High School Dropouts and Sexually Transmitted Infections

D. Mark Anderson
Department of Agricultural Economics and Economics
Montana State University
dwight.anderson@montana.edu
www.dmarkanderson.com

Claus C. Pörtner
Department of Economics
Albers School of Business and Economics
Seattle University
&
Center for Studies of Demography and Ecology
University of Washington
cportner@seattleu.edu
www.portner.dk

August 2012

Abstract

People who drop out of high school fare worse in many aspects of life. We analyze whether there is an effect of dropping out of high school on the probability of contracting a sexually transmitted infection (STI). Previous studies on the relationship between dropout status and sexual outcomes have not empirically addressed self-selection effects. Using individual fixed effects estimations we find strong evidence that dropping out increases the risk of contracting an STI for females. Furthermore, we present evidence that illustrates differences between the romantic partners of dropouts versus enrolled students. These differences suggest that female dropouts may be more susceptible to contracting STIs because they partner with significantly different types of people than non-dropouts. Our results point to a previously undocumented benefit of encouraging those at risk of dropping out to stay in school longer.

Keywords: Education; High school dropouts; Risky behavior; Sexually transmitted infections; Add Health
JEL Classification: I10; I20; J13

I. Introduction

People who drop out of high school do substantially worse compared to those who graduate. Dropouts earn less (Oreopoulos 2006), report lower levels of happiness (Oreopoulos 2007), commit more crimes (Lochner and Moretti 2004; Anderson 2012), and suffer from poorer health (Lleras-Muney 2005; van Kippersluis et al. 2011; Kemptner et al. 2011). Although the research on the consequences of dropping out is substantial, significantly less attention has been paid to the relationship between dropout status and sexual behavior. We fill this gap in the literature by examining the impact of dropping out of high school on the likelihood of contracting a sexually transmitted infection (STI).

From a policy and social perspective, it is crucial to understand the determinants of STIs because of their health and economic consequences (Eng and Butler 1997; Weinstock et al. 2004). Certain STIs can lead to cancer, infertility, or even death. Treating STIs and their complications places a substantial stress on health expenditures; the costs of STIs have been estimated at $17 billion per year (Eng and Butler 1997).[1] A focus on young people is vital because nearly half of STIs contracted occur among individuals who are 15 to 24 years old (Weinstock et al. 2004), and teens and young adults account for nearly 30 percent of new HIV infections annually (Hall et al. 2008).

Although several studies have focused on the relationship between leaving high school early and sexual outcomes, this paper is the first to examine the effect of dropping out on the risk of STI contraction. Prior research suggests that high school dropouts are

[1] This estimate includes the costs due to HIV infections. If these costs are omitted, the annual estimate is approximately $10 billion.

more likely to lose their virginity earlier (Brewster et al. 1998), become pregnant (Manlove 1998), and give birth (Manlove et al. 2000), and are less likely to use contraception (Darroch et al. 1999). Although these studies contribute pertinent information to the discussion on the consequences of dropping out, none have empirically addressed the fact that the dropout decision is endogenously determined.[2] Our research improves upon this literature by controlling for self-selection effects.

In addition, much of the research on the relationship between dropping out and risky behaviors has focused on issues that are relevant primarily for one gender. For example, teenage pregnancies have a much greater impact on females and males are more likely to engage in criminal activities. In fact, there is little research on the relationship between dropping out and risky behavior where the outcome of interest affects both males and females at similar rates. Focusing on STIs allows us to analyze the impact of dropping out for both sexes and to also observe potential short run effects of dropping out. Most adverse health outcomes do not manifest until the person is much older.

To estimate the relationship between high school dropout status and the risk of contracting an STI, we employ data from the National Longitudinal Study of Adolescent Health (Add Health). Longitudinal data allow us to use a fixed effects approach to account for important sources of unobserved heterogeneity. We include school fixed effects to account for the possibility that adolescents who grow up in disadvantaged neighborhoods may be simultaneously more likely to drop out of high school and engage in risky sexual behaviors. We also include individual fixed effects to control for

[2] The lone exception appears to be Black et al. (2008). These authors exploited variation in compulsory schooling laws to estimate the relationship between education and teenage births.

characteristics such as household poverty, individual tastes and preferences, and parenting style that may not only predict an individual's dropout status but may also determine his or her decision to engage in risky sex. As a result, our estimates are less likely to be biased due to sources of unobserved heterogeneity that have plagued previous studies on education and sexual behavior. Another advantage of using the Add Health data is that they contain information on behavior and characteristics of individuals. Previous studies within economics have relied on aggregated STI data, despite the fact that the choice to engage in risky sex is made at the individual level.[3] Lastly, the nature of the data allows us to circumvent issues of reverse causality.

When omitting controls for unobserved neighborhood and individual characteristics, our results suggest that both male and female dropouts have substantially higher STI rates than non-dropouts. Using school and individual fixed effects to account for selection biases, we still find strong evidence that dropping out increases the risk of contracting an STI for females. A complementary analysis suggests that female dropouts may face a higher STI risk because of their post-dropout choice of sexual partners; female dropouts tend to match with significantly older and more abusive partners.

II. Conceptual Framework

How might dropping out of high school influence an adolescent's risky sexual behavior? More specifically, what are the important mechanisms that underlie the relationship between high school dropout status and risky sex?

3 Klick and Stratmann (2003, 2008) and Girma and Paton (2011) estimated the effects of easier abortion access and emergency birth control on STI rates among young people. Chesson et al. (2000), Carpenter (2005), and Cook and Clark (2005) estimated the effects of alcohol policies on STI rates.

First, dropping out may impede accumulation of human capital; thus, lowering expected income and reducing the opportunity cost of risky behaviors. Contraction of an STI may therefore be less costly for individuals with fewer years of schooling.[4] Additionally, youth may learn important values in school that alter their tastes for engaging in risky sex (Arrow 1997).

Second, school may have an incapacitating effect; school attendance leaves less time and opportunity for potentially detrimental activities (Jacob and Lefgren 2003; Luallen 2006; Black et al. 2008; Anderson 2012). Adolescents are also more likely to be monitored in school as opposed to elsewhere.

Lastly, dropping out may change the make-up of one's social circle and thereby affect behavioral outcomes. No longer being around high school students may alter the pool of available sexual partners, and the new pool of partners may be more likely to include older individuals and persons with less desirable characteristics.

Although we are able to provide suggestive evidence as to which underlying factors are important, we cannot pin down the exact mechanism(s) through which dropping out influences sexual behavior. We are, however, able to address to what extent the probability of STI contraction is due to (unobservable) individual characteristics.[5]

---

[4] This argument has been used to explain why HIV rates are higher in African than the United States (Oster 2012).

[5] Because of the problems associated with identifying causal effects, the recent literature has focused on understanding the impacts of policies designed to affect the likelihood of dropping out. Researchers have, for example, studied the effects of compulsory schooling laws (Lochner and Moretti 2004; Lleras-Muney 2005; Oreopoulos 2007; Black et al. 2008; Anderson 2012). Policy experiments, however, do not necessarily reveal the relative importance of individual-level factors that influence future outcomes.

Dropouts may have intrinsic characteristics that simultaneously make them more likely to dropout and engage in risky sex. If this is the case, dropping out serves as a marker for (potentially unobservable) characteristics, such as lower intellectual ability, a higher discount rate, or inferior social skills.

If individual characteristics are truly intrinsic and do not change with more schooling, individuals who are likely to drop out because of those characteristics will do poorly even if forced to finish high school. Addressing the importance of these individual characteristics can provide an indication of how much we might expect policies aimed at reducing dropout rates to affect risky sexual behavior and, more generally, enhance future outcomes (Eckstein and Wolpin 1999).[6] If, instead, human capital, incapacitation, and/or social interaction effects are important, then policies that successfully lower the probability of dropping out will likely have positive social effects.

III. Data

The data used in this paper come from the National Longitudinal Study of Adolescent Health (Add Health). The Add Health is a nationally representative sample of adolescents in the United States who were in grades 7 through 12 during the 1994-95 school year. Data collection started with the identification of over 26,000 schools that included an 11^{th} grade and enrolled more than 30 students. From this sample frame, 80 high schools were selected to ensure representation of schools with respect to region of country, urbanicity, size, type, and ethnicity. Participating high schools were asked to

[6] Eckstein and Wolpin (1999) found that the effect of individual characteristics were so strong that even very restrictive dropout policies would not materially impact graduation rates or other outcomes.

identify feeder schools that included a 7th grade and sent at least five graduates to that high school. Feeder schools chosen to participate in the study were selected with probability proportional to the number of students it contributed to the high school. After including feeders, the total number of participating schools was 132.

In Wave I, data were collected from adolescents, their parents, siblings, friends, relationship partners, fellow students, and school administrators. The Add Health cohort has been followed with three subsequent in-home surveys in 1996, 2000-2001, and 2007-2008. The data includes information on respondents' social, economic, psychological, and health status. In addition to individual-level information, Add Health also contains contextual data on the family, neighborhood, school, and adolescents' peer networks.[7]

We use data from the in-home surveys from Waves I and II because they contain information on self-reported sexual behaviors. By Wave II, a substantial number of students reported having dropped out of school. We therefore use Wave II for the cross-sectional analysis and incorporate Wave I when exploiting the panel nature of the data. Waves III and IV were fielded when respondents were adults and are therefore not used here.

We limit the sample to those aged 15 to 19 at the time of Wave I. The lower limit is set at age 15 because some of the variables used in this analysis were constructed from questions that were only asked to respondents who were at least 15 years old. The upper age limit is 19 because individuals older than this were more likely to have been in high school at Wave I for atypical reasons. We also exclude individuals who reported having been married at any time during Waves I or II and females who were not going to school

[7] For a full description of the Add Health data, see Udry (2003)

because they were pregnant. Lastly, we omit respondents who had missing or inconsistent information on the key outcome variables used in the analysis. The sample sizes are 4554 males and 4470 females, who we observed in both Waves I and II.

Measures of sexual behavior

Our measure of sexual behavior is based on respondent self-reports of having been diagnosed with an STI. The questions used to construct the binary variable for STI status were straightforward. In Wave I, respondents were asked to identify if they had ever been told by a doctor or nurse that they had any of the following STIs: chlamydia, syphilis, gonorrhea, HIV/AIDS, genital herpes, genital warts, trichomoniasis, and hepatitis B. Females were also asked if they had ever been diagnosed with bacterial vaginosis or non-gonococcal vaginitis. In Wave II, the respondents were asked if a positive diagnosis had occurred since the date of the first interview. There is approximately one year between the two waves.

For cross-sectional analyses, only the data from Wave II are used. For individual-level panel analyses, the STI variable captures whether the respondent had been diagnosed with at least one STI before the beginning of the survey for Wave I, and having been diagnosed with an STI between Waves I and II. Although the time period covered by the question is not exactly the same for the two waves, any discrepancy would cause our estimates to be biased downward. Our definition of STI means that it is possible to be counted as infected in period 1 and not in period 2.

Dropout indicator

The explanatory variable of interest is a binary indicator for whether or not the respondent is a high school dropout. First, the adolescent was asked if he/she was currently attending school.[8] If the respondent was not attending school, the interviewer asked a follow-up multiple-choice question as to why the respondent was not in school. An available choice was the individual had "dropped out."

Table 1 illustrates the means of the dependent variables by dropout status. It is immediately apparent that dropouts are quite different than enrolled students by STI status. The problem with simple means is that dropouts are possibly different from non-dropouts along other dimensions than sexual behavior. These other differences, both observed and unobserved to the researcher, could be the driving factors behind the striking differences shown in Table 1. The goal of this paper is to determine the extent to which these disparities can be attributed to a causal effect of dropping out of high school.

[Table 1 about here.]

Covariates

The Add Health data enable us to control for a rich set of covariates that may be associated with dropout status and sexual behavior. These variables are described in Table 2. To retain sample size, we use dummy variables to indicate the respondents who were missing information for specific variables. The explanatory variables are grouped into three categories. The first group pertains to standard individual characteristics that

[8] If the respondent was interviewed in the summer, the question was worded in a retrospective manner. In this case, the respondent was able to identify if he/she had dropped out for the entire previous school year or only part of the year.

describe age, ethnicity, race, and whether the respondent was born in the United States. The second group includes family attributes that are likely to be important for both the dropout decision and the decision to engage in sexual activity. These measures include whether the family moved between survey waves, whether the respondent was an only child, rates of church attendance, parental education and income indicators, whether the adolescent's biological father was living in the household at the time of the Wave I interview, and a dummy variable indicating whether the mother strongly disapproved of her child engaging in sexual intercourse.

[Table 2 about here.]

The third group includes additional individual-level characteristics that are potentially important for dropout status and sexual behavior. Here, we include the respondent's score on the Add Health Picture and Vocabulary Test to serve as a proxy for cognitive ability. Variables for college aspirations and life expectancy are included because they may reflect levels of future orientation. Lastly, we also consider a measure for whether the respondent went with his/her "gut feeling" and did not think about consequences when making decisions. This variable is a proxy for impulsive behavior.

IV. Estimation strategy

Ordinary least squares

To model the effect of dropping out, this paper begins by estimating the following equation:

$$STI_i = \alpha + \boldsymbol{X_i\beta_1} + \beta_2 Dropout_i + c_i + \varepsilon_i, \tag{1}$$

where *i* indexes the individual respondent.

In Equation (1), *STI* refers to whether the respondent has been diagnosed with an STI since the date of last interview.[9] The vector ***X*** contains the personal and family characteristics described in Table 2. The variable *Dropout* is a binary indicator equal to one if the individual has dropped out of high school and zero otherwise. The coefficient of interest, β_2, measures the effect of dropping out on sexual behavior. The variables *c* and ε represent unobserved individual effects and an error term, respectively.

Equation (1) is estimated with ordinary least squares (OLS) for ease of interpretation.[10] Standard errors are clustered at the school level for all estimations. Clustering at the school level is conservative because it takes into account any dependence of errors within schools. All regressions are weighted by the sample weights provided with the Add Health data.

If dropping out of high school was exogenous, then simple OLS regressions would yield consistent estimates of the influence of dropping out on the probability of STI contraction. Clearly, the assumption of exogeneity is unrealistic. Two possible sources of endogeneity exist when estimating Equation (1). First, there may be unobserved factors in the error term that are correlated with dropout status and sexual behavior. For example, an adolescent that discounts the future heavily may choose not to

9 The date of last interview was approximately one year prior to the Wave II interview.

10 Estimating Equation (1) with a probit model to explicitly account for the dichotomous nature of the dependent variable yielded qualitatively similar results to the OLS estimates. Results are available on request.

invest in education while concurrently caring little about the potential consequences of sexual behavior. Alternatively, one can imagine how home or school environments might simultaneously increase the likelihood a youth drops out and engages in sexual activity. Second, structural endogeneity in the form of reverse causality is a concern. That is, students engaging in risky sexual behaviors and incurring the costs associated with their actions may be more likely to leave school than their classmates who abstain from such activities.

We take several approaches to control for endogeneity. First, we include in our models a rich set of covariates designed to capture important personal characteristics and to account for home and family environment. We compare models with and without these variables to gauge the extent to which the relationship between dropping out and STI diagnosis is influenced by contextual and background factors. Second, we employ a series of fixed effects models to eliminate time-invariant unobserved heterogeneity. Lastly, we discuss why reverse causality is not likely to be a concern and consider robustness checks to address this issue.

School fixed effects

To account for school-level characteristics that may be correlated with both dropout status and sexual behavior, we expand Equation (1) to include school fixed effects:

$$STI_{is} = \alpha + \boldsymbol{X_i\beta_1} + \beta_2 Dropout_{is} + c_i + \boldsymbol{\delta_s} + \varepsilon_{is}, \qquad (2)$$

where s indexes the school respondent i attended during Wave I. School fixed effects are represented by $\boldsymbol{\delta}$, a vector of school indicators. In addition to accounting for unobserved time-invariant heterogeneity at the school level, school fixed effects also serve as a proxy for important neighborhood and community characteristics that may bias estimates. The school fixed effects regressions are weighted by the sample weights provided with the Add Health data and standard errors are clustered at the school level.

Individual fixed effects

Though school fixed effects absorb important differences, results may still be biased because of family or individual heterogeneity. Adolescents from poor family environments with parents who place less value on their child's education may be more prone to engage in risky behaviors. Individuals with high rates of time preference, low expectations of positive future outcomes, or lack of motivation for academic pursuits may also be more likely to engage in sexual activities. There may also be systematic differences in the ability to recognize symptoms or be medically diagnosed as having an STI. To account for these sources of bias, we use Add Health data from Waves I and II and employ an individual fixed effects method that amounts to estimating the following first-differenced equation:

$$STI_{i(t+1)} - STI_{it} = \alpha + (\boldsymbol{X_{i(t+1)}} - \boldsymbol{X_{it}})\boldsymbol{\beta_1} + \beta_2(Dropout_{i(t+1)} - Dropout_{it}) + (\varepsilon_{i(t+1)} - \varepsilon_{it}). \quad (3)$$

Equation (3) is the preferred specification because it focuses explicitly on the changes in STI status that occur between Waves I and II, the time period where we observe individuals dropping out. In other words, Equation (3) estimates the causal impact of dropping out by removing individual specific characteristics that are correlated with contracting an STI. As above, the individual fixed effects regressions are weighted by the sample weights provided with the Add Health data and standard errors are clustered at the school level.

This method, however, does not purge our estimates of bias due to unobserved time-variant heterogeneity. Yet, because Waves I and II were only one year apart, it is likely that most unobserved personal characteristics that are correlated with STI and dropout status are constant over this period. Despite this, we examine the robustness of our results to the inclusion of time-varying observables. This allows us to assess how sensitive our estimates are to factors that vary between survey periods. Lastly, an identifying assumption of Equation (3) is that the sexual outcomes we study do not cause individuals to drop out of high school. We address this issue of reverse causality in the robustness checks below.[11]

[11] An instrumental variables strategy could, in principle, be used to control for reverse causality and unobserved time-varying heterogeneity. This approach was attempted; however, given the Add Health data, identifying a valid set of instruments was difficult. Following other research (Campolieti et al. 2010), we used county and local unemployment rates as instruments, but these variables performed poorly in predicting dropout status in the first stage. We also considered county- and local-level industry composition rates, government expenditures on education, labor force participation rates by age, proportions of the population with high school diplomas, and lagged school-level dropout rates. None of these variables predicted adolescent dropout status sufficiently enough to avoid the problems common to weak instruments (Bound et al. 1995; Staiger and Stock 1997).

Fixed effects propensity score matching

As a final estimation method, the robustness of the individual fixed effects results to a propensity score matching technique is examined.[12] After adolescents are matched, fixed effects estimates are obtained. This exercise essentially amounts to re-estimating Equation (3) on a matched sample, a subset of the general sample.[13] Because the propensity score represents the probability of becoming a dropout between Waves I and II, we exclude individuals from observation who reported being dropouts at Wave I.[14] As a result, one might think of our coefficient of interest as akin to the general difference-in-differences estimator. That is, we compare the change in STI status of an individual before and after he/she drops out to the change in STI status of a continuously enrolled individual, who had a similar probability of dropping out, over the same time span.

Propensity score matching involves comparing dropouts and non-dropouts by a single index that captures information from observable pre-dropout characteristics (i.e. Wave I characteristics). The goal is to find a group among our comparison population (i.e. continuously enrolled students) that look as similar as possible to our treatment group (i.e. dropouts). Because there are relatively few treated individuals in the Wave II sample, it is all the more important that care is taken when selecting a control group. Accordingly, we identify the probability of dropping out, i.e. the propensity score, using the following probit model:

---

[12] See Sabia (2006) and Sabia (2007) for applications of this method to topics on sexual behavior.

[13] Observations “off the common support” are excluded from the analysis.

[14] For our sample, this amounted to dropping 44 individuals from observation.

$$P(Dropout_{i(t+1)} = 1) = \Phi(\boldsymbol{X_{it}B}), \tag{4}$$

where $\boldsymbol{X}$ represents a vector of pre-dropout personal characteristics. Letting $P_{i(t+1)}$ denote the predicted probability of dropping out at time $t + 1$ for adolescent i, we match, by nearest propensity score, those whose dropout status did change between Waves I and II to those who remained enrolled for both periods. More formally, for each dropout i, a non-dropout j is selected such that

$$\lambda \;>\; \left|P_{i(t+1)} - P_{j(t+1)}\right| \;=\; \min_{k\in(Dropout=1)} \left\{\left|P_{i(t+1)} - P_{j(t+1)}\right|\right\}, \tag{5}$$

where λ is a pre-specified scalar.[15] This matching technique is referred to as the "within-caliper" method. Matching is done with and without replacement.[16]

V. Results

OLS estimates

Table 3 presents the baseline OLS estimates for males and females from Wave II. Each cell represents a separate regression; we only report the coefficient estimate on *Dropout*, our variable of interest. The full results are shown in Table A1 of the

15 For our analysis, adolescents are matched where the difference between each treated and untreated individual's propensity score is no greater than .10.

16 Matching with replacement is arguably the more flexible approach when working with smaller samples as it allows each treated individual to be paired with his/her closest match, without limiting the set of possible matches (Sabia 2006).

Appendix. To inspect the extent to which background characteristics explain the association between dropout and STI status, we examine increasingly rich specifications.

[Table 3 about here.]

Columns (1) and (2) consider models with only the baseline covariates listed in Table 2. Here, there exists a positive and statistically significant relationship between dropout and STI status. Male and female dropouts are 5.2 and 12.2 percentage points more likely to report having been diagnosed with an STI, respectively.

Columns (3) and (4) add parental and family covariates to the baseline specification. The coefficient estimates remain positive in sign; however, the estimate for males is no longer statistically significant. For females, including measures of family background explain part of the relationship between dropping out and STI contraction; yet, the magnitude of the coefficient estimate remains quite large.

Columns (5) and (6) include additional individual-level covariates likely to be correlated with dropout status. In these specifications we include the Add Health Picture and Vocabulary Test score to proxy cognitive ability, aspirations to attend college, life expectancy, and the degree to which adolescents make decisions based on their "gut feeling." The addition of these covariates decreases the magnitude of the coefficient estimate by slightly over one percentage point for females. The result remains significant at the 5% level and suggests that dropping out is associated with a 9.5 percentage point increase in the likelihood a female will contract an STI. Again, the result for males is positive in sign, but not statistically significant at conventional levels.

School and individual fixed effects

Table 4 presents the fixed effects estimates. Columns (1) and (2) include school fixed effects to control for unobserved heterogeneity at the school and neighborhood levels. Estimates are similar to those presented in Columns (5) and (6) of Table 3.

[Table 4 about here.]

Columns (3) and (4) illustrate results based on estimation of Equation (3) and utilize data from Waves I and II of the Add Health in-home survey.[17] Initially, we exclude time-varying controls from the estimating equations. These fixed effects results are purged of any bias due to unobserved time-invariant heterogeneity at the individual level. Intuitively, we would expect the type of unobserved differences that are important to this relationship to bias our OLS estimates upward.

For females, the estimate is slightly smaller in magnitude than the OLS results at 8.7 percentage points and is statistically significant at the 10 percent level. For males, the result is no longer large in magnitude and remains insignificant. Because contracting an STI requires having had sex, we also considered models where virgins were excluded from the analysis. The results changed little under this alternative specification.

The identifying assumption underlying the individual fixed effects procedure is that $E(\varepsilon_{(t+1)} - \varepsilon_t \,|\, Dropout_{(t+1)} - Dropout_t) = 0$. This assumption is violated if there are time-varying unobservable characteristics that are correlated with dropout status and the likelihood of STI contraction. As mentioned above, because Waves I and II were

[17] Recall that the question in the Wave I interview asked if the respondent had ***ever*** been diagnosed with an STI, while the Wave II question asked if a positive diagnosis had occurred since the date of last interview. Fortunately, any bias due to this slight discrepancy would cause our estimates to understate the truth.

conducted only one year apart, it is likely that most unobserved personal characteristics that are correlated with sexual behavior and dropout status are relatively constant over this period. In Columns (5) and (6), a set of time-varying controls are included to assess the sensitivity of the individual fixed effects models. While the estimate for females is slightly larger in magnitude than the estimate shown in Column (4), the results remain very much the same with or without time-varying controls. We also included time-varying characteristics for peer environment, alcohol use, and drug use in separate regressions. The results for females remained robust to the inclusion of these additional, arguably endogenous, variables. This provides confidence that our estimates are not sensitive to unobserved characteristics that change over time.

Robustness of individual fixed effects

In Table 5 we analyze the robustness of the individual fixed effects results. The baseline estimates from Columns (5) and (6) of Table 4 are shown in the first row of Table 5 for reference.

[Table 5 about here.]

In Row 2 we modify the baseline specification to exclude cases of HIV/AIDS from the definition of having an STI.[18] This is done primarily to address concerns regarding reverse causality. Though evidence suggests that increases in sexual activity can have adverse impacts on academic performance (Sabia 2007; Sabia and Rees 2011), we believe there is little reason to be concerned that being diagnosed with a "less serious"

[18] We also considered models where we excluded anyone having been diagnosed with HIV from the sample entirely. The results changed little under this specification.

STI would actually cause a youth to leave school entirely.[19] Dropping out of high school may be a likely outcome for an individual who has contracted a more serious STI such as HIV. The result for females is slightly smaller in magnitude than the baseline estimate but remains statistically significant at the 10 percent level.

In Row 3 we alter the baseline sample by excluding those individuals who had dropped out for only part of the school year as opposed to the entire year. This restriction also addresses some concerns about reverse causality in that it minimizes the likelihood an individual was diagnosed with an STI before dropping out. For females, the result is slightly larger in magnitude than the baseline estimate and is significant at the 5 percent level.

Next, we combine the restriction in Row 3 with the constraint that individuals who were out of school for reasons other than having "dropped out" are excluded from the sample.[20] These results are presented in Row 4. The purpose of this exercise is to compare dropouts with only individuals that have been in school the entire year. In this case, the coefficient estimate in the female equation remains large at 7.6 percentage points, but loses statistical significance.

In Row 5 we take a different approach and redefine the term "dropout" to include those individuals that were expelled from school. This helps us to better understand if it is the simple act of being out of school that matters for STI contraction. This

[19] Of the individuals in Wave II of Add Health that tested positive for an STI, around 65 percent were diagnosed with an STI that is curable with simple antibiotics. Roughly 20 percent tested positive for herpes or genital warts. While not curable, herpes and warts are readily treatable and can be managed quite effectively.

[20] Students could have been out of school for the following other reasons: suspended/expelled, sick/injured, or on leave.

specification suggests that being out of school leads to a 9.5 percentage point increase in the likelihood a female is diagnosed with an STI.

In Row 6 of Table 5 we exclude individuals who were dropouts at the time of the in-home Wave I interview in addition to excluding respondents who had dropped out for only part of the school year between survey waves. This restriction ensures that individuals are the same with respect to dropout status at baseline. An identifying assumption of the individual fixed effects model is that those who drop out and those who stay continuously enrolled share common unobserved time trends (Sabia 2007). This assumption may be violated if the effects of dropping out are cumulative. The findings in Row 6 suggest that dropping out leads to a 7 percentage point increase in the probability of STI diagnosis for females. Although this point estimate remains relatively large in magnitude, it is not significant at conventional levels.

In the last row of Table 5, we estimate Equation (3) conditional on the respondents having had sex. Because STI contraction requires sexual activity, limiting the sample to those who report having had sex may be more appropriate than considering the sample in its entirety.[21] In making this restriction, the estimate for females remains very similar to the baseline result.

Table 6 investigates the sensitivity of the individual fixed effects results to the propensity score matching procedure outlined above.[22] An important feature of this technique is that it allows us to consider selection into sexual activity in a flexible

---

[21] It is important to note, however, that many of the STIs we study can be transmitted through means other than vaginal sex (e.g. oral sex).

[22] The results from the probit specification used to construct the propensity scores are available from the authors upon request.

manner. When estimating the propensity scores, we included virginity status as an observed characteristic on the right-hand-side of Equation (4). This method is arguably preferred to a basic individual fixed effects model that excludes virgins from observation (i.e. the model estimated in Row 7 of Table 5) because the latter approach does not explicitly model selection into sexual activity.[23]

Columns (1) and (2) show results where matching is done without replacement. Here, the coefficient estimate for females is quite large in magnitude though not statistically significant. Columns (3) and (4) consider results where matching is done with replacement. As mentioned above, matching with replacement is arguably the preferred approach when working with samples where the number of treated persons is relatively small as it allows each treated individual to be paired with his/her closest match, without limiting the set of possible matches. In this case, our results indicate that dropping out of high school increases the risk of STI diagnosis for females by 9.5 percentage points.

[Table 6 about here.]

Sexual activity, partner characteristics, and dropout status

Our results clearly show that female dropouts are more likely to contract STIs, even after controlling for unobservable individual characteristics, whereas there is no statistically significant causal effect for males. The question is what explains these results. We outlined three possible theoretical explanations above. First, the lower

[23] However, it is worth mentioning that results were similar when excluding the indicator for virginity status from Equation (4).

human capital of dropouts makes it less costly to engage in risky sex. Second, no longer being in school means that the adolescent is less likely to be supervised and have less time and opportunity for risky behavior. Third, the make-up of the adolescent's social circle changes and the new pool of potential partners may make it more likely to contract STIs. Here, we provide some suggestive evidence as to which underlying factors are important.

One way to examine whether dropping out makes it less costly to engage in risky sex or lowers the degree of supervision is to estimate how dropping out impacts an adolescent's sexual behavior. Table 7 presents individual fixed effects results of dropping out on whether the respondent used a condom during last intercourse and whether the respondent had sex one or more times during the three months before each survey round. If lower human capital leads to more risky sex, we would expect to see a decrease in condom use as a result of dropping out. While males do have slightly lower rates of condom use, neither the effect for males nor females is statistically distinguishable from zero. If being in school leads to an incapacitation effect or more supervision, we would expect dropouts to be more likely to engage in sex. While male dropouts appear to be more sexually active, neither the effect for males nor females is statistically distinguishable from zero.

[Table 7 about here.]

To further examine what drives the increase in STIs, Table 8 presents results where the following outcome variables are regressed on dropout status: number of reported sexual partners since Wave I, number of sexual partners since Wave I (conditional on having had sex), average age of romantic partner, and whether the respondent reports being physically or verbally abused by their romantic partner.[24] Because of Wave I data limitations, individual fixed effects analyses were not possible for these outcomes. The results are instead from simple OLS regressions that are based on outcomes measured during Wave II and control for basic individual demographics. Of course, these results come with the caveat that the effect of dropping out cannot be purged of bias due to unobserved heterogeneity at the individual level.

Row 1 of Table 8 indicates that female dropouts have just over one sexual partner more than females who stay enrolled. Row 2, however, shows that this result loses statistical significance when conditioning on virginity status; as we saw above, female dropouts are no more likely to have been sexually active over the last three months. For males, it appears there is little difference in the number of sexual partners for dropouts versus enrolled students.

Row 3 highlights that female dropouts' partners are nearly one and a half years older than those of female high school students. Male dropouts' partners are slightly younger than the partners of males who stay in school, although the effect is nowhere near significant.[25] Even more striking is the difference in whether a female reports

---

24 Ideally, we would like to observe how a dropout's peer group changes upon leaving school. While the Add Health allows each respondent to nominate peers, only data is available for peers that are themselves apart of the survey. As a result, it is not possible to observe the characteristics of a peer if that individual is a high school dropout who did not attend a school in the Add Health sample.

25 These results fit the expected effect of dropping out. Lower human capital makes somebody a less attractive partner, but the reduction in lifetime income is especially a problem for males, who are valued more for their earnings abilities than females. In addition, males tend to match with younger females and females tend to match with older males. After dropping out, this pattern is exacerbated with the male dropouts revealing themselves as less desirable matches and female dropouts finding partners in a new and older pool outside school.

having been verbally or physically abused by a romantic partner. The estimate in Row 4 illustrates that female dropouts are over 10 percentage points more likely to report abuse.[26] Although the results in Table 8 are descriptive in nature, they indicate that female dropouts engage in romantic relationships with significantly different types of people. Their partners are not only older on average but are also more likely to be abusive. STI rates are increasing in age for this group and previous research has documented that females in abusive relationships are more worried about contracting STIs (Wingood and DiClemente 1997). This serves as a possible explanation for the disparity in STI status between female dropouts and female high school students.[27]

[Table 8 about here.]

VI. Conclusion

This paper makes three important contributions to the literature. First and foremost, it contributes to our understanding of the impact of dropping out of high school by analyzing an outcome that has not been previously studied and which can help us recognize the relative impacts of causal factors versus individual characteristics. We argue that STIs are superior to other outcomes that can be used to examine the consequences of dropping out: both males and females are sexually active and contract STIs at relatively high rates, STIs are less likely to suffer from reverse causality, and

[26] The sample sizes for the equations estimated in Rows 3 and 4 of Table 7 are greatly reduced because the “Relationship Information” module in the Add Health data set is fraught with missing data. In addition, numerous individuals reported not having been in a romantic relationship.

[27] In a related study, Harawa et al. (2003) find that differences in the characteristics of sexual partners do not explain the racial gap in STIs between young white and black women.

STIs are observed in a longitudinal data set that allows us to control for unobservable characteristics. Second, the paper highlights a previously unexplored pathway through which dropping out can affect future prospects, that of partner choice and availability. Third, we contribute to the economic literature on the determinants of STIs by focusing on the link between education and risky sexual behavior and by using individual-level reports of STI status rather than relying on state-level STI rates.[28]

Both male and female dropouts exhibit substantially higher risks of contracting an STI than those who stay in school if self-selection effects are ignored. Using individual fixed effects to eliminate unobservable individual characteristics that are simultaneously correlated with the risk of dropping out and risky sexual behavior, we find that female dropouts still face a significantly higher risk of contracting a sexually transmitted infection, whereas there appears to be no causal effect of dropping out on STI risk for males. The findings for females are robust to a number of different specifications and the use of different samples. Overall, the results indicate that dropping out of high school increases the likelihood that a female student will contract an STI by around 9 percentage points. This estimate implies that over 70 percent of the difference in the mean rate of STI status between female dropouts and females who stay in school can be explained by dropping out.

What accounts for the difference in effects between males and females? One explanation is that females are more likely to see a doctor regularly for OB-GYN check-ups and infections are therefore more likely to be detected. This explanation, however, is

[28] See, e.g., Klick and Stratmann (2003); Klick and Stratmann (2008); Girma and Paton (2011); Chesson et al. (2000); Carpenter (2005); Cook and Clark (2005).

unlikely to account for the large difference in effects. A more compelling argument is that peer groups change and dropouts enter into significantly different romantic relationships than non-dropouts. We find suggestive evidence in support of this argument. Female dropouts not only match with substantially older males, but they are also more likely to be in physically and verbally abusive relationships. Both of these effects have potentially large detrimental impacts on future life outcomes. Additionally, these differences suggest that romantic partner and peer effects are important mechanisms underlying the relationship between dropping out and STI status. To the extent that sexual behavioral outcomes either directly or through other pathways affect future prospects, our results indicate an important role for public policy in inducing those at risk from dropping out to remain in school longer.

What do these results imply for the scope and effectiveness of policies aimed at reducing dropout rates? Eckstein and Wolpin (1999) argued in their seminal paper on high school dropouts that the effect of individual characteristics is so strong that even very restrictive policies will not materially impact graduation rates or other outcomes. Hence, they argued that policies that do not affect individual traits will have little impact on future labor market outcomes of youths who are kept in school. Contrary to Eckstein and Wolpin (1999) and taken at face value, our results indicate that most of the effect of dropping out for females can be explained by causal factors; for females there is almost no difference between the estimated risk of contracting an STI with and without controlling for unobservable individual characteristics. For males the opposite is the case; when we control for unobservable characteristics there is no difference between dropouts and those who stay in school in terms of STI risk.[29]

Our estimates highlight that policies aimed at reducing dropout rates have important heterogeneous effects. Given our findings, we expect policies that focus on keeping students from dropping out to be substantially more effective for females than for males. To the extent these effects translate to other outcomes, future research on dropout policies should consider the differences between males and females as well as other potentially important sources of heterogeneity. Future research will also benefit from identifying factors that lead to differential effects between genders.

[29] It is clear, however, that there still is a strong individual characteristics effect for male dropouts on the risk of contracting an STI. Furthermore, it is possible that these results mask a causal decrease in human capital for males.

Acknowledgements

We would like to thank Benjamin Hansen, Shelly Lundberg, David Ribar, and Erdal Tekin for helpful comments and suggestions. This research uses data from Add Health, a project directed by Kathleen Mullan Harris and designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill, and funded by grant P01-HD31921 from Eunice Kennedy Shriver National Institute of Child Health and Human Development, with cooperative funding from 23 other federal agencies and foundations. Special acknowledgement is due to Ronald R. Rindfuss and Barbara Entwisle for assistance in the original design. Information on how to obtain the Add Health data files is available on the Add Health website (http://www.cpc.unc.edu/addhealth). No direct support was received from grant P01-HD31921 for this analysis. Partial support for this research came from a Eunice Kennedy Shriver National Institute of Child Health and Human Development research infrastructure grant, 5R24HD042828, to the Center for Studies in Demography and Ecology at the University of Washington.

Appendix

[Table A1 about here.]

References

Anderson, D. Mark. 2012. "In School and Out of Trouble? The Minimum Dropout Age and Juvenile Crime." Working Paper, Montana State University.

Arrow, Kenneth. 1997. "The Benefits of Education and the Formation of Preferences." In Jere Behrman and Nevzer, eds., *The Social Benefits of Education*. Ann Arbor, MI: University of Michigan Press.

Black, S. E., Devereux, P. J., & Salvanes, K. G. 2008. Staying in the Classroom and out of the Maternity Ward? The Effect of Compulsory Schooling Laws on Teenage Births. *Economic Journal* 118: 1025-1054.

Bound, John; Jaeger, David and Baker, Regina. 1995. "Problems with Instrumental Variables Estimation When the Correlation Between the Instruments and the Endogenous Explanatory Variable is Weak." *Journal of the American Statistical Association* 90: 443-450.

Brewster, Karin; Cooksey, Elizabeth; Guilkey, David and Rindfuss, Ronald. 1998. "The Changing Impact of Religion on the Sexual and Contraceptive Behavior of Adolescent Women in the United States." *Journal of Marriage and Family* 60: 493-504.

Carpenter, Christopher. 2005. "Youth Alcohol Use and Risky Sexual Behavior: Evidence from Underage Drunk Driving Laws." *Journal of Health Economics* 24: 613-628.

Campolieti, Michele; Fang, Tony and Gunderson, Morley. 2010. "Labour Market Outcomes and Skill Acquisition of High-School Dropouts." *Journal of Labor Research* 31: 39-52.

Chesson, Harrell; Harrison, Paul and Kassler, William J. 2000. "Sex Under the Influence: The Effect of Alcohol Policy on Sexually Transmitted Disease Rates in the United States." *Journal of Law and Economics* 43: 215-238.

Cook, Robert L. and Clark, Duncan B. 2005. "Is There an Association Between Alcohol Consumption and Sexually Transmitted Diseases? A Systematic Review." *Sexually Transmitted Diseases* 32: 156-164.

Darroch, Jacqueline; Landry, David and Oslak, Selene. 1999. "Age Differences Between Sexual Partners in the United States." *Family Planning Perspectives* 31: 160-167.

Eckstein, Zvi and Wolpin, Kenneth. 1999. "Why Youths Drop Out of High School: The

Impact of Preferences, Opportunities, and Abilities." *Econometrica* 67: 1295-1339.

Eng, Thomas R. and Butler, William T. 1997. "The Hidden Epidemic: Confronting Sexually Transmitted Diseases." Institute of Medicine (U.S.). Committee on Prevention and Control of Sexually Transmitted Diseases. Washington, D.C.: National Academy Press.

Girma, Sourafel and Paton, David. 2011. "The Impact of Emergency Birth Control on Teen Pregnancy and STIs." *Journal of Health Economics* 30: 373-380.

Hall, H. Irene; Song, Ruiguang; Rhodes, Philip; Prejean, Joseph; An, Qian; Lee, Lisa; Karon, John; Brookmeyer, Ron; Kaplan, Edward; McKenna, Matthew and Janssen, Robert. 2008. "Estimation of HIV Incidence in the United States." *Journal of the American Medical Association* 300: 520-529.

Harawa, Nina; Greenland, Sander; Cochran, Susan; Cunningham, William and Visscher, Barbara. "Do Differences in Relationship and Partner Attributes Explain Disparities in Sexually Transmitted Disease Among Young White and Black Women?" *Journal of Adolescent Health* 32: 187-191.

Jacob, Brian A. and Lefgren, Lars. 2003. "Are Idle Hands the Devil's Workshop? Incapacitation, Concentration, and Juvenile Crime." *American Economic Review* 93: 1560-1577.

Kemptner, Daniel; Jurges, Hendrik and Steffen Reinhold. 2011. "Changes in Compulsory Schooling and the Causal Effect of Education on Health: Evidence from Germany." *Journal of Health Economics* 30: 340-354.

van Kippersluis, Hans; O'Donnell, Owen and Eddy Van Doorslaer. 2011. "Long-Run Returns to Education: Does Schooling Lead to an Extended Old Age?" *Journal of Human Resources* 46: 695-721.

Klick, Jonathan and Stratmann, Thomas. 2003. "The Effect of Abortion Legalization on Sexual Behavior: Evidence from Sexually Transmitted Diseases." *Journal of Legal Studies* 32: 407-434.

Klick, Jonathan and Stratmann, Thomas. 2008. "Abortion Access and Risky Sex Among Teens: Parental Involvement Laws and Sexually Transmitted Diseases." *Journal of Law, Economics, and Organization* 24: 2-21.

Lleras-Muney, Adriana. 2005. "The Relationship Between Education and Adult Mortality in the United States." *Review of Economic Studies* 72: 189-221.

Lochner, Lance and Moretti, Enrico. 2004. “The Effect of Education on Crime: Evidence from Prison Inmates, Arrests, and Self-Reports.” *American Economic Review* 94: 155-189.

Luallen, Jeremy. 2006. "School's Out...Forever: A Study of Juvenile Crime, At-risk Youths and Teacher Strikes." *Journal of Urban Economics* 59: 75-103.

Manlove, Jennifer. 1998. "The Influence of High School Dropout and School Disengagement on the Risk of School-Age Pregnancy." *Journal of Research on Adolescence* 8: 187-220.

Manlove, Jennifer; Terry, Elizabeth; Gitelson, Laura; Papillo, Angela and Russel, Stephen. 2000. "Explaining Demographic Trends in Teenage Fertility, 1980-1995." *Family Planning Perspectives* 32: 166-175.

Oreopoulos, Philp. 2006. “Estimating Average and Local Average Treatment Effects of Education when Compulsory Schooling Laws Really Matter.” *American Economic Review* 96: 152-175.

Oreopoulos, Philip. 2007. “Do Dropouts Drop Out Too Soon? Wealth, Health and Happiness from Compulsory Schooling. *Journal of Public Economics* 91: 2213-2229.

Oster, Emily. 2012. HIV and Sexual Behavior Change: Why Not Africa? *Journal of Health Economics* 31: 35-49.

Sabia, Joseph. 2006. “Does Sex Education Affect Adolescent Sexual Behaviors and Health?” *Journal of Policy Analysis and Management* 25: 783-802.

Sabia, Joseph. 2007. “Reading, Writing, and Sex: The Effect of Losing Virginity on Academic Performance.” *Economic Inquiry* 45: 647-670.

Sabia, Joseph and Rees, Daniel. 2011. "Does the Number of Sex Partners Affect Educational Attainment? Evidence from Female Respondents to the Add Health." Forthcoming in *Journal of Population Economics*.

Staiger, Douglas and Stock, James. 1997. “Instrumental Variables Regression with Weak Instruments.” *Econometrica* 65: 557-586.

Udry, J. Richard. 2003. "The National Longitudinal Study of Adolescent Health (Add Health), Waves I & II, 1994-1996; Wave III, 2001-2002 [machine-readable data file and documentation]. Chapel Hill, NC: Carolina Population Center, University of North Carolina at Chapel Hill.

Weinstock, Hillard; Berman, Stuart and Willard, Jr, Cates. 2004. “Sexually Transmitted Diseases Among American Youth: Incidence and Prevalence Estimates, 2000.” *Perspectives on Sexual and Reproductive Health* 36: 6-10.

Wingood, Gina and DiClemente, Ralph. 1997. “The Effects of an Abusive Primary Partner on the Condom Use and Sexual Negotiation Practices of African-American Women.” *American Journal of Public Health* 87: 1016-1018.

Table 1: Means of dependent variable by dropout status

	High school students		High school dropouts	
	Males	Females	Males	Females
Diagnosed with STI	0.021	0.040	0.077	0.166
	(0.003)	(0.005)	(0.026)	(0.045)
	[4373]	[4347]	[181]	[123]

Notes: (1) Means are based on data from Wave II of the National Longitudinal Study of Adolescent Health. (2) Standard errors of the means are reported in parentheses. (3) Sample sizes are in brackets. (4) The STI variable refers to having been diagnosed with an STI since the Wave I interview.

Table 2: Descriptive statistics

Variable	Description	Males Mean	Males Std. Dev.	Females Mean	Females Std. Dev.
Dependent variable					
STI	Equal to 1 if respondent has been diagnosed with an STI, 0 otherwise.	0.023	0.151	0.045	0.208
Baseline respondent characteristics					
Dropout	Equal to 1 if respondent has dropped out of high school, 0 otherwise.	0.047	0.212	0.042	0.200
Age	Respondent's age.	17.55	1.078	17.43	1.005
Hispanic	Equal to 1 if respondent is Hispanic, 0 otherwise.	0.128	0.334	0.115	0.319
Black	Equal to 1 if respondent is black, 0 otherwise.	0.166	0.372	0.173	0.378
White	Equal to 1 if respondent is white, 0 otherwise.	0.730	0.444	0.727	0.445
Other non-white	Equal to 1 if respondent is other non-white race, 0 otherwise.	0.146	0.354	0.138	0.345
U.S. born	Equal to 1 if respondent was born in the United States, 0 otherwise.	0.938	0.241	0.922	0.268
Family characteristics					
Moved	Equal to 1 if family moved between Waves I and II, 0 otherwise.	0.061	0.239	0.062	0.242
Only child	Equal to 1 if respondent is the only child in the family, 0 otherwise.	0.193	0.395	0.196	0.397
Only child missing	Equal to 1 if only child information is missing, 0 otherwise.	0.006	0.075	0.004	0.064
Church 1	Equal to 1 if did not attend church in last year, 0 otherwise.	0.279	0.448	0.226	0.419
Church 2	Equal to 1 if went to church less than once per month, 0 otherwise.	0.188	0.391	0.191	0.393
Church 3	Equal to 1 if went to church at least once per month, 0 otherwise.	0.511	0.500	0.567	0.495
Church missing	Equal to 1 if church information is missing, 0 otherwise.	0.022	0.148	0.015	0.122
Mother's education 1	Equal to 1 if mother has less than high school degree, 0 otherwise.	0.144	0.351	0.175	0.380
Mother's education 2	Equal to 1 if mother has high school degree or GED, 0 otherwise.	0.329	0.470	0.329	0.470
Mother's education 3	Equal to 1 if mother has more schooling than a high school degree, 0 otherwise.	0.408	0.491	0.411	0.492
Mother's education missing	Equal to 1 if mother's education is missing, 0 otherwise.	0.119	0.324	0.085	0.279
Father's education 1	Equal to 1 if father has less than high school degree, 0 otherwise.	0.109	0.311	0.114	0.318
Father's education 2	Equal to 1 if father has high school degree or GED, 0 otherwise.	0.226	0.418	0.224	0.417
Father's education 3	Equal to 1 if father has more schooling than a high school degree, 0 otherwise.	0.352	0.478	0.330	0.470
Father's education missing	Equal to 1 if father's education is missing, 0 otherwise.	0.313	0.464	0.332	0.471
Bio father present	Equal to 1 if biological father was present during Wave I, 0 otherwise.	0.522	0.500	0.512	0.500

Table 2 (continued): Descriptive statistics

Variable	Description	Males Mean	Males Std. Dev.	Females Mean	Females Std. Dev.
Bio father present missing	Equal to 1 if information on biological father's presence is missing, 0 otherwise.	0.120	0.325	0.127	0.333
Parental income 1	Equal to 1 if total household income is less than 40k, 0 otherwise.	0.388	0.487	0.386	0.487
Parental income 2	Equal to 1 if total household income is between 40k and 80k, 0 otherwise.	0.298	0.457	0.290	0.454
Parental income 3	Equal to 1 if total household income is greater than 80k, 0 otherwise.	0.093	0.290	0.100	0.300
Parental income missing	Equal to 1 if total household income information is missing, 0 otherwise.	0.221	0.415	0.225	0.417
Mother's disapproval	Equal to 1 if child thinks mother would strongly disapprove of him/her having sex, 0 otherwise	0.370	0.483	0.540	0.498
Mother's disapproval missing	Equal to 1 if information is missing about mother's disapproval of child having sex, 0 otherwise.	0.072	0.258	0.056	0.229
Additional respondent characteristics					
PVT score	Add Health Picture and Vocabulary Test score.	97.16	25.28	96.33	24.21
PVT score missing	Equal to 1 if Picture and Vocabulary Test score is missing, 0 otherwise.	0.045	0.207	0.039	0.194
College aspirations	Scale of 1 to 5 measure of how much respondent wants to go to college. 1=low aspirations. 5=high aspirations.	4.214	1.441	4.130	1.446
College aspirations missing	Equal to 1 if college aspirations information is missing, 0 otherwise.	0.036	0.186	0.053	0.224
Life expectancy	Scale of 1 to 5 measure reflecting odds respondent thinks he/she will live to be 35. 1=almost no chance. 5=almost certain.	4.214	0.950	4.323	0.891
Life expectancy missing	Equal to 1 if life expectancy information is missing, 0 otherwise.	0.003	0.053	0.004	0.060
Gut feeling	Scale of 1 to 5 measure reflecting if respondent, when making decisions, goes with his/her "gut feeling" without thinking about consequences. 1=Strongly agrees he/she goes with "gut feeling." 5=Strongly disagrees he/she goes with "gut feeling."	2.951	1.155	3.176	1.153
Gut feeling missing	Equal to 1 if "gut feeling" information is missing, 0 otherwise.	0.003	0.055	0.003	0.056

Note: All variables are from Wave II of the National Longitudinal Study of Adolescent Health with the exception of PVT score and the dummies for parental education, parental income, presence of the adolescent's biological father, and whether the respondent is an only child. These variables are from the Wave I survey.

Table 3: Dropping out and sexually transmitted infections (OLS estimates)

	(1) Males	(2) Females	(3) Males	(4) Females	(5) Males	(6) Females
Dropout	0.052**	0.122***	0.038	0.107**	0.036	0.095**
	(0.026)	(0.044)	(0.025)	(0.044)	(0.025)	(0.045)
	[4554]	[4470]	[4554]	[4470]	[4554]	[4470]
1.) Baseline covariates	YES	YES	YES	YES	YES	YES
2.) Family covariates	NO	NO	YES	YES	YES	YES
3.) Additional individual-level covariates	NO	NO	NO	NO	YES	YES

Notes: (1) Each cell represents a separate regression and all regressions are weighted by the sample weights provided by Add Health. (2) Standard errors are in parentheses and are clustered at the school level. (3) Sample sizes are in brackets. (4) Covariates included in the models are described in Table 2. (5) *, significant at 10% level; **, significant at 5% level; ***, significant at 1% level.

Table 4: Dropping out and sexually transmitted infections (Fixed effects results)

	School FE		Individual FE w/o time varying controls		Individual FE w/ time varying controls	
	(1)	(2)	(3)	(4)	(5)	(6)
	Males	Females	Males	Females	Males	Females
Dropout	0.033	0.102**	0.002	0.087*	0.001	0.091*
	(0.024)	(0.046)	(0.036)	(0.048)	(0.036)	(0.048)
	[4554]	[4470]	[9108]	[8940]	[9108]	[8940]

Notes: (1) Each cell represents a separate regression and all regressions are weighted by the sample weights provided by Add Health. (2) Standard errors are in parentheses and are clustered at the school level. (3) Sample sizes are in brackets. (4) The school FE models include the same covariates as the OLS models in the last two columns of Table 3. The individual FE models in Columns 5 and 6 include a set of time-varying controls. The time-varying controls include church attendance, whether the respondent moved, sentiment of mother towards sexual behavior, college aspirations, life expectancy, and an indicator for making decisions based on "gut feeling." (5) *, significant at 10% level; **, significant at 5% level; ***, significant at 1% level.

Table 5: Dropping out and sexually transmitted infections (Robustness of individual FE)

	(1) Males	(2) Females
1.) Baseline estimates	0.001	0.091*
	(0.036)	(0.048)
	[9108]	[8940]
2.) Define "STI" to exclude cases of HIV/AIDS	0.002	0.074*
	(0.036)	(0.043)
	[9108]	[8940]
3.) Exclude those that dropped out for only part of the year	0.004	0.105**
	(0.038)	(0.053)
	[9036]	[8882]
4.) Restriction 3.) and exclude those who were out of school for other reasons	0.017	0.076
	(0.044)	(0.056)
	[8634]	[8596]
5.) Restriction 3.) and define "dropout" as having dropped out or been expelled	0.034	0.095**
	(0.029)	(0.046)
	[9036]	[8882]
6.) Restriction 3.) and exclude those who were dropouts at wave I	0.021	0.070
	(0.040)	(0.057)
	[8970]	[8820]
7.) Conditional on having had sex	-0.010	0.087*
	(0.047)	(0.053)
	[3692]	[3492]

Notes: (1) Each cell represents a separate regression and all regressions are weighted by the sample weights provided by Add Health. (2) Standard errors are in parentheses and are clustered at the school level. (3) Sample sizes are in brackets. (4) All models include a full set of time-varying controls. The time-varying controls include church attendance, whether the respondent moved, sentiment of mother towards sexual behavior, college aspirations, life expectancy, and an indicator for making decisions based on "gut feeling." (5) *, significant at 10% level; **, significant at 5% level; ***, significant at 1% level.

Table 6: Dropping out and sexually transmitted infections (Matching analysis)

	Matching without replacement		Matching with replacement	
	(1)	(2)	(3)	(4)
	Males	Females	Males	Females
Dropout	-0.017	0.120	0.003	0.095*
	(0.048)	(0.085)	(0.037)	(0.049)
	[714]	[476]	[6824]	[7588]

Notes: (1) Each column represents a separate regression and all regressions are weighted by the sample weights provided by Add Health. (2) Standard errors are in parentheses and are clustered at the school level. (3) Sample sizes are in brackets. (4) All models include a full set of time-varying controls. The time-varying controls include church attendance, whether the respondent moved, sentiment of mother towards sexual behavior, college aspirations, life expectancy, and an indicator for making decisions based on “gut feeling.” (5) *, significant at 10% level; **, significant at 5% level; ***, significant at 1% level.

Table 7: Differences between dropouts and non-dropouts (Individual FE results)

	(1) Males	(2) Females
Condom use during last intercourse	-0.019 (0.071) [2860]	0.028 (0.069) [2956]
Sex during last 3 months	0.075 (0.054) [8406]	-0.003 (0.066) [8608]

Notes: (1) Each column represents a separate regression and all regressions are weighted by the sample weights provided by Add Health. (2) Standard errors are in parentheses and are clustered at the school level. (3) Sample sizes are in brackets. (4) All models include a full set of time-varying controls. The time-varying controls include church attendance, whether the respondent moved, sentiment of mother towards sexual behavior, college aspirations, life expectancy, and an indicator for making decisions based on "gut feeling." (5) *, significant at 10% level; **, significant at 5% level; ***, significant at 1% level.

Table 8: Differences between dropouts and non-dropouts (OLS results)

	(1) Males	(2) Females
Number of sexual partners since Wave I interview	0.223 (0.515) [4557]	1.081*** (0.348) [4574]
Number of sexual partners since Wave I interview (conditional on having had sex)	0.316 (0.508) [1814]	0.681 (0.457) [1786]
Average age of romantic partner	-0.085 (0.153) [3035]	1.472*** (0.316) [3357]
Physically or verbally abused by romantic partner	0.066 (0.053) [3062]	0.102** (0.044) [3415]

Notes: (1) Results are based on data from Wave II of the National Longitudinal Study of Adolescent Health. (2) Each cell represents a separate regression and all regressions are weighted by the sample weights provided by Add Health. (3) Standard errors are in parentheses and are clustered at the school level. (4) Sample sizes are in brackets. (5) All models control for age, race, and whether the respondent was born in the United States. (6) *, significant at 10% level; **, significant at 5% level; ***, significant at 1% level.

Table A1: Dropping out and sexually transmitted infections (Full OLS estimates)

	(1) Males	(2) Females	(3) Males	(4) Females	(5) Males	(6) Females
Dropout	0.052**	0.122***	0.038	0.107**	0.036	0.095**
	(0.026)	(0.044)	(0.025)	(0.044)	(0.025)	(0.045)
Age 16	-0.001	-0.013	-0.005	-0.012	-0.004	-0.014
	(0.011)	(0.019)	(0.012)	(0.019)	(0.012)	(0.019)
Age 17	0.006	-0.002	0.003	-0.003	0.005	-0.005
	(0.012)	(0.020)	(0.013)	(0.021)	(0.013)	(0.021)
Age 18	-0.006	0.016	-0.007	0.014	-0.004	0.015
	(0.011)	(0.019)	(0.012)	(0.020)	(0.013)	(0.020)
Age 19	0.028	0.024	0.023	0.017	0.028	0.023
	(0.020)	(0.025)	(0.019)	(0.026)	(0.019)	(0.027)
Hispanic	0.013	-0.006	0.003	-0.010	0.001	-0.015
	(0011)	(0.016)	(0.012)	(0.015)	(0.013)	(0.016)
Black	0.036***	0.049***	0.032**	0.046***	0.024*	0.041***
	(0.013)	(0.013)	(0.013)	(0.013)	(0.012)	(0.013)
Other non-white	0.004	0.019	0.005	0.021	0.004	0.023
	(0.009)	(0.018)	(0.008)	(0.018)	(0.008)	(0.017)
U.S. born	0.004	0.026**	0.002	0.026*	0.008	0.026*
	(0.014)	(0.013)	(0.014)	(0.013)	(0.013)	(0.014)
Moved	---	---	-0.010	-0.003	-0.009	-0.008
			(0.011)	(0.021)	(0.011)	(0.021)
Only child	---	---	0.009	0.026**	0.012	0.027**
			(0.008)	(0.011)	(0.008)	(0.011)
Only child missing			-0.029***	-0.005	-0.027***	-0.009
			(0.010)	(0.017)	(0.010)	(0.018)
Church 2	---	---	-0.003	0.018	-0.002	0.020*
			(0.012)	(0.012)	(0.012)	(0.012)

Table A1 (continued): Dropping out and sexually transmitted infections (Full OLS estimates)

Church 3	---	---	-0.013	-0.002	-0.012	0.000
			(0.010)	(0.009)	(0.010)	(0.009)
Church missing			-0.021	-0.063***	-0.020	-0.055***
			(0.020)	(0.017)	(0.020)	(0.015)
Mother's educ. 2	---	---	-0.030*	-0.011	-0.027*	-0.009
			(0.017)	(0.015)	(0.016)	(0.015)
Mother's educ. 3	---	---	-0.046***	-0.013	-0.040**	-0.007
			(0.017)	(0.014)	(0.016)	(0.014)
Mother's educ. missing			-0.053	0.022	-0.051**	0.022
			(0.021)	(0.027)	(0.020)	(0.027)
Father's educ. 2	---	---	0.008	0.002	0.010	0.003
			(0.012)	(0.017)	(0.012)	(0.017)
Father's educ. 3	---	---	0.009	-0.008	0.013	-0.004
			(0.014)	(0.015)	(0.014)	(0.016)
Father's educ. missing			0.018	0.004	0.018	0.007
			(0.017)	(0.016)	(0.017)	(0.016)
Bio father present	---	---	-0.003	0.012	-0.002	0.013
			(0.008)	(0.014)	(0.008)	(0.014)
Bio father present missing			-0.030*	0.004	-0.025*	0.004
			(0.016)	(0.020)	(0.015)	(0.020)
Parental income 2	---	---	-0.003	-0.001	-0.002	0.000
			(0.006)	(0.010)	(0.006)	(0.010)
Parental income 3	---	---	0.008	-0.008	0.011	-0.004
			(0.009)	(0.014)	(0.008)	(0.014)
Parental income missing			0.021	0.000	0.018	-0.001
			(0.016)	(0.015)	(0.015)	(0.015)
Mother's disapproval	---	---	0.006	-0.023**	0.007	-0.020**
			(0.008)	(0.010)	(0.008)	(0.010)

Table A1 (continued): Dropping out and sexually transmitted infections (Full OLS estimates)

Mother's disapproval missing			0.040**	0.006	0.038**	0.012
			(0.019)	(0.036)	(0.019)	(0.036)
PVT score/10	---	---	---	---	-0.005**	0.000
					(0.002)	(0.004)
PVT score missing					-0.064**	-0.032
					(0.028)	(0.039)
College aspirations	---	---	---	---	0.001	-0.011
					(0.003)	(0.007)
College aspirations missing					-0.013	-0.074**
					(0.017)	(0.035)
Life expectancy	---	---	---	---	-0.008**	-0.010
					(0.004)	(0.006)
Life expectancy missing					-0.054**	-0.084***
					(0.024)	(0.026)
Gut feeling	---	---	---	---	-0.008***	-0.004
					(0.003)	(0.004)
Gut feeling missing					-0.057***	-0.058*
					(0.014)	(0.029)
N	4554	4470	4554	4470	4554	4470

Notes: (1) Each column represents a separate regression and all regressions are weighted by the sample weights provided by Add Health. (2) Standard errors are in parentheses and are clustered at the school level. (3) Dummy variables for missing observations are included, but not reported. (4) *, significant at 10% level; **, significant at 5% level; ***, significant at 1% level.