The Link Between School Attendance and Sexual Activity in Malawi:

A Search for Mechanisms

Margaret Frye

University of California, Berkeley

Graduate Group in Sociology and Demography

Over the past two decades, two dramatic transformations have occurred in several countries across sub-Saharan Africa: the rise in the proportion of adolescent sexual behavior taking place outside of marriage, and the widespread increase in school enrollment. Demographic and Health Surveys and other data sources show that throughout the region, current adolescents are much more likely to remain unmarried until after their eighteenth birthday than were earlier cohorts (B. S Mensch, Singh, and Casterline 2005). Rates of adolescent sexual activity have remained relatively constant, leading to an increasing proportion of first sexual experiences to occur outside of marriage (Barbara S. Mensch, Grant, and Blanc 2006). Second, school enrollment rates have increased quickly and dramatically across the region, in the wake of the global Education for All Initiative (Lloyd 2006). In Malawi, in response to the eradication of primary school fees in 1994, school enrollment almost doubled in the two years following the legislation, and continued to expand in subsequent years (Samer and Hassan 2007).

In this context of rapid social transformation, it is essential to unpack the complex relationship between schooling status and premarital sexual behavior. Widespread evidence indicates that a lower proportion of students are sexually active than their non-schooling counterparts, and wait longer to have sex for the first time (Kaufman et al. 2004; Lloyd 2005; Hargreaves et al. 2008; McGrath et al. 2009). Yet we know little about the mechanisms underlying this relationship. In this paper, I explore four pathways that might explain the process through which being in school leads to lower rates of sexual activity and delayed sexual initiation: the development of critical thinking and cognitive abilities leading to different patterns of decision-making, changes in attitudes towards sexual practices and HIV risk as a result of school-based sexual education programs, the anticipation of future opportunities which leads to avoidance of

sexual experiences, and the separation of in-school and out-of-school peer networks, resulting in a different set of norms and social influences regarding sexual behavior. Using new data from a longitudinal survey on youth in Malawi, I test whether any of these mechanisms can be show to mediate the relationship between educational experiences and sexual behavior.

First, I explore whether the accumulation of knowledge, critical thinking, and cognitive skills might lead to unique patterns of decision-making for those with more schooling. To examine this relationship, I model the effects of education in two ways: first, as a continuous variable measuring years of schooling completed, and second, as a categorical variable measuring current school status: out of school, in primary school, or in secondary school. If education is operating primarily through increasing the cognitive ability or critical thinking of students, we can expect a "dose-response" relationship, where students with more years of education would exhibit more substantial delays in onset of sexual activity.

Second, I examine whether the relationship between sexual activity and education might be operating through changes in attitudes toward sexual behavior or HIV risk. School-based programs are a major mode of transmission of information on HIV prevention (Douglas Kirby 2002). However, evidence on whether school-based "life skills programs" alter the sexual behavior of youth is mixed. A review of 11 studies examining the effects of school-based programs in sub-Saharan Africa concluded that while many studies show evidence of attitudinal changes, but no consistent effects on sexual behavior (Gallant and Maticka-Tyndale 2004)). On the other hand, a recent review of over 60 studies found evidence that certain school-based programs were associated with delayed sexual activity and increased contraceptive among students (D Kirby, Laris, and Rolleri 2007). To examine the extent to which knowledge of HIV is operating, I analyze a series of attitudinal questions about HIV risk and sexual behavior and

their association with both educational status and sexual behavior, testing for possible mediation effects.

Third, I examine the effect of youth's imagined futures on the relationship between school status and sexual behavior. In her dissertation on sexual behavior of in-school versus out of school youth in Malawi, Michelle Poulin attributes lower levels of risky sexual activity among in-school women to their "perceptions that a better future awaits" those who delay marriage and avoid premarital sex. Indeed, a set of qualitative interviews with schoolgirls conducted by the author in 2008 revealed that commitment to future ambitions is intimately linked to avoidance of sexual activity and romantic partnerships among this population. In addition, previous work in Malawi has shown that marital aspirations are linked to sexual activity among youth (Clark, Poulin, and Kohler 2009). If school status influences marital aspirations, this may be a component of the link between schooling and sexual behavior. I use two measures of respondent's future outlook: their expectations of remaining in school and their desired timing of marriage.

Finally, I examine the effect of peer networks- it is possible that in-school and out-of-school youth have distinct patterns of sexual activity because they have separate social communities, in which unique norms proliferate. A review of the literature on adolescent sexual activity in sub-Saharan Africa highlights the role of peers and other "proximate" social influences in shaping the sexual practices of youth, both positively and negatively (Eaton, Flisher, and Aarø 2003). In Malawi, Poulin (2007) finds that the sexual networks of youth are sharply delineated by school status: out-of-school youth rarely report interacting with school-goers. She also finds that while both in-school and out-of-school youth discuss sexual relationships amongst their friends, the content of these conversations was found to be quite

different: schoolgirls reported frequently advising each other against sexual relationships, and routinely discuss the dangers of premarital sex. I examine the effects of the reported sexual behavior of both intimate friendships and more distant social networks.

Methods

The data for the analysis come from Tsogolo la Thanzi (TLT), a longitudinal survey in Balaka, Malawi designed to study how young people navigate the transition to adulthood in an AIDS epidemic. The first wave of data collection took place between May and August 2009. Fifteen hundred female respondents were randomly selected from a sampling frame of 15 to 24 year olds living in census enumeration areas within 7 kilometers of Balaka, Malawi. The catchment area includes a mix of rural and peri-urban communities around Balaka, a growing town one and a half hours from the southern city of Blantyre. In addition to the core sample of women, TLT used respondent driven sampling (RDS) to enroll the romantic partners of our survey respondent into the study, and successfully enrolled about 450 partners. A random sample of men between the ages 15 and 24 (N=600) allows for comparisons between the partner sample and a random sample of men, in order to identify any key differences.

One particularly unique feature of TLT is the use of a centrally located research center for conducting interviews. Respondents were first contacted in their homes and asked to set up a time for an interview. On their assigned day (or more accurately close to it), respondents came to the research center and were interviewed in a private room where their responses could not be overheard by family members or neighbors. The survey took approximately one and a half hours to complete. Refusal at the time of making an appointment and passive refusal by not showing up at the research center were relatively rare (our current estimate is <5%).

A total of 1492 women completed the Wave 1 interview. Our analyses also include 444 male partners who enrolled in the study through RDS and a random sample of 575 men, who were selected for participation in exactly the same way the women's sample was drawn. The sample for this study is restricted to never-married respondents (N=1,311) who have a parity of 0 (N=1,210). Because of the need for a consistent sample size in our analysis, respondents with missing values for one of the key explanatory variables were also dropped from the sample (N=54) reducing the sample size to 1,156. In the future, we will explore the use of imputation techniques to keep respondents with incomplete data in our study sample.

Table 1 gives summary statistics of the sample. The unmarried, nulliparous sample is about 45% female. It is important to note that a large proportion of respondents are several years behind where they would be in schooling if they had begun at age 6, the target age for entering primary students, and continued without interruption. Thus, while the sample is restricted to respondents above age 15, an age at which all youth who are "on track" for their age would be in secondary school, more respondents are currently in primary school (N=466) than are in secondary school (N=430). Indeed, the average age of primary school students in the sample is 16, and 52 respondents (11% of those in primary school) were 18 or older at the time of the survey.

There are significant gender differences in our sample in terms of school attendance, with women substantially more likely to be in primary school, and men more likely to be out of school. This is likely to be a factor of the fact that the men's sample was partially made up of the romantic partners of women, who are more likely to be out of school. Out of those currently in school, men are more likely to be in secondary school than are women, though this difference is

not significant. Out-of-school respondents are significantly older than in-school, and primary school respondents are younger on average than secondary school respondents.

The differences in SES score between the three educational groups were all statistically significant at the .001 level using a one-tailed t-test, with secondary school respondents reporting the highest SES score on average, and respondents currently in primary school reporting the lowest SES score. This is not surprising, as those respondents currently in primary school are all several years behind the target grade level for their age.

The Yao are a matrilocal, predominantly Muslim ethnic minority in Malawi, living predominantly in the Southern part of the country, who have been documented to have both significantly lower rates of education, higher rates of HIV, and less stable marriages than other ethnic groups (Helleringer and Kohler 2005; Poulin 2007; Swidler and Watkins 2007). In this sample, Yao respondents are significantly less likely to be in secondary school, although the difference between proportion Yao in primary school and out of school was not statistically significant.

The proportion reporting that they had "ever had sex" was much higher for the out of school sample, with primary school respondents reporting the lowest levels of sexual activity. All differences in proportion sexually active between the three categories are statistically significant at the .01 level. We can see from these simple summary statistics that schooling status and sexual behavior are indeed linked in this sample. It is essential to unpack these differences further, to determine what mechanisms are mediating between educational status and sexual behavior.

I explore the relationship between schooling and sexual activity using two outcomes: first, binary variable for having ever had sex, and second, the age of onset of sexual activity. All

sexual activity data are self-reported for this study. This method carries with it a risk of unreliable reporting: a study comparing various methods of collecting data on sexual activity found face-to-face survey questionnaires to generate lower levels of reported sexual activity (Plummer et al. 2004). However, recent evidence indicates that by restricting the analysis t reports of having ever had sex, as opposed to more invasive questions, we may be minimizing the bias produced from this method of generating data. A study in Malawi that compared face-to-face interviewing with ACASI self-reported data, demonstrates that face-to-face interviewing produced higher rates of reporting for "ever had sex" and "sex with a boyfriend" (Barbara S Mensch et al. 2008). However, face-to-face interviewing was associated with lower reporting of sexual activity with older men and sexual intercourse with a stranger, indicating that the reports of risky sexual activity may be underestimated for this sample. To minimize rates of under-reporting of sexual activity, interviews were conducted in private rooms within the research center, and staff were extensively trained on methods of building rapport with respondents. All interviewers were trained and certified in Voluntary Testing and Counseling (VCT) methods.

To explore the impact of sexual education programs on sexual activity, I examine the relationship between response to several attitudinal questions and the difference in reported sexual activity by school status. The attitudinal questions examined here were all read as part of a long list of statements to which respondents were asked to either agree or disagree. These questions include four that probe the respondent's view of how AIDS is transmitted: "*AIDS is in the flour*," a common saying which communicates that AIDS is all around and there is no way to actively prevent AIDS, "*HIV is God's lottery*," a saying that communicates that God determines who will contract AIDS and who will not, "*Nowadays a man who gets AIDS is deliberately choosing death*," and "*If you are hated by somebody, they can create AIDS for you*," which

refers to the belief that AIDS can be caused by witchcraft. I also examine three statements that demonstrate attitudes towards sexual activity: *"Sex before marriage is acceptable if the couple loves each other," "A man feels proud if he has many sex partners,"* and *"Sex with condoms is not sweet,"* which evokes a common metaphor used to invoke feelings of sexual frustration related to condoms (Tavory and Swidler 2009).

To explore the effect of attitudes towards the future, I examine the impact of respondent's expectations for two different future events: desired timing of marriage and perceptions of their likelihood of continuing their education. I measure desired marriage timing using the following question, asked of all unmarried survey respondents: *"When do you want to get married?"* Possible answer choices include: as soon as possible, less than two years, two to three years, three to four years, four to five years, five or more years, and don't want to get married. Respondents answering "no preference" and "I don't know" were dropped from these analyses (N=5). I created a standardized scale for responses to this question, and used this in my analyses.

To examine the possibility that in-school youth have lower rates of sexual activity because of the "bright futures" that they imagine for themselves through schooling, I also analyzed the effect of probabilistic estimates of respondent's likelihood of continuing schooling. This question uses an innovative method, piloted in Malawi, to collect probabilistic estimates of future outcomes (Delavande and Kohler 2009). Respondents are asked choose a specific amount of beans from a pile, to represent the probability that a specific outcome will occur. This technique provides a visual representation of an otherwise abstract concept, and mimics the game of *bawo*, a popular pastime in Malawi, in which beans are moved from one dish to another. It has been shown to provide strikingly realistic estimates of child mortality and infection rates, demonstrating that rural Malawians can assess probabilities accurately, despite high rates of illiteracy (Delavande and

Kohler 2009, De Paula, Shapira and Todd 2008). The specific question analyzed here reads: *"How likely is it that you will be enrolled in school one year from now?"* All answers to this question were centered around the mean. Because expectations of future schooling is likely to affect sexual activity differently according to current educational status, I also analyzed the interaction between this variable and current educational status.

Finally, to examine the effect of peer networks, I explored two different measures. First, to attempt to investigate the effect of more diffuse social network ties, I looked at whether or not respondents agreed with the question: *"My friends have changed their sexual behavior to prevent HIV."* Second, to investigate the effects of closer, more intimate friendships, I looked at the question: *"How many men do you think your best female friend has slept with in the past 12 months?"* For use in the statistical models, responses to this question were standardized around the mean response.

Because this paper seeks to examine the potential mediating effects of various mechanisms on the relationship between educational status and sexual behavior, underlying differences between in-school and out-of-school samples with regard to individual-level background variables should be minimized. For this reason, four background-level control variables are used: birthyear, gender, socioeconomic status, and a dummy variable for Yao ethnicity. Birthyear, of course, is equivalent ot age at the time of the survey; this term is used to differentiate this variable from the age-at-risk measure used in the hazard analysis. To measure socioeconomic status, an index was created out of 11 household goods commonly owned in Balaka- the proportion of these goods owned by the respondent makes up the value of this index. The household goods store was standardized before entering it into regression models. Finally, because the Yao have been shown to have different levels of education and are also at a higher risk for HIV, I control for Yao ethnicity, as it seems

likely that underlying differences in culture, access to resources, or social networks between the Yao and the rest of the population in Balaka might also mediate the relationship between educational status and sexual behavior.

To model the relationship between independent variables and the binary variable for sexual activity, a simple logistic regression function was used. To examine the effect of independent variables on the timing of sexual intercourse, discrete-time hazard models are used, following the method proposed by Allison (Allison 1982). This model provides a way to estimate hazard models and also allows hazard rates to be estimated for time intervals that are large enough for multiple events to be recorded for the same time period. In this case, age at first sex is recorded in years; thus, the proportional hazard model would be inappropriate for this purpose. This method has been commonly used in the literature to study the onset of sexual activity (Brewster 1994; Whitbeck et al. 1999; Dupéré et al. 2008). For the hazard models, the effect of age at risk is represented by a quadratic smooth function, rather than entering each risk-age separately into the model; the smoothed approach minimizes the number of degrees of freedom lost, and allows for more parsimony (Treiman 2008)**.** All analysis was performed using the Stata 10 software package.

Results

Mechanism 1: The accumulation of Knowledge or Critical Thinking

To explore the possibility that schooling shapes sexual activity through the accumulation of knowledge or skills, which leads to different decision-making patterns, I explored whether or education has a "dose-response" effect on sexual behavior. I hypothesize that if schooling operates through the mechanism of accumulated skills or knowledge, that those who have been exposed to more education, and thus presumably have had more opportunities to improve their

critical thinking or cognitive ability, will have lower rates of sexual activity than those who have been exposed to less schooling. Thus, years of schooling, when measured as a continuous variable, should have a negative effect on the likelihood of having ever had sex, and a negative effect on the age of initiation of sexual intercourse.

I test the hypothesis that years of accumulated education has an effect on likelihood of having ever had sex using two logit regressions, examining the effect of the two different measures of school status on sexual activity, controlling for individual background factors including Yao ethnicity, age, gender, and socioeconomic status (Table 2). In the first model (first column), education is measured as a continuous variable of years of schooling completed, and in the second model (2nd column), education is measured as a categorical variable indicating current schooling status: out of school, in primary school, or in secondary school. The results show that accumulated schooling has no effect on the likelihood of ever having sex. A likelihood ratio test demonstrates that adding this measure of schooling to the model has no effect (LR Chi Squared = 0.1, p value=0.75). Further, this model has a lower overall significance, despite the lower number of model degrees of freedom. On the other hand, current schooling status, measured categorically, has a strong effect on likelihood of having ever had sex, as demonstrated by the likelihood ratio test of the effect of adding these variables to the model (LR Chi Squared =19.54, p value<0.01).

I test the hypothesis that years of accumulated education has an effect on the timing of first sexual intercourse by comparing two discrete time series logit models: one with years of completed education measured as a continuous variable, and the other with current educational status measured as a categorical variable (Table 2, Columns 3 and 4). The coefficient for the years of education variable is not significant, while the categorical variables for current

educational status yield highly significant results. Because observations for risk years are clustered by individual respondents, a likelihood ratio test is not appropriate for comparing model fit between discrete-time hazard models. For this reason, adjusted Wald tests are used and reported in Table 3. The effect of adding the continuous measure of years of education is negligible, while the effect of adding the categorical measures of current schooling status are highly significant.

These models yield consistent and convincing evidence that education, when measured as a continuous variable, does not have a significant effect on age at sexual intercourse. Further, when we examine the model coefficients for the two different current educational statuses, we see that they are almost identical. This indicates that it is something to do with going to school, rather than the level of education attained, that reduces sexual activity among students. Thus, for the remainder of this paper, education will be measured as a categorical variable, with distinctions made between those currently not schooling, those currently in primary school, and those currently attending secondary school.

Criteria for Establishing Mediation Effects

Now that we have ruled out the possibility of a dose-response relationship between education and sexual activity, it is time to examine the three remaining mechanisms: attitudes toward sex or the prevention of HIV/AIDS, anticipation of future events, or peer networks. All could hypothetically be influenced by current school status, and thus are consistent with our new focus on educational status as an indicator of current status, rather than accumulated years of schooling. In order to examine whether or not any of these mechanisms operate as a mediator between educational status and sexual behavior, I will employ the four criteria which are

commonly recognized as necessary in order for a mechanism to act as a mediator (Baron and Kenny 1986; MacKinnon, Fairchild, and Fritz 2007). I will examine each of the above criteria in turn, and rule out those mechanisms found to not fulfill each criterion.

(1) The independent variable (educational status) must significantly affect the dependent variable (sexual activity) in the absence of the mediator.

(2) The independent variable (educational status) must significantly affect the mediator.

(3) The mediator must have a significant unique effect on the dependent variable (sexual activity)

(4) The effect of the independent variable (educational status) on the dependent variable (sexual activity) must shrink upon the addition of the mediator to the model.

1st Criterion: Educational Status's Significant Effect on Outcome Variable

The results of the analyses performed to rule out the first mechanism, the accumulation of skills and knowledge, offer convincing evidence that educational status has a significant effect on sexual activity (Table 2). In the second column of this table, we see that, when sexual activity is measured as a dichotomous variable, distinguishing those who have never had sex from those who have had sex at least once, both being in secondary school and being in primary school have significant negative effects on the likelihood of having sex, independent of the individual-level control variables. While the coefficient for the effect of being in primary school is slightly larger in magnitude, the standard error for this term is also slightly larger: in general, there is no significant difference between the effect of primary school compared to secondary school on likelihood of having sex. As the fourth column tells us, when we examine the timing of sexual

onset, through the use of a discrete-time hazard model, we see that both current secondary school enrollment and current primary school enrollment are associated with significantly younger age at first intercourse. Once again, the coefficient of the primary school variable is slightly more negative, but the standard error is also slightly larger, and there is no significant difference between these two effects.

2nd Criterion: Effect of independent variable on the mediator

Next, we must establish that educational status and the mediator variable are significantly correlated. I present bivariate relationships between the values for potential mediators broken down by educational status, and test the significance using simple t-tests (Table 3). In the case of marital aspirations, because the mean value was not a meaningful metric, the median value was used, and the Wilcoxin median test was used to establish the significance of difference between the values for each educational status.

For the attitudinal measures, with the exception of the statement "*Nowadays a man who gets HIV is deliberately choosing death,*" for which no significant differences exist between categories of education, secondary school students respond differently to all attitudinal statements. In each case, the response given by secondary students is closer to the information likely to be communicated through life skills programs, and other school-based HIV prevention programs: secondary school students are less likely to say that AIDS comes from sources such as God, witchcraft, and fate, and less likely to say that sex with condoms is less pleasurable, and less likely to say that sex before marriage is acceptable. Thus, it seems that, in terms of attitudes communicated during the survey, there are marked differences between secondary school students and other survey responses.

In terms of anticipated future events, significant differences were found by educational status for both expected probability of future schooling and desired marital timing. It is clear that overall, Malawian youth are highly optimistic about their likelihood of being in school: even out of school youth give themselves a 40% chance, on average, of being in school next year, and in-school youth, at both the primary and secondary level, give themselves above an 80% chance on average. Yet despite this overall optimism, there is considerable variation in the sample with 114 in-school youth responding with a 50% chance or lower of being in school next year, and 78 out-of-school respondents estimating that they have above a 50% chance of returning to school. In terms of desired timing of marriage, in-school youth tend to want to get married in the relatively distant future, in 5 or more years, while those out-of-school generally intend to get married somewhat sooner. Once again, considerable variation exists, with 129 in-school respondents wanting to get married in the next three years, and 50 out-of-school respondents wanting to wait 5 or more years before getting married.

The bivariate examination of peer networks reveals mixed results. It is clear that responses to the number of sexual partners reported by the respondent's best friend varies substantially, with out of school youth reporting significantly higher numbers than in-school youth. On the other hand, levels of agreement with the statement *"My friends have changed their sexual behavior to prevent HIV"* do not differ significantly by educational status: between 75 and 80 percent of all respondents agree with this statement. Thus, it appears that the behavior of close friends, or the respondents' perception of their behavior, does seem to have a significant effect on sexual behavior, and does vary by socioeconomic status. However, the more general reference to friends' sexual behavior does not seem to be correlated. It seems that youth are less

strongly influenced by the sexual behavior of more distant social network ties than they are by the behavior of their closest friends.

Next, I examine the relationship between the independent variable and the mediator using regression analysis, to ensure that the relationships detected in the bivariate analyses are not driven primarily by underlying individual background factors. I constructed a regression analysis with each potential mediator as the outcome variable, with the following independent variables: current educational status, age, gender, socioeconomic status, and Yao ethnicity. Results of these analyses are presented in Table 4. At this point, we can rule out four of our potential mediating variables, due to their lack of a significant relationship between the mediators and the independent variable. Three are attitudinal: *"AIDS is in the flour," "Nowadays a man who gets HIV is deliberately choosing death,"* and *"Sex with condoms is not sweet*." We can also rule out the general measure friends' behavior change in response to HIV. All other measures analyzed here satisfy the second criteria: educational status has a significant effect on these outcomes, even when controlling for individual background variables.

3ʳᵈ Criteria: Mediators' Effects on Sexual Activity

The third criteria is that the mediator variables in question must have significant and unique effects on the outcome, sexual activity. To test whether this is true, all remaining candidates for mediation effect were added one by one to regression equations with school status and the four individual control variables. Results are presented in Table 5. At this point, we can rule out three more attitudinal measures, which do not have significant effects on either sexual activity outcome, when educational status is controlled for. These are: *"HIV is God's lottery," "If you are hated by somebody, they can create AIDS for you,"* and *"A man feels proud if he has*

many sex partners." At this point, only one attitudinal statement remains: *"Sex before marriage is acceptable if a couple loves each other*." While it is possible that attitudes specifically toward premarital sex mediate between education and school status, the broader hypothesis that school-based HIV prevention programs create a unique perspective on HIV risk and sexual activity, can at this point be effectively ruled out. All other outcomes had significant, independent effects on both measures of sexual activity, thus they remain candidates for mediation effects.

4th Criteria: Reduction in the Effect of Educational Status when Mediators are Added

The fourth criteria is that The effect of the independent variable (educational status) on the dependent variable (sexual activity) must shrink upon the addition of the mediator to the model. To test this criteria, we will first examine the size of the effect of schooling status when each potential mediator is added to the model (Table 6). The coefficients for each model are to be compared to the corresponding coefficient in the baseline model, the first row in table six (these numbers appear in bold). If the coefficients for the models with the mediator variables are not smaller, then the 4th criteria is not met. Immediately, we see that the final attitudinal variable, *"Sex before marriage is acceptable if a couple loves each other,"* can now ruled out as a possible mediator of the relationship between educational status and sexual behavior: the effect of educational status does not get smaller with the addition of this variable into the model, with the exception of primary school students in the model of ever having had sex. Even in this case, however, the decrease in the effect is minimal, thus we can assume that attitudinal effects do not mediate between educational status and sexual behavior.

To effectively establish a significant mediation effect, a more formal test is required. Because both outcomes are estimated using logit models, the comparison of regression

coefficients to determine the strength of the mediation effect is not possible, because the log-odds transformation mutes the effect of extreme values in the probability distribution (Buis 2008; MacKinnon et al. 2007). Thus, I will use a method introduced by Erikson et al. (2005) for composing the total effect in a logit model into direct and indirect effects. A version of this method was developed by Buis (2008) for use in the Stata software package.

This method should hypothetically work for any situation in which mediating variables are sought to explain an underlying relationship between two variables. Erickson originally developed this method to explain the relationship between social class, school performance, and the likelihood entering college . The basic idea underlying this technique is to compare the log odds of "success" in one category of the direct variable with the log odds of success of a counterfactual group. In this case, imagine that there is a simple mediating relationship predicted:

$$X \rightarrow Z \rightarrow Y$$

There are two different values of X, x1 and x2. According to this method, the log odds calculated for group x1 are compared to a counterfactual group, who is assigned the distribution of Z values for x1, but the probability of achieving outcome Y, conditional on Z, as the x2 group. To use an example from this paper, imagine that the out-of-school group is x1, and the mediating variable of interest, Z, is the respondent's desired marital timing. The counterfactual in this case would be an imaginary sample which is assigned the same distribution of responses for desired marital timing as out-of-school youth, but the probability of having sex, conditional on marital timing, as the in-school group. This way, the counterfactual and factual groups only differ with respect to X, their schooling status, and not with respect to Z, the mediating variable, and thus the indirect effect can be easily calculated.

In another version of the same method, it is the distribution of the mediating variable that changes, and the conditional probability with respect to Z that stays the same. To use the same example, the counterfactual in this case would be a group with the same conditional probability of having sex, given an specific response for desired marital timing, calculated for the out-of-school group, but the distribution of marital aspirations taken from the in-school sample. The two specifications yield very close, but not identical, results. Both Buis (2008) and Jackson et al. (2007) recommend computing the relative size of the indirect effect using each version, and reporting the average of the two; I follow this suggestion here.

To simplify interpretation of this method, and because primary school and secondary school has been found to have markedly similar effects throughout the analysis thus far, the school status outcome variable will be converted into a dummy, delineating in-school versus out-of-school status. This decomposition method allows for the use of either continuous or categorical mediator variables, and allows for the use of other covariates, so I include the four individual-level explanatory variables used in other models in these tests. Because this function does not allow the use of clustered models, and it is unreasonable to assume that observations from the same respondent are independent of one another in the hazard model, the mediation effects are explored only for the logit regression using the dichotomous variable of whether respondents have ever had sex as the outcome. I generated confidence intervals using bootstrap sampling, with 500 repetitions. Results from these analyses are reported in Table 7, with the proportion of the total effect of education explained through each mediating variable, modeled separately.

Attitudes about premarital sex were not found to be a significant mediator using this test, which confirms the intuition gained through comparing the effect size of the education variable.

The other three potential mediators were found to explain a substantial proportion of the total effect of being in school. In particular, marital aspirations were found to have the strongest effect, explaining over 60% of the total effect of being in school.

Summary of Results

We are left with three significant mediation effects: respondents' probabilistic estimate of their likelihood of being in school next year, their desired timing of marriage, and the number of sexual partners reported by their closest friend. Out of the four mechanisms proposed in this paper, these data offer evidence that two appear to be operating: anticipated timing of future events and the influence of peers. Table 8 presents the results of the logistic regression models of the "ever sex" variable for the baseline model, with each mediator added individually, and with all three mediators added together. It is clear that these variables do indeed appear to mediate the relationship between educational status and sexual activity. Table 9 presents a similar set of results, using the discrete-time hazard model to examine the effect of each mediator on the relationship between schooling and age at first sex.

To provide a visual picture of how these measures mediate the effect of educational status, a series of graphs of predicted probabilities calculated from the hazard models are presented in Table 9 (Figures 1-5). Figure 1 presents the predicted probabilities of first sex, by age at risk, for the baseline model. Figure 2 presents the model with educational expectations added, Figure 3 with marital aspirations, Figure 4 with the peer sexual activity measure, and Figure 5 with all 3 mediator effects added to the model. As one quickly see from comparing Figures 2-5 with Figure 1, the 3 mediator effects do indeed attenuate the difference in the distribution of age at first sex between the in-school and out-of-school populations. Further these

graphs demonstrate visually the fact that differences in the age pattern of sexual activity between primary and secondary school students are minimal, once background variables are controlled for.

Discussion

This study indicates that peer influence and future orientation may mediate the relationship between school status and sexual behavior. On the other hand, the other two mechanisms examined, attitudes about safe sexual practices and HIV risk and accumulated knowledge and cognitive skills, do not appear to operate in this context. Peer influence appears to operate primarily through close relationships, and not through more diffuse social networks.

These analyses are somewhat limited by the cross-sectional nature of the data at this point in the survey. Indeed, the evidence from this study should be taken to be indicative of possible relationships, rather than conclusive. Because all data were self-reported and collected at the same time, it is impossible to empirically verify the direction of causality. Further analyses with later waves of data will allow me to examine time lags between hypothesized causal factors and sexual activity.

Indeed, it is possible that rather than school status leading to different patterns of sexual activity, abstinence and delayed marriage are instead the primary causal mover in the link between school status and sexual activity. However, it is difficult for me to think of a causal pathway flowing in this reverse direction that would explain such dramatic differences in rates of sexual activity and not be linked to school achievement, rather than the state of either being a student or being out of school. For example, if students who are less likely to have sex are also more likely to work hard and be diligent about their studies, we would expect more students in

higher grades to have lower levels of sexual activity than students in lower grades. The fact that all in-school youth have a similar likelihood of sexual activity, regardless of grade level, indicates that there is something about the lived experience of being in school compared with being out of school that is driving the different patterns of sexual activity.

Works Cited

Allison, P. D. 1982. "Discrete-time methods for the analysis of event histories." *Sociological methodology* 13:61–98.

Baron, Reuben M., and David A. Kenny. 1986. "The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations.." *Journal of Personality and Social Psychology* 51:1173-1182.

Brewster, K. L. 1994. "Race differences in sexual activity among adolescent women: The role of neighborhood characteristics." *American Sociological Review* 59:408–424.

Buis, M. L. 2008. "Direct and indirect effects in a logit model." *The Stata Journal*.

Clark, Shelley, Michelle Poulin, and Hans-Peter Kohler. 2009. "Marital Aspirations, Sexual Behaviors, and HIV/AIDS in Rural Malawi." *Journal of Marriage and Family* 71:396-416.

Dupéré, V., É Lacourse, J. D Willms, T. Leventhal, and R. E Tremblay. 2008. "Neighborhood poverty and early transition to sexual activity in young adolescents: a developmental ecological approach." *Child Development* 79:1463–1476.

Eaton, Liberty, Alan J. Flisher, and Leif E. Aarø. 2003. "Unsafe sexual behaviour in South African youth." *Social Science & Medicine* 56:149-165.

Gallant, Melanie, and Eleanor Maticka-Tyndale. 2004. "School-based HIV prevention programmes for African youth." *Social Science & Medicine* 58:1337-1351.

Hargreaves, J R et al. 2008. "The association between school attendance, HIV infection and sexual behaviour among young people in rural South Africa." *Journal of Epidemiology and Community Health* 62:113-119.

Helleringer, Stéphane, and Hans-Peter Kohler. 2005. "Social Networks, Perceptions of Risk, and Changing Attitudes Towards HIV/AIDS: New Evidence from a Longitudinal Study Using Fixed-Effects Analysis." *Population Studies* 59:265-282.

Kaufman, Carol E., Shelley Clark, Ntsiki Manzini, and Julian May. 2004. "Communities, Opportunities, and Adolescents' Sexual Behavior in KwaZulu-Natal, South Africa." *Studies in Family Planning* 35:261-274.

Kirby, D, B Laris, and L Rolleri. 2007. "Sex and HIV Education Programs: Their Impact on Sexual Behaviors of Young People Throughout the World." *Journal of Adolescent Health* 40:206-217.

Kirby, Douglas. 2002. "The Impact of Schools and School Programs upon Adolescent Sexual Behavior." *The Journal of Sex Research* 39:27-33.

Lloyd, C. B. 2005. *Growing up global: the changing transitions to adulthood in developing countries*. National Academy Press.

Lloyd, C. B. 2006. "Schooling and adolescent reproductive behavior in developing countries." *background paper to Public Choices, Private Decisions: Sexual and Reproductive Health and the Millennium Development Goals, United Nations Millennium Project, New York.*

MacKinnon, David P., Amanda J. Fairchild, and Matthew S. Fritz. 2007. "Mediation Analysis." *Annual Review of Psychology* 58:593-614.

McGrath, N, M Nyirenda, V Hosegood, and M-L Newell. 2009. "Age at first sex in rural South Africa." *Sexually Transmitted Infections* 85:i49-i55.

Mensch, B. S, S. Singh, and J. Casterline. 2005. "Trends in the timing of first marriage among men and women in the developing world." *The changing transitions to adulthood in developing countries: Selected studies* 118–171.

Mensch, Barbara S, Paul C Hewett, Richard Gregory, and Stephane Helleringer. 2008. "Sexual behavior and STI/HIV status among adolescents in rural Malawi: an evaluation of the effect of interview mode on reporting." *Studies in Family Planning* 39:321-334.

Mensch, Barbara S., Monica J. Grant, and Ann K. Blanc. 2006. "The Changing Context of Sexual Initiation in Sub-Saharan Africa." *Population and Development Review* 32:699-727.

Plummer, M. L. et al. 2004. ""A bit more truthful": the validity of adolescent sexual behaviour data collected in rural northern Tanzania using five methods." *British Medical Journal* 80.

Poulin, Michelle. 2007. "The Sexual and Social Relations of Youth in Rural Malawi: Strategies for AIDS Prevention." *Ph.D. Dissertation*.

Swidler, A., and S. C Watkins. 2007. "Ties of dependence: AIDS and transactional sex in rural Malawi." *Studies in Family Planning* 38:147–162.

Tavory, I., and A. Swidler. 2009. "Condom Semiotics: Meaning and Condom Use in Rural Malawi." *American Sociological Review* 74:171–189.

Treiman, Donald J. 2008. *Quantitative data analysis*. Jossey-Bass.

Whitbeck, L. B, K. A Yoder, D. R Hoyt, and R. D Conger. 1999. "Early adolescent sexual activity: A developmental study." *Journal of Marriage and the Family* 61:934–946.

Table 1: Descriptive Statistics of the Sample

	Out of School	In Primary	In Secondary	Total Sample
N	260	466	430	1156
Proportion Male	0.58	0.4	0.43	0.45
Median Age	20	16	18	17
Mean SES score (out of 17)	12.01	11.6	12.8	12.12
Proportion Yao	0.33	0.29	0.23	0.28
Percent Sexually Active	0.76	0.39	0.48	0.51

Table 2: Effect of Education, Measured as Cumulative Years or Current Schooling Status, on Sexual Behavior

	Logit: Ever had Sex				Hazard: Age of 1st Sex			
	Years of Ed.		Current Ed.		Years of Ed.		Current Ed.	
	b/se		b/se		b/se		b/se	
Male	1.09	***	1.13	***	0.85	***	0.86	***
	(0.14)		(0.14)		(0.10)		(.10)	
Age	0.25	***	0.19	***	-0.11	***	-0.15	***
	(0.03)		(0.04)		(0.02)		(0.03)	
SES	-0.05		-0.03		-0.08	+	-0.07	
	(0.04)		(0.04)		(0.05)		(0.05)	
Yao Ethnicity	0.31	*	0.27	+	0.29	**	0.25	*
	(0.15)		(0.15)		(0.10)		(0.10)	
Years of Schooling	0.01				0.01			
	(0.03)				(0.02)			
Out of School			Ref.				Ref.	
			---				---	
In Primary			-0.86	***			-0.35	**
			(0.22)				(0.12)	
In Secondary			-0.81	***			-0.28	**
			(0.20)				(0.09)	
Age at Risk					1.31	***	1.34	***
					(0.16)		(.16)	
Age at Risk Squared					-.03	***	-.03	***
					(.01)		(.01)	
Constant	-4.39	***	-2.79	***	-3.17	***	-2.47	***
	(0.65)		(0.82)		-0.27		-0.34	
Log Likelihood	-687.54		-677.82					
LR Chi Sq. (df)	217.38 (5)		236.83 (6)					
N	1155		1155					
Wald F (p)[1]					0.1 (.75)		5.91 (.002)	
Model F					55.25		49.17	
Respondents					1155		1155	
Risk Years					8221		8221	

Table 3: Response to Questions of Interest by Current Educational Status

Atttiduinal Questions About AIDS		Out of school	In Primary	In Secondary
"AIDS is in the flour"		0.24	0.23	0.18*
"HIV is God's lottery"		0.17	0.19	0.06***
"Nowadays a man who gets HIV is deliberately choosing death"		0.75	0.73	0.72
"If you are hated by somebody, they can create AIDS for you"		0.324	0.369	0.2***
Attitudinal Questions About Sexual Activity				
"Sex before marriage is acceptable if the couple loves each other."		0.27	0.285	.14***
"A man feels proud if he has many sex partners."		0.29	0.3	0.22*
"Sex with condoms is not sweet"		0.25	0.26	.18***
Timing of Future Events				
Probability of being in school next year (mean, out of 10)				
	Female	4.4***	8.6	8.5
	Male	3.7***	8.7	8.5
Desired time of marriage (median)				
	Female	2-3 years***	5+ years	5+ years
	Male	3-4 years***	5+ years	5+years
Peer Networks				
# of sexual partners of best friend (Mean)		1.46***	.83	.78
"My friends have changd their sexual behavior to prevent HIV"		.76	.76	.79

+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001, all starred results were significant when compared with both other   categories, using t-tests or Wilcoxon median tests, when appropriate.

Table 4: Results from Regression Analysis of Effect of Educational Status on Mediator Variables

Outcome	Logit Regression b (se)	
Attitudinal Questions About AIDS	Primary School	Secondary School
"AIDS is in the flour"	-.12 (.23)	-.24 (.21)
"HIV is God's lottery"	-.18 (.27)	-1.00*** (.27)
"Nowadays a man who gets HIV is deliberately choosing death"	.22 (.22)	.07 (.19)
"If you are hated by somebody, they can create AIDS for you"	-.07 (.21)	-.67** (.19)
Attitudinal Questions About Sexual Activity		
"Sex before marriage is acceptable if a couple loves each other."	-.07 (.23)	-.89*** (.23)
"A man feels proud if he has many sex partners."	-.29 (.22)	-.48* (.20)
"Sex with condoms is not sweet"	.09 (.23)	-.34 (.21)
	OLS Regression b (se)	
Timing of Future Events		
Probability of being in school next year	1.25*** (.08)	1.23*** (.07)
Desired time of marriage	.95*** (.08)	.93*** (.07)
Peer Networks		
# of sexual partners of best friend	-.40*** (.09)	-.44*** (.08)
"My friends have changed their sexual behavior to prevent HIV"	-.07 (.15)	0.22

Table 5: Test of Criteria 3, Unique and Significant Effects of Mediator on Outcomes

Mediator	Outcome	
	Ever Sex	Age 1st Sex
Attitudinal Questions About AIDS	b (se)	b (se)
"HIV is God's lottery"	-.01 (.19)	-.08 (.10)
"If you are hated by somebody, they can create AIDS for you"	.19 (.14)	.10 (.10)
Attitudinal Questions About Sexual Activity		
"Sex before marriage is acceptable if a couple loves each other."	.73*** (.16)	.41*** (.10)
"A man feels proud if he has many sex partners."	.18 (.15)	.06 (.11)
Timing of Future Events		
Probability of being in school next year	-.06* (.03)	-0.035* (0.016)
Desired time of marriage	-.52*** (.09)	-.27*** (.05)
Peer Networks		
# of sexual partners of best friend	.625*** (.09)	.20*** (.05)

+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

All coefficients reported are adjusted for gender, age, socioeconomic status, and Yao ethnicity.

Table 6: Test of Criteria 4

	Change in Effect Size of Educational Status				Mediation Test[1]	
Outcome	Age at First Sex		Ever had Sex		Proportion of the total effect of being in school explained by the mediator	
Covariate	Primary	Secondary	Primary	Secondary		
	b (se)	b (se)	b (se)	b (se)		
Baseline Model	**-.35** (.12)**	**-.28** (.09)**	**-.86*** (.22)**	**-.81*** (.20)**	Proportion	(95% CI)
Mediator being tested						
Attitude toward premarital sex	-.41** (.15)	-.35** (.12)	-.91*** (.22)	-.74*** (.20)	.045	(-.02, .11)
Educational Expectations	-.24 (.17)	-.27+ (.14)	-.64** (.24)	-.58** (.22)	.33*	(0.005, .65)
Desired time of marriage	-.17 (.16)	-18 (.13)	-.43+ (.23)	-.38+ (.21)	.61***	(.28, .94)
Friend's # of Sex Partners	-.31* (.15)	-.35** (.13)	-.74** (.22)	-.60** (.20)	.32**	(.124, .516)

---

[1] Method for testing mediation effect for logit analysis described in Erickson et al. (2005), further developed in Buis (2008).

Table 8: Logistic Regression Results for Likelihood of Being Sexually Active with Mediator Effects

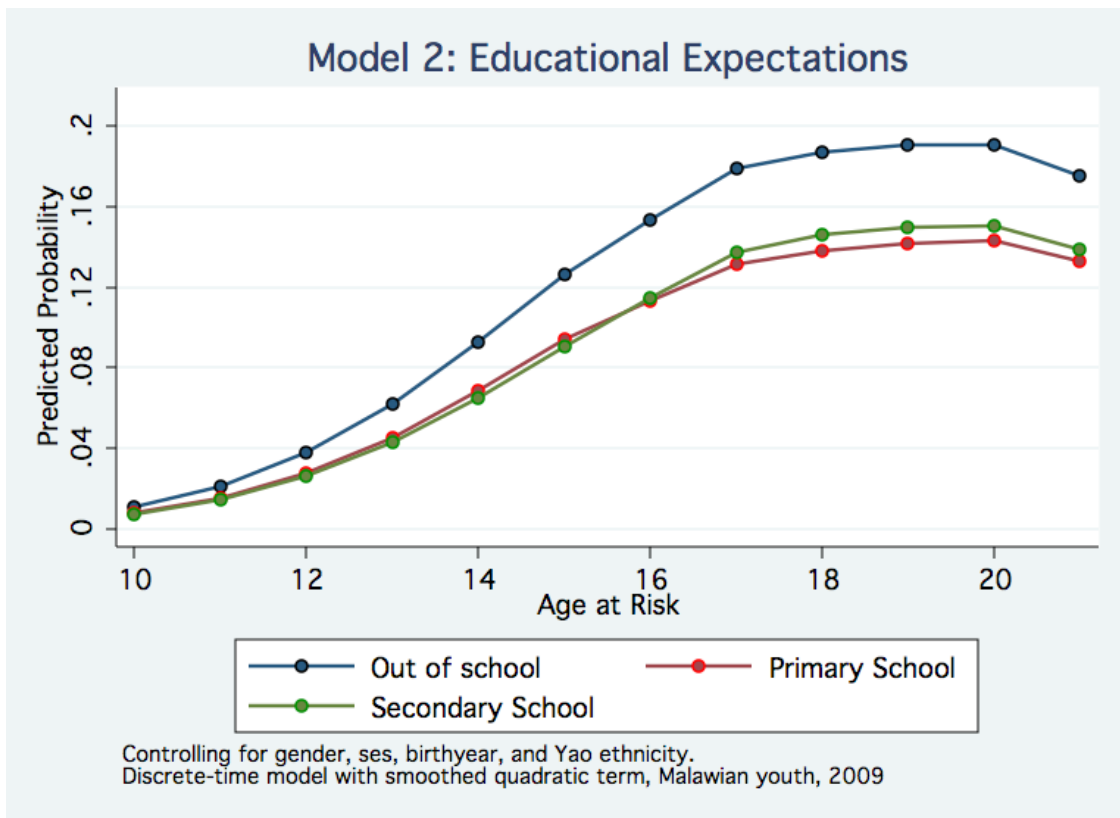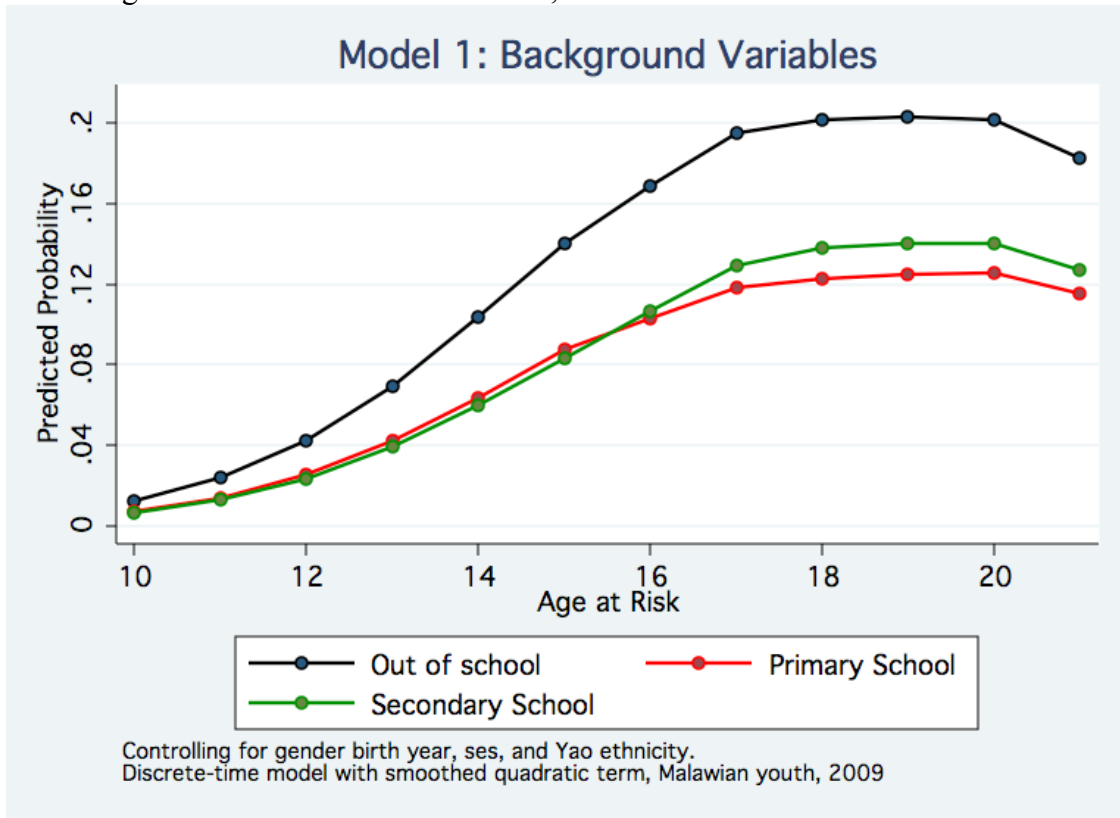|	Baseline		Educ. Exp.		Mar. Asp		Friend's # Partners		All Mediators	
	b/se		b/se		b/se		b/se		b/se	
In primary	-0.89	***	-0.64	**	-0.43	+	-0.74	**	-0.3	
	-0.22		-0.24		-0.23		-0.23		-0.26	
In secondary	-0.83	***	-0.58	**	-0.38	+	-0.6	**	-0.18	
	-0.2		-0.22		-0.21		-0.2		-0.23	
Male	1.13	***	1.15	***	1.32	***	0.99	***	1.15	***
	-0.14		-0.14		-0.14		-0.14		-0.15	
Age	0.18	***	0.18	***	0.13	***	0.18	***	0.13	***
	-0.04		-0.04		-0.04		-0.04		-0.04	
SES	-0.03		-0.02		-0.02		-0.02		-0.01	
	-0.04		-0.04		-0.04		-0.04		-0.04	
Yao ethnicity	0.27	+	0.25	+	0.24		0.23		0.2	
	-0.15		-0.15		-0.15		-0.15		-0.15	
Educ. expectations			-0.06	*					-0.03	
			-0.03						-0.03	
Mar. aspirations					-0.52	***			-0.39	***
					-0.09				-0.09	
Friend's # partners							0.63	***	0.54	***
							-0.09		-0.09	
Constant	-2.7	***	-2.85	***	-2.33	**	-2.73	**	-2.53	**
	-0.82		-0.82		-0.83		-0.83		-0.85	
			-		-		-		-	
Log Likelihood	-680.53		677.36		661.63		650.03		638.79	
N	1155		1155		1155		1155		1155	

+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

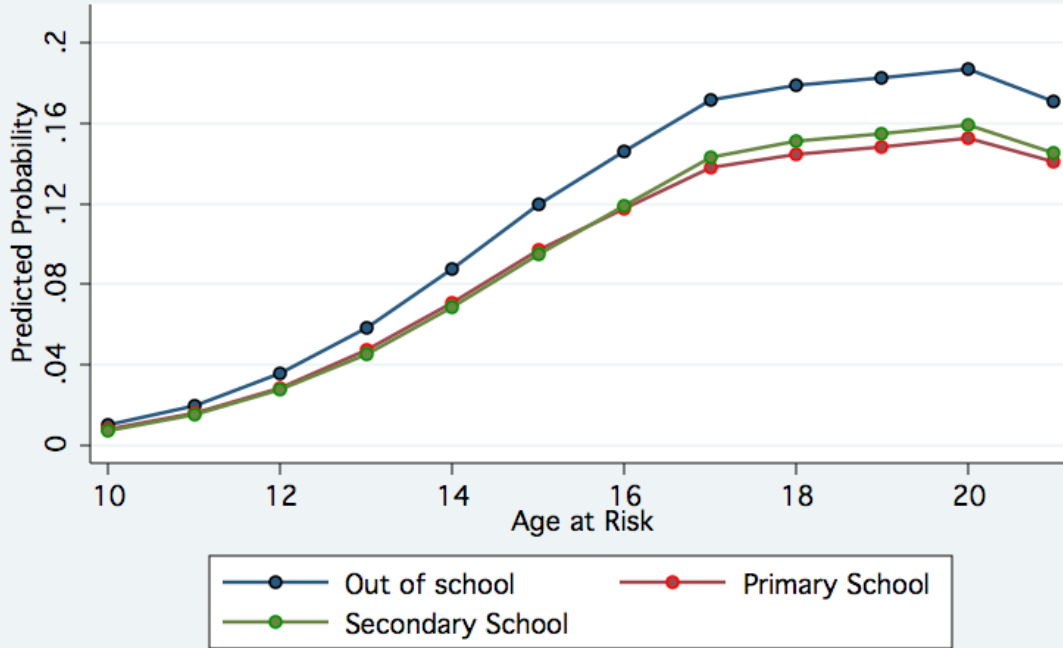Table 9: Results for Discrete-Time Hazard Models for Age of First Sex with Mediator Effects

	Baseline		Educ. Exp.		Mar. Asp		Friend's # Partners		All Mediators	
	b/se		b/se		b/se		b/se		b/se	
In primary	-0.4	**	-0.24		-0.17		-0.31	*	-0.07	
	-0.15		-0.17		-0.16		-0.15		-0.17	
In secondary	-0.41	***	-0.27	+	-0.18		-0.35	**	-0.1	
	-0.12		-0.14		-0.13		-0.13		-0.15	
Male	0.86	***	0.86	***	0.97	***	0.77	***	0.88	***
	-0.1		-0.1		-0.1		-0.1		-0.1	
Age	-0.15	***	-0.15	***	-0.19	***	-0.15	***	-0.18	***
	-0.03		-0.03		-0.03		-0.03		-0.03	
SES	-0.07		-0.06		-0.06		-0.05		-0.04	
	-0.05		-0.05		-0.05		-0.05		-0.05	
Yao ethnicity	0.25	*	0.24	*	0.23	*	0.25	*	0.23	*
	-0.1		-0.1		-0.1		-0.1		-0.1	
Age at risk	1.34	***	1.34	***	1.34	***	1.35	***	1.35	***
	-0.16		-0.16		-0.16		-0.16		-0.16	
Squared age at risk	-0.03	***	-0.03	***	-0.03	***	-0.03	***	-0.03	***
	-0.01		-0.01		-0.01		-0.01		-0.01	
Educ. expectations			-0.04	*					-0.01	
			-0.02						-0.02	
Mar. aspirations					-0.27	***			-0.23	***
					-0.05				-0.06	
Friend's # partners							0.2	***	0.18	***
							-0.05		-0.05	
Constant	-12.43	***	-12.48	***	-12.03	***	-12.52	***	-12.21	***
	-1.34		-1.34		-1.35		-1.36		-1.37	
F (df)	49.17	(8)	43.42	(9)	43.56	(9)	42.32	(9)	34.71	(9)
Respondents	1156		1156		1156		1156		1156	
Risk Years	8221		8221		8221		8221		8221	

+ $p<0.10$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$

Figures 1-5: Predicted Probabilities, Discrete Time Series Hazard Models

Model 1: Background Variables

Controlling for gender birth year, ses, and Yao ethnicity.
Discrete-time model with smoothed quadratic term, Malawian youth, 2009

Model 2: Educational Expectations

Controlling for gender, ses, birthyear, and Yao ethnicity.
Discrete-time model with smoothed quadratic term, Malawian youth, 2009
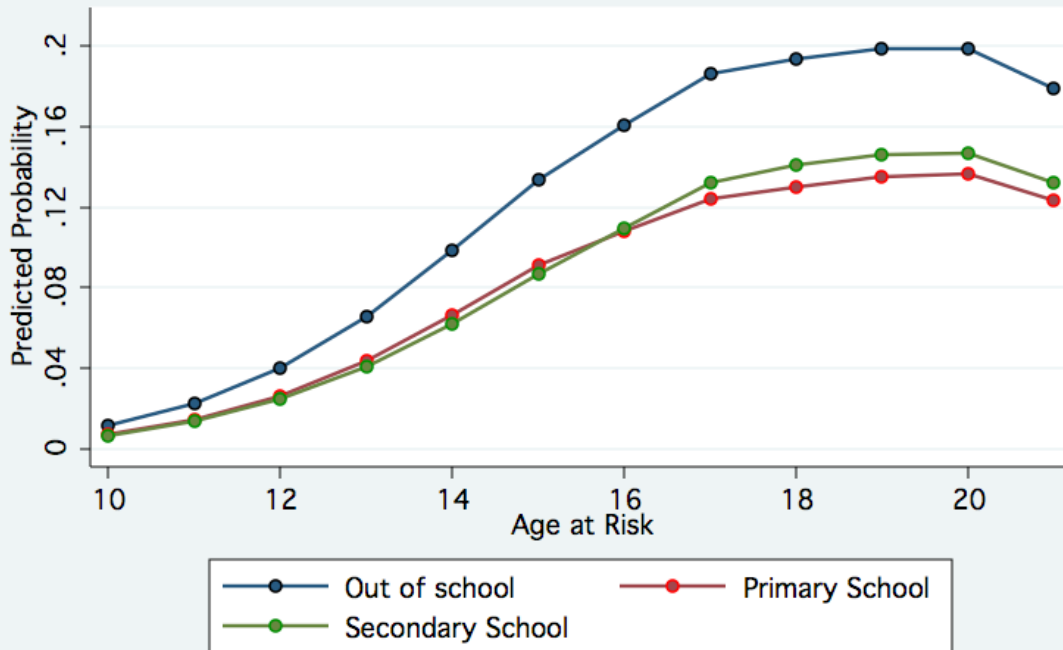
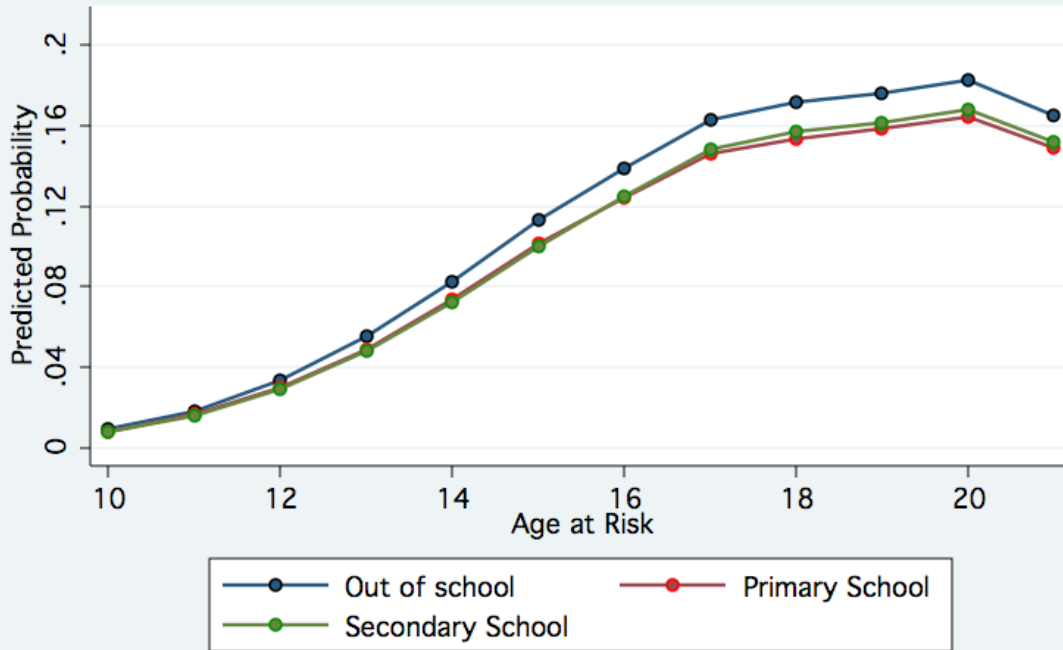# Model 3: Marital Aspirations

Controlling for gender, ses, birthyear, and Yao ethnicity.
Discrete-time model with smoothed quadratic term, Malawian youth, 2009

Model 4: Peer Sexual Behavior

Controlling for gender, ses, birthyear, and Yao ethnicity.
Discrete-time model with smoothed quadratic term, Malawian youth, 2009

Model 5: Combined Mediation Effect

Controlling for gender, ses, birthyear, and Yao ethnicity.
Discrete-time model with smoothed quadratic term, Malawian youth, 2009