For example, Cain and colleagues (2011) recommend using interval regressions, such as the “intreg” command in Stata, to accommodate left and right censored epidemiological processes, such as disease seroconversion. In our analysis, the age at which a respondent tested positive for HSV-2 would be modelled as the dependent variable, with left and right censoring indicated in the interval variables.

We will also use these analyses to test the sensitivity of the association between school enrollment and HSV-2 seroconversion to the timing of variable measurement. These analyses will compare the results that are estimated when using school enrollment status over the inter-
survey period immediately preceding HSV-2 testing to results that are obtained when school status variables are lagged one survey round prior to HSV-2 testing.

**Preliminary Results**

Table 1 presents the distribution of the HSV-2 and school enrollment variables across survey rounds and by sex. At all survey rounds, HSV-2 prevalence was higher for females than for males. In Round 4, the first time respondents were tested for HSV-2, almost 10 percent of males and 15 percent of females tested positive. Prevalence more than doubled two years later; at the sixth survey round 21 percent of males and 32 percent of females were positive.

By the fourth survey round, when respondents were 17-20 years old, 63 percent of females had already left school as compared to 42 percent of males. An additional 14 percent of males and 12 percent of females left school between the third and fourth survey rounds. By the sixth survey round, when the respondents were 20-23 years old, only 7 percent of females and 21 percent of males remained enrolled in school for the entire period from Round 5 to Round 6.

For males, there was no clear pattern of HSV-2 infection by recent school enrollment status. Females who were already out of school, however, had a much higher HSV-2 prevalence. At the time of the fourth survey round, almost 19 percent of out-of-school girls were HSV-2 positive, as compared to 7 percent of in-school girls and 9 percent of girls who left school during the previous inter-survey period. By the sixth survey round, almost 36 percent of out-of-school girls were infected, as compared to only 10 percent of in-school girls and 21 percent of girls who left school during the previous inter-survey period.

Table 2 presents the probit and Heckman selection models for HSV-2 prevalence in Round 4, by sex. Controlling for all other covariates, we find a positive and statistically significant association between school enrollment since the last survey round and HSV-2 infection for males, but a negative and statistically significant association between recent school enrollment and HSV-2 infection for females. Girls who left school during the previous inter-survey period were also significantly less likely to be infected with HSV-2 relative to girls who had already left school by the previous survey round. In the Heckman selection model, both males and females who were interviewed at a distance from Balaka and Machinga districts—the districts where our sample resided in 2007—were significantly less likely to have been tested for HSV-2 because of the difficulty of preserving the samples for testing. Males who were enrolled in school during an exam year were significantly more likely to have been tested for HSV-2, contrary to our expectations. Despite the significant association between these variables and HSV-2 testing, the inclusion of the selection equation did not substantively modify the pattern of association between our key variables and infection status. The post-estimation test that rho=0 was non-significant indicating that the error terms of the prevalence and selection equations were uncorrelated; in this situation, the regression which estimates prevalence is the same as a simple probit model and suggests that HSV-2 testing refusal may not be a source of significant bias in these data.

Future analyses, focusing on seroconversion, will investigate whether the negative association between current school status and HSV-2 prevalence for girls observed in Table 2 reflects a protective effect of schooling or the reverse association, to wit, that sex among schoolgirls, which we measure by HSV-2 infection, increases the likelihood of school dropout, perhaps because of pregnancy and/or impending marriage.
<table>
<thead>
<tr>
<th></th>
<th>Male</th>
<th></th>
<th>Female</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Round 4</td>
<td>Round 5</td>
<td>Round 6</td>
<td>Round 4</td>
</tr>
<tr>
<td>HSV-2 prevalence</td>
<td>9.9</td>
<td>16.7</td>
<td>21.4</td>
<td>15</td>
</tr>
<tr>
<td>School enrollment</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Out of school by prior round</td>
<td>41.9</td>
<td>56.5</td>
<td>63.7</td>
<td>65.4</td>
</tr>
<tr>
<td>In school, current round</td>
<td>44.4</td>
<td>39.8</td>
<td>20.6</td>
<td>22.3</td>
</tr>
<tr>
<td>Left school between prior round and current round</td>
<td>13.7</td>
<td>3.4</td>
<td>15.7</td>
<td>12.3</td>
</tr>
<tr>
<td>HSV-2 prevalence by school enrollment</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Out of school by prior round</td>
<td>9.5</td>
<td>17.6</td>
<td>19.7</td>
<td>18.9</td>
</tr>
<tr>
<td>In school, current round</td>
<td>10.4</td>
<td>15.4</td>
<td>24</td>
<td>7.2</td>
</tr>
<tr>
<td>Left school between prior round and current round</td>
<td>9.2</td>
<td>17.9</td>
<td>24.3</td>
<td>9.4</td>
</tr>
<tr>
<td>N</td>
<td>892</td>
<td>941</td>
<td>852</td>
<td>952</td>
</tr>
</tbody>
</table>
Table 2. Probit and Heckman selection models, HSV-2 Prevalence, Malawi Schooling and Adolescent Study, 2010 (Round 4)

<table>
<thead>
<tr>
<th></th>
<th>Male Probit</th>
<th>Male Heckman probit with selection</th>
<th>Female Probit</th>
<th>Female Heckman probit with selection</th>
</tr>
</thead>
<tbody>
<tr>
<td>Current school enrollment (ref. Out of school by prior round)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In school, current round</td>
<td>0.52 *</td>
<td>0.49 *</td>
<td>-0.73 **</td>
<td>-0.72 **</td>
</tr>
<tr>
<td>Left school between prior round and current round</td>
<td>0.30</td>
<td>0.29</td>
<td>-0.75 ***</td>
<td>-0.74 ***</td>
</tr>
<tr>
<td>Highest grade attended</td>
<td>-0.09 +</td>
<td>-0.09 +</td>
<td>0.19 ***</td>
<td>0.19 ***</td>
</tr>
<tr>
<td>English literacy</td>
<td>0.07</td>
<td>0.07</td>
<td>-0.26 +</td>
<td>-0.26</td>
</tr>
<tr>
<td>Numeracy</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.09 **</td>
<td>-0.09 **</td>
</tr>
<tr>
<td>Age</td>
<td>0.27 ***</td>
<td>0.26 ***</td>
<td>0.09</td>
<td>0.09</td>
</tr>
<tr>
<td>Marital status (ref. Never married)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>-0.34</td>
<td>-0.34</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>Previously married</td>
<td>0.29</td>
<td>0.28</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Mother's education (ref. None)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary</td>
<td>0.06</td>
<td>0.05</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>Secondary</td>
<td>0.34</td>
<td>0.33</td>
<td>0.11</td>
<td>0.11</td>
</tr>
<tr>
<td>Father's education (ref. None)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary</td>
<td>0.03</td>
<td>0.03</td>
<td>-0.01</td>
<td>-0.01</td>
</tr>
<tr>
<td>Secondary</td>
<td>0.04</td>
<td>0.04</td>
<td>-0.08</td>
<td>-0.08</td>
</tr>
<tr>
<td>Mother's co-residence (ref. Dead)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Co-resident</td>
<td>0.11</td>
<td>0.13</td>
<td>-0.16</td>
<td>-0.16</td>
</tr>
<tr>
<td>Not co-resident</td>
<td>-0.11</td>
<td>-0.11</td>
<td>0.15</td>
<td>0.15</td>
</tr>
<tr>
<td>Father's co-residence (ref. Dead)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Co-resident</td>
<td>-0.25</td>
<td>-0.24</td>
<td>0.09</td>
<td>0.09</td>
</tr>
<tr>
<td>Not co-resident</td>
<td>-0.28</td>
<td>-0.28</td>
<td>-0.23</td>
<td>-0.23</td>
</tr>
<tr>
<td>Yao</td>
<td>0.17</td>
<td>0.17</td>
<td>0.14</td>
<td>0.14</td>
</tr>
<tr>
<td>Household assets</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.02</td>
<td>-0.02</td>
</tr>
<tr>
<td>Urban</td>
<td>-0.54 *</td>
<td>-0.51 +</td>
<td>-0.30</td>
<td>-0.29</td>
</tr>
<tr>
<td>Ever moved (ref. Never)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>At least once</td>
<td>0.22</td>
<td>0.23</td>
<td>0.21</td>
<td>0.21</td>
</tr>
<tr>
<td>Missing</td>
<td>-0.27</td>
<td>-0.24</td>
<td>-0.02</td>
<td>-0.01</td>
</tr>
<tr>
<td>Ever unpaid work</td>
<td>-0.03</td>
<td>-0.04</td>
<td>-0.03</td>
<td>-0.03</td>
</tr>
<tr>
<td>Ever paid work</td>
<td>-0.20</td>
<td>-0.20</td>
<td>0.06</td>
<td>0.06</td>
</tr>
<tr>
<td>Term</td>
<td>Coefficient</td>
<td>p-value</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------------------</td>
<td>-------------</td>
<td>---------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean HSV-2 prevalence</td>
<td>0.05</td>
<td>***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-5.62</td>
<td>***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tested for HSV</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ramadan</td>
<td>-0.30</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exam year</td>
<td>0.27</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Long distance interview</td>
<td>-0.62</td>
<td>***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Currently married</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.90</td>
<td>***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rho</td>
<td>-0.31</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uncensored observations</td>
<td>892</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Censored observations</td>
<td>253</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log likelihood</td>
<td>3905.3</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Uncensored observations     892 892 951 951
Censored observations       253 243
Log likelihood              -3905.3 -12963.0 -3704.2 -10495.2


Reniers, G. *Divorce and remarriage in rural Malawi*. *Demographic Research*. Available at www.demographic-research.org


detection of antibodies against herpes simplex virus type 2 in African populations. *Journal of Clinical Microbiology*, 42(7), 2961-2965.

