Schooling and sexual behavior in South Africa: The role of peer effects

David Lam (corresponding author)
University of Michigan
davidl@umich.edu
734-763-3009

Letícia Marteleto
University of Texas
marteleto@prc.utexas.edu

Vimal Ranchhod
University of Cape Town
vimal.ranchhod@gmail.com

Extended abstract prepared for the submission to the
Sixth African Population Conference
Ouagadougou, Burkina Faso
5-9 December 2011

Date of draft: 24 June 2011

Support for this research was provided by the U.S. National Institute of Child Health and Human Development (Grants R01HD39788 and R01HD045581) and the Fogarty International Center of the U.S. National Institutes of Health (D43TW000657). Excellent research assistance was provided by Kendra Goostrey.

ABSTRACT

This study examines the influence of exposure to older peers on sexual debut in urban South Africa. The study analyzes data from the Cape Area Panel Study (CAPS), a longitudinal survey of young adults in metropolitan Cape Town. The combination of early sexual debut, high rates of school enrollment into the late teens, and high rates of grade repetition create an environment in which young people who progress through school ahead of their cohorts interact with classmates who may be several years older. We construct a measure of cumulative exposure to classmates at least two years older, and show that this measure has a statistically significant positive effect on sexual debut of adolescent girls. It also increases the age difference of the first sexual partner for those girls, and helps explain a significant fraction of the earlier sexual debut of African girls compared to coloured and white girls in Cape Town.

INTRODUCTION

This paper looks at the relationship between schooling and sexual debut in South Africa. The paper builds on two important features of adolescent lives in urban South Africa that have been well documented in previous research. First, school enrollment rates are high through at least age 18. Significant proportions of African (black) South Africans continue to be enrolled in secondary school beyond age 20, a result of high rates of grade repetition and high payoffs to completing grade 12 (Anderson et al. 2001, Lam et al. 2010). The second important feature is that most African adolescents become sexually active by age 18. As shown by Dinkelman et al. (2007), 72% of 17-18 year-old African females in Cape Town reported having had sex in 2005. The combination of these two patterns means that most young people become sexually active while they are still in school.

Research from a number of other African countries has argued that school enrollment has the "protective effect" of delaying sexual debut (Lloyd 2005, Lloyd 2007, Darabi et al. 2008). These studies argue that schools have the capacity to enhance success in all transitions to adulthood, mainly through the acquisition of knowledge and skills. They also note, however, that schools can be a place of conflict and socialization to undesirable behaviors. Not only teachers and principals, but also peers have an important impact on young people's schooling experience and on how those will relate to subsequent transitions to adulthood. Belonging to a positive peer group is likely to lead to a positive school effect on adolescents' outcomes. At the same time, being exposed to an older and therefore more sexually active peer group might influence adolescents to become sexually active themselves.

The "protective effect" of schools may be more complicated when most adolescents are becoming sexually active at ages before they leave school. One of the intriguing results found in

previous analyses of sexual debut in South Africa (using the same data used in this paper) was a positive effect of baseline grade attainment on subsequent sexual debut, controlling for baseline age (Dinkelman et al. 2007, 2008; Marteleto et al. 2008). Estimating probit regressions of sexual debut on a number of individual and household characteristics, Marteleto et al. (2008) found that the number of grades completed by 2002 had a significant positive effect on sexual debut between 2002 and 2005 for both males and females who were age 14-16 in 2002. These positive effects of schooling on sexual debut, controlling for age, are surprising, since we might expect that young people who are ahead of their age group in school would be less likely to become sexually active. The estimates imply that a girl age 14-16 with one additional grade completed in 2002, given her age, was 6.6 percentage points more likely to sexually debut by 2005. The effect for boys was slightly larger at 8 percentage points.

One possible interpretation of this result is that young people who are ahead of their cohort in school interact with an older and more sexually active group of girls and boys. High rates of grade repetition in South Africa mean that students in any given grade in secondary school span a wide range of ages, especially in African schools. As we will see below, African girls who were age 16 in 2002 were distributed from Grade 6 to Grade 12, implying large differences in the age distribution of their classmates. Adolescents could be influenced by the behavior of older same-sex peers and by interactions with older opposite-sex peers.

The goal of this paper is to explore these possible peer effects in greater detail. Our analysis is similar in spirit to recent research in the United States that tries to identify "contagion" effects of interacting with older peers. A recent paper by Argys and Rees (2008), for example, uses variation in the mandated age at which children begin school across U.S. states as an exogenous source of variation in exposure to older classmates during the teenage years. Other papers, such

as Eisenberg (2004) have used the variation in whether ninth graders are grouped with grade 10-12 students or with grade 7-8 students.

Significant research has gone into the peer effects on behavior and life decisions of students at various levels in the United States. Exploiting the random assignment of college roommates, Sacerdote (2001) found evidence for peer effects among roommates on academic outcomes and decisions to join social groups. Analyzing 10th grade classmates in the National Education Longitudinal Survey (NELS), Gaviria and Raphael (2001) found evidence of peer effects at the school level on substance use, church attendance and dropping out of high school. Using data from the National Longitudinal Survey of Adolescent Health, Sieving et al. (2006) found that there was a significant positive relationship between the proportion of a youth's friends who were sexually active in Wave 1 and the probability of sexual debut between Waves 1 and 2.

Our analysis uses the fact that there is a high variance in age-for-grade distributions in South Africa, especially in predominantly non-white schools. Using retrospective schooling histories we estimate the exposure of respondents to older classmates beginning at age 12. The results suggest that the apparent positive effect of grade completion on sexual debut is in fact due to the increased exposure to older classmates that results from being further ahead in school. We estimate a statistically significant positive impact of our exposure measure (explained below) on the probability of sexual debut for females, with inclusion of the exposure variable causing the apparent impact of grade completion to become much smaller and statistically insignificant. As further evidence on the impact of older classmates, we show that high school girls who pass their grade in 2002 are more likely to become sexually active than girls who fail their grade.

We analyze three outcomes in addition to sexual debut. We show that our peer exposure variable has a positive impact on the age of the first sexual partner for females, consistent with

the view that they are affected by interactions with older classmates. We also look at the impact of our exposure variable on smoking and drinking. While we estimate positive point estimates for the impact of exposure to older classmates on smoking and drinking, the estimates are small in magnitude and are statistically insignificant. While this might appear to contradict our hypothesis of contagion effects of older peers, we show that the age gradient for smoking and drinking is much less steep than the age gradient for sexual debut. Smoking and drinking rates are particularly low for Africans, reducing the opportunity for contagion effects to operate.

DATA: THE CAPE AREA PANEL STUDY

We use data from Waves 1-4 of the Cape Area Panel Study (CAPS), a longitudinal survey of young people in metropolitan Cape Town. Details about the design of CAPS are provided in Lam et al. (2008).[1] Wave 1, which was conducted in 2002, included a household questionnaire along with a young adult questionnaire administered to up to three young adults aged 14-22. The young adult questionnaire collected data on a wide range of topics, including sexual behavior, schooling, and employment. The young adult questionnaire also included a literacy and numeracy evaluation and a life history calendar that provides retrospective information on living arrangements, schooling, and pregnancy.

An important feature of Cape Town is that there continue to be large socieconomic disparities and a high degree of spatial segregation across the three major population groups – African, coloured, and white (the Cape Town population in the 2001 census was 48% coloured, 32% African, and 19% white). These three population groups were subject to very different treatment under apartheid. Whites had advantages in most areas, including significantly higher expenditures on schooling, privileged access to the labor market, unrestricted residential

[1] The Cape Area Panel Study is a collaborative project of the University of Michigan and the University of Cape Town, funded by the U.S. National Institutes of Health. Additional details and technical documentation are available at www.caps.uct.ac.za.

mobility, and better access to social services. Africans had the least access to services and the most restrictions on work and migration, with a large gap in expenditures on schooling. The coloured population, which is heavily concentrated in the Western Cape (including Cape Town), occupied an intermediate status under apartheid, with higher expenditures on schooling, fewer restrictions on residential mobility, and better access to jobs than Africans.[2]

CAPS was designed using a two-stage probability sample of households, with an oversampling of African and white households in order to get large enough samples to make meaningful comparisons across groups. The baseline wave of CAPS surveyed 4,751 young adults living in 3,304 households. As in most South African household surveys, response rates were high in African and coloured areas and low in white areas. Household response rates were 89% in African areas, 83% in coloured areas, and 46% in white areas.[3] Young adult response rates, conditional on participation of the household, were high, even in white areas. Given household participation, response rates for young adults were 93% in African areas, 88% in coloured areas, and 86% in white areas (Lam et al. 2008).

Table 1 shows the sample size by population group and provides information on attrition between waves. We show information for the full sample aged 14-22 in 2002 and for the subset that was aged 14-17 in 2002, the sample we use for our regressions. As seen in Table 1, the original Wave 1 sample included roughly equal numbers of African and coloured respondents, as

[2] The coloured population, which was about 9% of the country in the 2001 census, includes descendants from indigenous Khoisan people and Dutch slaves from Malaysia and other areas. It is predominantly Afrikaans speaking, with about 20% identifying as Muslim in CAPS. Under apartheid the coloured population, concentrated in Cape Town, was distinct from the Indian population, concentrated in Durban, as well as from the roughly 80% of the population classified as African/black.

[3] As discussed in Lam et al. (2008), household response rates were lower in high-income areas. Sample weights adjust for differential response rates within sample clusters, which partially accounts for differential response rates that are correlated with sample cluster characteristics such as income. In practice results are very little affected by sample weights when race dummies are included in regressions.

intended in the sample design. The weighted percent column shows that when sample weights are used to adjust for the sample design and differential response rates, the weighted sample is 28% African, 53% coloured, and 19% white, proportions that are similar to those found for the same age group in Cape Town in the 2001 South African census (Lam et al. 2008).

Sample attrition

CAPS Wave 3, collected in 2005, provides most of the longitudinal information used in this paper. As seen in Table 1, 3,531 of the 4,751 original respondents were successfully interviewed in Wave 3. In Wave 4, collected in 2006, we located almost 400 additional respondents that were missed in 2005. Since we collect retrospective data on variables such as schooling and sexual activity since the last interview, we can use Wave 4 data to fill in information on 2005 outcomes for respondents who were in Wave 4 but not Wave 3. The effective sample for 2005 outcomes, then, is 3,916, implying a 17.6% overall attrition rate between 2002 and 2005.

As seen in Table 1, attrition differs significantly by race. The African attrition rate is 20%, with proxy reports indicating that most attrition is due to migration back to the rural Eastern Cape province, the main sending region for Africans living in Cape Town. The coloured population has strong roots in Cape Town, a factor in its lower 10% attrition rate. The 34% attrition rate for whites includes both migration out of Cape Town (including out of South Africa) and a significant number of refusals. The bottom panel of Table 1 shows the sample size and attrition rates for the sample that was aged 14-17 in 2002, the group we use in our regressions. Attrition for this group is considerably lower than for the full sample, 12% overall, a reflection of the positive relationship between age and attrition.

We have analyzed attrition using probit regressions of 2002-2005 attrition on baseline characteristics (not shown). In addition to being correlated with age and race, attrition is

negatively correlated with baseline grade attainment and school enrollment. It is not significantly correlated with whether the respondent was sexually active at baseline, once we control for variables we will include in our regressions such as age, race, household income, and parental education. As a robustness test for the regressions reported below we have estimated the regressions using inverse probability weights following the approach of Fitzgerald et al. (1998). The results are virtually identical in the weighted regressions, giving us confidence that sample attrition does not have a major impact on our results.[4]

DISTRIBUTION OF KEY VARIABLES

In order to see contagion effects from exposure to older classmates we need to see two patterns in the data. First, there needs to be a fairly steep age gradient in the behavior for which we expect there to be contagion effects. Students need to be exposed to significantly different behavior when they interact with, say, 17-year-old classmates than when they interact with 15-year-old classmates. Second, we need to have variation in the degree to which students are exposed to older peers, with some students experiencing significant exposure to older peers. In this section we present evidence about the age gradient in the outcomes we are studying and about the age-for-grade distributions in African, white, and coloured schools. We also define the measure of exposure to older peers that will be used in our regression analysis below.

[4] For the approach of Fitzgerald et al. (1998) we assume that some variables (household size, quadratic in head's age, whether the head was born in Cape Town, and whether the respondent had a high expectation of living in Cape Town in three years) are correlated with attrition but not with unobserved determinants of sexual debut. While these variables are significant predictors of attrition, we are not confident assuming that they (or any other variables in the data) are uncorrelated with unobserved determinants of sexual debut. We therefore do not include the adjustments for attrition in the results presented below. While we cannot rule out that attrition is affecting our results, we think it is unlikely that attrition is playing an important role given that attrition in the relevant age group is only 12%, attrition is uncorrelated with sexual activity at baseline, and our results are unaffected by adjustments for attrition.

Variables and potential biases

Our key outcome variables are self reports of sexual debut, age of first sexual partner, and whether the respondent had smoked cigarettes or drunk alcohol in the previous month. For sexual debut we use two questions, the first on whether the respondent ever had sex and the second on the age at which the respondent first had sex. For most respondents we simply look at whether they report having had sex at Wave 3 (2005), conditional on not having had sex at Wave 1 (2002). For respondents captured in Wave 4 (2006) but not in Wave 3 we use their report of age at first sex to determine whether they had sex by 2005. We also use information on the age of the first sexual partner, a question that appears in Wave 3 and Wave 4.

The CAPS sexual behavior questions are similar to surveys such as the Demographic Health Surveys. As shown by studies such as Nnko et al. (2004) and Gersovitz (2005), sexual behavior may not be reported accurately, even when interviews are done in a private setting by well-trained interviewers. While there are no doubt inaccuracies in these reports, we think it unlikely that there are systematic biases driving our results. The biggest concern would be that exposure to older classmates leads individuals to report having had sex when they had not. This could be a peer effect in its own right, with respondents wanting to seem more like their older classmates. While we cannot rule this out, we note that there is a great deal of variation in reported sexual activity by age. It is certainly not the case that 15 year-olds all report being sexually active when they have a high proportion of sexually active classmates. As shown below, female respondents exposed to older classmates report both earlier sexual debut and older partners at first sex. While peer effects might cause respondents to falsely claim that they are sexually active or understate their age of sexual debut, it seems unlikely that peer effects would cause them to inflate the age of their first sexual partner. While we cannot verify the self reports of sexual behavior, we think it unlikely that inaccurate reports are responsible for the empirical patterns we document below.

Age profiles for key outcomes

We begin by documenting the age profiles in three outcomes – sexual debut, cigarette smoking, and drinking alcohol. In order for there to be contagion effects from interacting with older classmates, one necessary condition is that there is a steep age gradient in the behavior being considered. Figure 1 shows age profiles by race and gender for these three outcomes from age 14 to 22 as reported in CAPS Wave 1. As shown in the top panel, the age gradient for sexual debut is very steep. The proportion of African girls who reported having had sex rose from less than 5% at age 14 to 32% at age 16 to 68% at age 18.[5] A 16-year-old whose classmates are age 14-16 would have a very different exposure to sexually active classmates than a 16-year-old whose classmates are age 16-18. As seen in Figure 1, coloured and white teenagers start sexual activity somewhat later than do Africans, although the age gradient is still very steep.

A useful summary measure of the slope of the age gradients in Figure 1 is a simple OLS regression of the binary outcome on age. Table 2 shows this slope coefficient for each of the 18 sex*race combinations in Figure 1. We use the sample aged 14-20 since that is the range most relevant for our analysis of peer exposure during secondary school. As seen in Table 2, the probability of sexual debut rises by 12 and 14 percentage points per year of age for African males and females, respectively, and about 10 percentage points per year for the other groups.

The second panel of Figure 1 shows the proportion who said they smoked a cigarette in the month prior to the Wave 1 survey. Smoking rates for African females are extremely low, under 4% at all ages from 14-22. Smoking rates for African males are higher, increasing from 4% at age 14 to 20% at age 18. Coloured males and females have the highest smoking rates of the

[5] African males appear to have somewhat earlier sexual debut than African females, although more females report being sexually active at age 15 and 17. There are about 100 observations in each age-sex cell for African and coloured results in Figure 1, about 40 per cell for whites. Given these cell sizes, many of the gender differences in single-year age groups are not statistically significant, and we focus more on the overall age gradient for each sex than on the sex differences at a given age.

three population groups, with the rate rising from 23% to 54% between age 14 and age 18 for coloured males. These racial patterns in teenage smoking are consistent with previous work in South Africa that has documented the lower rates of smoking among African adolescents compared to coloured and white adolescents (Swart et al. 2001).[6] As shown in the second panel of Table 2, the age gradient in smoking is considerably lower than the age gradient for sexual debut. The smoking rate rises 5-7 percentage points per year of age for males, and at lower rates for females. For African females the smoking rate rises at only 0.2 percentage points per year, a rate that is statistically not significantly different than zero.

The third panel of Figure 1 shows drinking rates by age and race. Self-reported drinking is very uncommon among African males and females, with the rate only reaching 40% among 22-year-old African males. Drinking rates are somewhat higher among coloured youth, and are highest among whites, where 70% of males and females report drinking at age 19 and up. Looking at the regression coefficients in the bottom panel of Table 2, the age gradient in drinking is much lower than the age gradient in sexual debut for African males, African females, and coloured females. The age gradient in drinking for whites is similar to the age gradient in sexual debut, about 10 percentage points per year of age. While this could generate peer effects of the kind we are interested in, whites have very little dispersion in age for grade, and therefore have little opportunity for contagion effects from interacting with older classmates.

Distribution of age by grade

The second important component of our argument is that some students are exposed to classmates spanning a wide age range. In this section we present evidence on the age-for-grade and grade-for-age distributions for students in the age range where contagion effects may be

[6] Although current legislation bans the sale of cigarettes to minors, Swart et al. (2001) found that almost two thirds of current smokers were not refused cigarettes because of their age.

important. Figure 2 shows the distribution of current grade for respondents who were age 14, 15, and 16 in Wave 1.[7] Large racial differences in grade-for-age are immediately apparent. Among white 14-year-olds, 57% were in grade 8 and 97% were between grade 7 and grade 9. Among African 14-year-olds, only 28% were in grade 8, 25% were in grade 7, and 18% were below grade 7. Similar patterns are observed at age 15 and 16, with Africans having much greater dispersion in grade-for-age than white or coloured students.

Figure 3 looks at this a different way, showing the age distribution of students in grades 9, 10, and 11. Naturally the grade-for-age dispersion in Figure 2 is reflected in the age-for-grade distributions in Figure 3. Among white 10th graders, for example, 46% are 16 and 95% are 15-17. Among African 10th graders, only 20% are 16, 18% are 17, and 44% are older than the 15-17 age range that might be considered normative for grade 10. If the African grade-for-age distributions are typical of all African schools in Cape Town, an African 15-year-old in grade 10 would be at least two years younger than 62% of her classmates and at least three years younger than 44% of her classmates. By contrast, a white 15-year-old in grade 10 would be at least two years younger than only 5% of her classmates and at least three years younger than only 1% of her classmates, assuming she were in a school with a grade-for-age distribution represented by the white students in CAPS. The potential for contagion effects from interacting with older classmates is clearly very large for African students who are not behind in school.

Table 3 shows the percentage of students who are two or more years older and three or more years older than the normative age for each grade among students in grades 9-11 (we assume that the normative ages are 15 for grade 9, 16 for grade 10, and 17 for grade 11). Among African

[7] Age-for-grade and grade-for-age distributions depend on the time of the school year at which the measures are calculated. We calculate each respondent's age and grade as of July 1, 2002 to construct Figures 2 and 3. This is roughly the middle of the South African school year.

males in grades 9-11, 57% are at least two years older than the normative age for their grade. This compares to 43% for African females. While grade repetition is common for both males and females in predominantly African schools, repetition rates are considerably higher for males (Lam et al. 2010). Coloured males and females are considerably less likely than Africans to be two or more years older than the normative age for grade. This reflects both lower rates of grade repetition and the fact that coloured students are more likely than Africans to drop out of school when they fail a grade (Lam et al. 2010)[8]. As shown Figures 2 and 3, white students are rarely two or more years older than the normative grade. Only 4% of white males and zero white females in our sample were two or more years older than the normative age in grades 9-11. Looking at the final two columns in Table 3, 33% of African males are three or more years older than the normative age for grade, a dramatic demonstration of the wide distribution in age for grade in this group. This compares to only 4% of coloured males and 1% of white males.

Table 4 looks at how key characteristics vary with age for students in a single grade – grade 9. We focus on Africans since they have by far the widest dispersion in age for grade. As shown in the bottom row, African ninth graders are fairly evenly distributed from age 14 to age 19-22. Younger students in a class are much more likely to be female than are older students. While 62% of 14-year-olds in grade 9 are female, only 38% of those age 19-22 are female. Rows 2 and 3 provide information about why there is such a wide age distribution in a given grade, showing both the age of starting school and the number of grades failed by 2002 (both estimated using the retrospective life history calendar). It is clear that many students are behind in school because

[8] Coloured youth have have better job opportunities than African youth, due in part to the legacy of coloured labor market preferences under apartheid. In addition, failing a grade is a better predictor of future failure in coloured schools than in African schools, a result of differences in the quality of evaluation (Lam et al. 2010). Both factors contribute to the higher coloured dropout rate, especially in response to failing a grade.

they started school late. The mean age at starting school for 18 year-olds in grade 9 was 7.8, compared to 6.1 for the 15-year-olds in grade 9.[9] Grade repetition also plays an important role, with 18 year-olds having failed 1.1 more grades than 15 year-olds.

Lines 4-6 of Table 4 show how our three outcomes vary across age among 9th graders. We see a steep age gradient in sexual debut. While only 3% of 14-year-olds reported having had sex, 69% of 18 year-olds and 87% of 19-22 year-olds (these two older groups accounting for 28% of African ninth graders) reported having had sex by 2002. This gradient is key to the peer effects we are focusing on. For a 14 year-old in grade 9 almost half of her classmates are already sexually active. The comparable number for coloured 9th graders is 7% and for white 9th graders is 4% (not shown). As we would expect based on the results above, the gradients for smoking and drinking are much flatter than the gradient for sexual debut. Only 10% of ninth-graders report smoking and only 9% reported drinking in the month prior to the time of the 2002 survey. Even among the 19-22 year-olds, only 24% had smoked and 22% had drunk alcohol, suggesting much smaller potential for an influence on younger classmates.

Lines 7-12 show how individual and household characteristics vary with age. Younger ninth graders performed much better on the literacy and numeracy test, with the mean for 14 year-olds being one standard deviation above the mean for 19-22 year-olds. Younger students have better educated parents, presumably indicating that better educated parents get their children into school earlier and help them make normal grade progression. Younger students are more likely to live with their mothers, although differences in household income between younger and older students are surprisingly small. All of these characteristics are likely to affect sexual debut, so we include all of these variables in our regressions below.

[9] According to the South African Schools Act of 1996, schooling is compulsory from age 7 until age 15 or the completion of grade 9.

We see from Table 4 that an African student who is fortunate enough to reach grade 9 by age 15 will have a substantial proportion of classmates who are older and sexually active. The older students will also tend to have poor academic performance. This lower achievement among the older students might have indirect peer effects that go beyond those related to sexual activity. The poorer school performance of older students may have a negative impact on the school performance of younger students, with spillover effects to sexual behavior. We will be looking at the combination of all the influences of older students and will not be able to disentangle the specific mechanism at work. While we are not able to identify exactly what it is about older students that have an effect on the behaviors of younger classmates, we answer the important question of whether exposure to older classmates has an effect.

A measure of exposure to older peers

While we cannot observe the actual distribution of ages in the schools attended by our respondents, we can use the patterns in Figure 3 to estimate the age distribution faced by students in a given grade. We take advantage of CAPS' complete schooling history for all respondents. This allows us to generate a race-specific age-for-grade distribution for each grade, pooling the retrospective histories for all respondents. We can use these to generate an estimate of the age distribution of students that each respondent experienced at every age since starting school.

A key assumption of our measure is that African students face the grade-for-age distribution displayed by all African students. In other words, we implicitly assume that the African age-for-grade distribution is identical in all African schools and that Africans only attend African schools. Analogous assumptions are made for coloured and white students. While these are obviously strong assumptions, there are several reasons to think that they are a reasonable approximation to reality. First, schools in Cape Town, like schools all over South Africa,

continue to by highly segregated. Lam et al. (2010) show using CAPS data that only 11% of Africans in grade 8 or 9 in 2002 attended historically coloured schools and only 3% attended historically white schools. While all African schools are not identical, the differences *between* African and white schools in all dimensions, including age-for-grade distributions, are vastly greater than differences *within* the group of African (or white) schools. Even if we had the specific age-for-grade distribution for each school we would not necessarily be better off using it. Since the choice of a specific school is endogenous, using the overall age-for-grade distribution of each racial group is in many ways a more valid exogenous predictor of exposure to older peers than would be the actual school-specific distribution.

Our measure of exposure to older peers is constructed as follows: We use the retrospective schooling histories to construct age-for-grade distributions for every grade for each of the three racial groups. For each respondent we look at the grade they were attending at age 12 using the retrospective schooling history. We take the race-specific age-for-grade distribution for that grade and calculate the percentage of students who would have been at least two years older than the respondent. For example, if an African respondent were in Grade 8 at age 12, we take the percentage of 8th graders who are 14 and older in the typical African age-for-grade distribution (67%) and assign that value as the percentage who were at least two years older than the respondent when she was 12. We make the same calculation at each age up to the age of the respondent in Wave 1, using only respondents who were age 14-17. Respondents who are not enrolled in school are given a zero for the exposure measure for that age.[10]

We sum these age-specific exposure measures across years from age 12 through their Wave

[10] The assumption that there is no exposure to older peers when young people are out of school is fairly unimportant since 96% of the sample in our regressions was continuously enrolled since age 12. To the extent that our measure misses peer exposure when young people are out of school we will be biased against finding an effect of our exposure measure.

1 age. For example, if the student were in grade 8 at age 12, repeated grade 8 at age 13, advanced to grade 9 at age 14, and was age 14 in Wave 1, she would have a total exposure of 0.67 +0.42 +0.43 =1.52. All African respondents with the same schooling history will get the same value. This can be thought of as a measure of person-years of exposure to classmates who were at least two years older from age 12. For a 14-year-old in 2002 this has a theoretical maximum of 3, implying that 100% of students were at least two years older than the respondent in every grade since age 12. For 14 year-old Africans in Wave 1 the mean of our exposure measure is 0.62, the standard deviation is 0.44, with a range from 0 to 1.98. The distribution for 14 year-old whites is very different, with a mean of only 0.07, a standard deviation of 0.13, and a range from 0 to 0.93.

We will use this exposure measure in regressions to see if it predicts sexual debut between 2002 and 2005. We are particularly interested in whether including this variable changes the positive sign on the "highest grade completed" variable that we found in our earlier studies. Note that in order to do this we require that grade completion in 2002 is not perfectly correlated with our exposure measure. In a regime in which all students progress one grade per year, both highest grade completed and our exposure measure might have some variation for students of a given age due to differences in the age at which students began school. This is the source of variation exploited by Argys and Rees (2008), who use differences in the mandated age at starting school across U.S. states as an instrument for exposure to older classmates. The grade variable and the exposure variable will move together as students progress, however, making it almost impossible to estimate separate effects of the two variables. In our case we take advantage of the high levels of grade repetition, especially in African and coloured schools. This means that two 16-year-olds in grade 7 in 2002 may have had very different grade trajectories since age 12. While the correlation between our exposure measure and highest grade completed ranges

between 0.85 and 0.91 for Africans for each age from 14 to 17, we will see below that we are able to estimate a statistically significant effect of the exposure measure.

Another important consideration for our analysis is the finding of Lam et al. (2010) that grade repetition is poorly linked to actual learning, especially for Africans. They find that there is a stochastic component to grade advancement that is uncorrelated with learning, suggesting that some component of our exposure measure may be unrelated to school performance. Since we will also be including grade attainment and a measure of literacy and numeracy in Wave 1, we will be able to isolate the effect of older peers from whatever association may exist between school performance and sexual debut.

One limitation of our measure is that it is only an estimate of students who were in the same grade as the respondent in a given year. It ignores the potential effect of interacting with older students from other grades or with those who are not in school. While we could construct a measure that includes students in other grades (using typical age-for-grade distributions and the grade grouping of South African schools), such a measure would be highly correlated with the measure we have constructed based on a single grade. We assume that our measure picks up both the effect of older students in the same grade and the effect of older students in other grades. If we assume that a 10th grade student is more likely than a 9th grade student to interact with 11th graders, then a 16-year-old in grade 10 will have an additional source of exposure to older peers when compared to a 16-year-old in grade 9.

Another limitation of our measure is that it does not distinguish between older male classmates and older female classmates. While we can construct separate measures for exposure to older males versus exposure to older females using the approach described above, the male and female measures are so highly correlated that it is in practice impossible to estimate separate

effects of the two measures. Given our measure, students with high exposure to older males will also have had high exposure to older females. While we would be better able to understand the mechanisms if we could separate the effect of older males from the effect of older females, it is not something we can do in practice.[11]

EMPIRICAL RESULTS

In this section we present regressions analyzing the impact of our peer exposure measure on four outcomes – (1) sexual debut between 2002 and 2005; (2) the age difference of the first sexual partner for those who become sexually active between 2002 and 2005; (3) whether the respondent smoked in the month before the 2002 interview; (4) whether the respondent drank alcohol in the month before the 2002 interview. The analysis uses CAPS respondents who were aged 14-17 in 2002. In order to control carefully for age we include a quadratic of age in months. We include an indicator for 2002 school enrollment, the highest grade attained in 2002, and the standardized score on the literacy and numeracy exam administered in 2002. Since there is variation in the time between Wave 1 and Wave 3 interviews we include a control for the number of months between interviews. We also include a number of household characteristics. These include mother's and father's education (these were collected from the youth respondent even when the parent was not coresident); log of per capita household income in 2002; dummies for coloured and white; dummies to indicate whether the mother and father were coresident with the young adult in 2002; dummies to indicate that parental education is missing.

Descriptive statistics

Table 5 presents descriptive statistics of key variables, broken down by gender and

[11] Only a small minority of students in South Africa attend single-sex schools. We have explored trying to identify respondents in single-sex schools using CAPS data on school names, but the number we can identify is very small. The single-sex schools we can identify are mainly elite private schools catering to high-income whites, so this provides very little information about peer effects in the full population.

population group. We see large racial differences in sexual activity by 2002. About 30% of African males and females reported having had sex in CAPS Wave 1, compared to 4-14% for coloured and white youth. Our analysis of sexual debut is restricted to the sample that had not had sex by 2002. Within this group, 68% of African females and 61% of African males become sexually active by 2005. This compares to 37% of coloured females, 30% of white females, 40% of coloured males, and 36% of white males. We will also analyze the age difference of the first sexual partner for those who become sexually active. This difference ranges from 2.2 to 2.6 years for females (meaning the male partner is older), and from 0.12 to -0.5 years for males.

Table 5 includes several key schooling variables. School enrollment is well over 90% for all groups, but we see large racial differences in grade attainment.[12] The largest differences are for males, with African males aged 14-17 having completed 6.8 grades, compared to 8.1 and 8.6 for coloured and white males. Our measure of exposure to older peers, explained above, has a mean of 0.96 for African females, 0.46 for coloured females, and 0.09 for white females. We see large racial differences in performance on the literacy and numeracy evaluation that was administered in Wave 1. This was a self-administered written test taken after completion of the young adult questionnaire. The test had 45 questions and took about 20 minutes to complete. Respondents could choose to take the test in English or Afrikaans. There was no version in Xhosa, the home language of most African respondents. The English language test was taken by 99% of African respondents, 43% of coloured respondents, and 64% of white respondents.[13] We use the score as a measure of cumulative learning as of Wave 1, with performance on the test reflecting factors

[12] As previously noted, schooling is compulsory until age 15 or the completion of grade 9. While there are some apparent violations of this in CAPS and other South African surveys, most young people stay in school beyond the legal minimum.

[13] Although it is important to keep in mind that Africans took the test in a second language, it must also be noted that English is the official language of instruction in African schools and is used for important tests such as the grade 12 matriculation exam.

such as innate ability, home environment, and the quantity and quality of schooling to that point. As seen in Table 5, African females have a mean score that is 1.6 standard deviations below the mean score for white females. As shown in Lam et al. (2010), the distribution of test scores for Africans and whites barely overlap. There are also enormous racial differences in income. Household income per capita is almost ten times as high in the households of white 14-17 year-olds as African 14-17 year-olds. Income in coloured households is about twice the income in African households.

Determinants of sexual debut

Table 6 presents probit regressions in which the dependent variable is equal to 1 if the respondent became sexually active between 2002 and 2005, using the sample that had not had sex by 2002. We present marginal effects (evaluated at the sample mean for all variables) along with robust standard errors adjusted for sample clustering (in brackets). Given sample size limitations we pool the population groups but estimate separate regressions for males and females. Columns 1 and 4 leave out our peer exposure measure and the literacy/numeracy store. We estimate a positive impact of grade attainment on sexual debut, consistent with previous estimates using CAPS. The estimated effect of grade attainment implies that a girl with one additional year of schooling in 2002 (controlling for age) would be 3.6 percentage points more likely to become sexually active by 2005. The effect for boys is similar – 4.4 percentage points per year of schooling. We estimate a negative but statistically insignificant effect of being in school in 2002 on sexual debut over the next three years. We get very large negative marginal effects on the coloured and white dummies, indicating that the variables included in the regression do not explain the large racial differences in early sexual debut.

Columns 2 and 5 add the literacy and numeracy evaluation (LNE) score to the regressions.

The LNE score itself has a negative statistically significant effect on sexual debut for both males and females. A one standard deviation increase in the test score is associated with an 8.3 percentage point reduction in the probability of sexual debut for females. Also noteworthy is that including the LNE score causes the effect of highest grade completed to become even more positive for both males and females. This is consistent with our hypothesis that the grades completed variable is partially picking up the effect of exposure to older peers. When we don't include the LNE score the highest grade variable picks up two offsetting effects. The first effect is that students who are doing better in school and are more committed to school may be less likely to become sexually active. The second effect is the influence of older peers, which tends to encourage sexual debut. When we include the LNE score it picks up some of the first effect, leaving the highest grade variable to pick up more of the second effect.

Columns 3 and 6 introduce our measure of exposure to peers at least two years older since age 12. This variable is estimated to have a statistically significant positive effect on sexual debut for females. The marginal effect of 0.138 implies that an increase in cumulative exposure by 1.0 would increase the probability of sexual debut by 13.8 percentage points. An increase in cumulative exposure of 1.0 could result from an increase in the percentage of classmates who were at least two years older by 25 percentage points in each of four years since age 12, an increase by 50 percentage points in each of two years, or any other combination that adds up to 1.0. The standard deviation of this variable for Africans is 0.7, so an increase of 1.0 is an empirically plausible example. The estimated effect of the peer exposure variable is only about half as large for males and is not statistically significant. This is similar to the results of Argys and Rees (2008), who find significant peer effects for females but not for males in the U.S..

Another important result from Table 6 is that including the peer exposure variable causes the

estimated effect of grades completed to become much smaller for both males and females. For females it also loses its statistical significance, while for males the coefficient estimate is now only significant at the 10% level. For females the estimated marginal effect of grades completed falls from 0.058 in Regression 2 to 0.015 in Regression 3. This supports our hypothesis that the apparent positive effect of grades completed on sexual debut is due to an effect of exposure to older peers. It is also striking that the coefficient for coloured drops by about 25% for females when the peer exposure variable is added to the regression. The coefficient for white drops by over 60% and becomes statistically insignificant. This suggests that the much higher exposure of African girls in secondary school to peers who are at least two years older plays a substantial role in explaining the earlier sexual debut of African girls compared to coloured and white girls.

Since variation in our peer exposure measure depends on variation in the age of starting school, interruptions in schooling, and grade repetition, it is likely to be correlated with characteristics such as the student's (and parents') commitment to school, the student's academic ability, and neighborhood characteristics. Controlling for baseline variables such as grade attainment, literacy/numeracy scores, and household income should remove much of this correlation, but the exposure variable may still be correlated with unobserved determinants of sexual debut.[14] Most of these effects would lead us to expect that students who are farther ahead in school (and thus have high values of the peer exposure variable) would be *less* likely to become sexually active. The bias in our estimates, then, should work against finding a positive effect of exposure to older peers on sexual debut. The fact that we do estimate a positive effect

[14] In regressions not shown we have also tried including measures of school quality (pupil-teacher ratios and the percentage of teachers hired by the parent governing body, taken from matching the 2000 School Register of Needs to school names reported in CAPS) and neighborhood youth unemployment (based on 2001 census data at the "small area" level). The results are almost identical, with a slightly larger and statistically more significant estimate of the impact of the peer exposure measure.

gives us confidence that the effect we are measuring is a real effect of peer exposure.

Peer effects and the age of first sexual partner

The impact of older classmates on sexual debut could work through a number of channels. The simplest version of the "contagion effect" hypothesis is that interacting with peers who are sexually active, whether same sex or opposite sex, may make it more likely that an individual decides to experiment with sexual activity. Another possible channel is that individuals become sexually active with their classmates (or friends of their classmates). This might be especially important for girls, who, as shown in Table 5, have first sexual partners who are on average two to three years older. We might expect, then, that exposure to older classmates would have an effect on the age of the first sexual partner in addition to having an effect on sexual debut. CAPS collected information on a number of characteristics of the first sexual partner, including age. Table 7 presents regressions in which the dependent variable is the age difference of the first sexual partner, using only the sample that became sexually active between 2002 and 2005.

Looking at Regression 1 in Table 7, we estimate a statistically significant positive effect of our peer exposure variable on the age difference of the first sexual partner for females. An increase in the cumulative peer exposure of 1.0 is associated with an increase in the age difference of the first sexual partner of 0.87 years. This provides additional evidence that exposure to older peers is affecting sexual behavior. While we cannot tell whether the first sexual partner is a classmate (or a friend of a classmate), the results in Tables 6 and 7 suggest that girls with older classmates have both earlier sexual debut and older first sexual partners. The estimated effect of exposure to older peers for males is also positive in sign, but is much smaller (0.06) and not statistically significant.

The impact of older peers on smoking and drinking

Research on the impact of older peers often includes analysis of smoking and drinking, two outcomes thought to be sensitive to contagion effects (Eisenberg 2004; Argys and Rees 2008). CAPS includes relatively simple questions about whether the respondent smoked any cigarettes or consumed any alcohol "over the past month." As shown in Figure 1 and Table 5, there are large racial differences in smoking and drinking among teenagers. Only 1% of African girls aged 14-17 report smoking in the last month, compared to 28% of coloured girls and 17% of white girls. Only 3% of African girls aged 14-17 report drinking alcohol in the last month, compared to 12% of coloured girls and 32% of white girls. As previously noted, we see from Figure 1 that the age gradient for these behaviors is much less steep than the age gradient for sexual debut, making it less likely that we will see an impact of exposure to older peers.

Table 8 presents marginal effects from probit regressions for smoking and drinking, estimated separately for males and females. While the estimated effect of our peer exposure variables is positive in all four regressions, the effects are small in magnitude and are never statistically significant. We do estimate a statistically significant negative association between being enrolled in school in 2002 and smoking for girls and for both smoking and drinking for boys. We also estimate statistically significant negative effects of the literacy/numeracy score on smoking for both girls and boys.

While the absence of an impact of our peer exposure variable on smoking and drinking might be seen as weakening our argument about peer effects, we see these results as easily explained by the patterns shown in Figures 1-3. In order for there to be an effect of older classmates on behavior, there needs to be a steep age gradient for that outcome over the relevant ages. The age gradient for sexual debut is much steeper than the age gradients for smoking and drinking, with the possible exception of the white pattern for drinking. Whites are much less

likely to be exposed to older peers as classmates, however, given their much lower variance in age-for-grade. Only the sexual debut outcome has a steep age gradient for the groups that experience significant exposure to older peers as classmates. It is therefore entirely consistent that we see significant effects of older peers on sexual debut but not on smoking and drinking.

The impact of passing a grade on sexual debut

One somewhat extreme implication of our hypothesis that the presence of older classmates increases the probability of sexual debut is that students who pass a given grade, especially in high school, are more likely to become sexually active than students the same age who fail the grade. For this to occur the impact of older classmates would have to be strong enough to overcome what is presumably a tendency for students who pass a grade to be students with a stronger commitment to school and an associated lower probability of becoming sexually active. In this section we examine this directly by looking at the impact of passing a grade on subsequent sexual debut.

An advantage of using grade advancement as an independent variable is that it is a more direct measure than our cumulative peer exposure measure. It has larger within-race variation, since the cumulative exposure measure is based on average age-for-grade distributions by race. The disadvantage of the grade advancement measure is that it is even more subject to endogeneity bias than the cumulative exposure measure, since students who pass a given grade will be positively selected on many characteristics that are likely to deter sexual debut. The biases created by the endogeneity of grade advancement will therefore tend to work against us finding a positive relationship between grade advancement and sexual debut.

Table 9 presents regressions using grade advancement as an independent variable in our sexual debut regressions. We estimate separate regressions by race and gender, taking advantage

of the fact that there is substantial within-race variation in grade advancement. The dependent variable equals 1 if a respondent was in school in 2002 and was in school in 2003 in a grade at least one grade higher than 2002. The grade advancement variable equals 0 if the respondent was in school in 2002 and was in school in 2003 in a grade at or below the 2002 grade. As shown in Table 5, the 2002 pass rate was 88% for African males and females, 92% for coloured females, 90% for coloured males, and 99% for white males and females. Because of the almost 100% pass rate for whites we exclude them from the analysis in Table 9.[15]

We restrict the analysis to students in grades 9-11 in 2002, meaning they are in grades 9-12 in 2003. The dependent variable is the same as in Table 6 – it equals 1 if the respondent becomes sexually active between 2002 and 2005 and equals 0 if the respondent does not begin sexual activity by 2005. As before, we control for age with a quadratic in age in months. Regressions 1 and 2 show that for African and coloured females, those who pass in 2002 have a statistically significant higher probability of sexual debut than those who fail. African females who advance one grade from 2002 to 2003 are 21 percentage points more likely to become sexually active by 2005 than African females who remain in the same grade in 2003. The difference for coloured females is 14 percentage points. The point estimates for males are negative and not statistically significant. For African females we continue to get a positive and statistically significant impact of the peer exposure variable, on top of the impact of advancing a grade from 2002 to 2003.[16]

Passing the grade in 2002 is obviously not a randomly assigned treatment. Girls who pass are likely to differ in many respects from girls who fail, even after controlling for the individual and household characteristics included in the regressions in Table 9. Most of those differences,

[15] In order to maximize sample size the regressions in Table 9 use all respondents in grades 9-11 regardless of age. If we restrict the analysis to those under age 17 the results are very similar.

[16] Note that the peer exposure variable is only measured through 2002 and is unaffected by the outcome of the 2002 grade.

such as unmeasured attachment to school, school quality, and home and neighborhood environmental influences, would lead us to expect that girls who fail the grade would be more likely, not less likely, to become sexually active. One variable working in the direction of the effect we see is that girls who pass the grade will be exposed to older classmates in the following year. The results in Table 9 suggest that this peer exposure effect is strong enough to cause a positive relationship between passing a grade and subsequent sexual debut.

SUMMARY AND CONCLUSIONS

South Africa's combination of early sexual debut, high rates of school enrollment through late teenage years, and high rates of grade repetition create an environment in which moving through school faster than one's age-mates means being exposed to significant numbers of older classmates who are already sexually active. Previous research provided suggestive evidence of such peer effects, identifying a surprising positive relationship between grade attainment and subsequent sexual debut, controlling for age. This paper attempts to provide clearer evidence about the existence of peer "contagion" effects, taking advantage of several features of the Cape Area Panel Study. We document two important features of schooling and sexual debut that create the potential for contagion effects. First, we show that there is a steep age-gradient in sexual debut for males and females in all three of the population groups we study – African, coloured, and white. This gradient is much steeper than the gradient observed for smoking and drinking. Second, we show that high rates of grade repetition and secondary enrollment that continues even beyond age 20 lead to high variance in the age-for-grade distribution, especially for Africans. Using the retrospective schooling histories in CAPS, we generate race-specific age-for-grade distributions for all grades and use these to estimate the history of exposure to classmates who are at least two years older beginning at age 12.

Our probit regressions indicate that our measure of cumulative exposure to older peers in

2002 has a positive and statistically significant impact on sexual debut between 2002 and 2005, controlling for baseline age, grade attainment, literacy and numeracy competence, and a number of household background variables. Being exposed to 50% more classmates who are at least two years older for two years increases the probability that a girl becomes sexually active between 2002 and 2005 by 14 percentage points. In addition, inclusion of our peer exposure variable causes the estimated effect of baseline grade completion to drop from a statistically significant positive effect to a much smaller and statistically insignificant effect. The exposure variable also causes the estimated marginal effect of being coloured rather than African to fall by 25% and the estimated effect of being white to drop by over 60% and become statistically insignificant. This is provocative evidence that the earlier sexual debut of African girls may be partly due to the much higher degree of exposure to older classmates experienced in African schools. As further evidence of this effect, we show that girls in grades 9-11 who passed their grade in 2002 were significantly more likely to become sexually active than girls who did not pass the grade.

Our measure of exposure to older peers is also estimated to have a statistically significant positive effect on the age of the first sexual partner for females. This is consistent with the argument that exposure to older peers is a factor in encouraging earlier sexual debut. We do not find statistically significant effects of our peer exposure measure on smoking and drinking behavior. We argue that this is consistent with the fact that smoking and drinking have much lower age gradients than sexual debut, especially in the African sample where there is the highest exposure to older classmates.

We are not able to identify the precise mechanisms driving the effect of exposure to older classmates. Data limitations mean that we cannot separate the effect of exposure to older girls versus older boys. It is possible that the results are driven in part by girls having sex with older

male classmates. Alternatively, the results may indicate that girls are being influenced to adopt the behavior of their older female classmates, a contagion effect in the narrow sense. Whatever the mechanism, the results suggest that South Africa's high rates of grade repetition create an environment in which exposure to older classmates is contributing to earlier sexual debut.

Our findings are a potentially troubling caveat to the view that schooling is protective in terms of delaying sexual debut. Although girls who are in school are less likely to become sexually active than girls who are not in school, girls who are ahead of their cohorts in advancing through school face the potentially negative consequences of interacting with classmates who may be four or five years older. This may be an important factor that should be taken into account in evaluating the high rates of grade repetition in disadvantaged South African schools.

REFERENCES

Anderson, Kermyt, Anne Case and David Lam. 2001. "Causes and Consequences of Schooling Outcomes in South Africa: Evidence from Survey Data." *Social Dynamics* 27(1): 1-23.

Argys, Laura M. and Daniel I. Rees. 2008. "Searching for Peer Group Effects: A Test of the Contagion Hypothesis," *Review of Economics and Statistics* 90(3): 442-458.

Darabi, Leila, Akinrinola Bankole, Kalundi Serumaga, Stella Neema, Richard Kibombo, Humera Ahmed and Paul Banoba. 2008. "Protecting the next generation in Uganda: New evidence on adolescent sexual and reproductive health needs." New York City: The Guttmacher Institute.

Dinkelman, Taryn, David Lam, and Murray Leibbrandt. 2007. "Household and community income, economic shocks and risky sexual behavior of young adults: evidence from the Cape Area Panel Study 2002 and 2005." *AIDS*, November 2007, (21 Suppl 7):S49-S56.

Dinkelman, Taryn, David Lam, and Murray Leibbrandt. 2008. "Linking Poverty and Income Shocks to Risky Sexual Behavior: Evidence from a Panel Study of Young Adults in Cape Town," *South African Journal of Economics*, May 2008, 76(S1):53-74.

Eisenberg, Daniel. 2004. "Peer Effects for Adolescent Substance Use: Do They Really Exist?" UC-Berkeley School of Public Health working paper.

Fitzgerald, John, Peter Gottschalk and Robert Moffitt. 1998. "The Impact Of Attrition In The Panel Study Of Income Dynamics On Intergenerational Analysis," *Journal of Human Resources* 33(2): 300-344.

Gaviria, Alejandro, and Steven Raphael. 2001. "School-Based Peer Effects and Juvenile Behavior," *The Review of Economics and Statistics*, May 2001, 257-268.

Gersovitz, Mark. 2005."The HIV Epidemic in Four African Countries Seen through the Demographic and Health Surveys," *Journal of African Economies* 14(2): 191-246.

Lam, David, Cally Ardington, Nicola Branson, Anne Case, Murray Leibbrandt, Alicia Menendez, Jeremy Seekings and Meredith Sparks. 2008. The Cape Area Panel Study: Overview and Technical Documentation of Waves 1-2-3-4. The University of Cape Town, October 2008.

Lam, David, Cally Ardington and Murray Leibbrandt. 2010. "Schooling as a lottery: Racial differences in school advancement in Urban South Africa," *Journal of Development Economics*, forthcoming.

Lloyd, Cynthia. 2005. *Growing Up Global: The Changing Transitions to Adulthood in Developing Countries.* Washington: National Academies Press.

Lloyd, Cynthia. 2007. "The role of schools in promoting sexual and reproductive health among adolescents in developing countries." Poverty, Gender, and Youth Working Paper No. 6, New York: Population Council.

Marteleto, Letícia, David Lam, and Vimal Ranchhod. 2008. "Sexual Behavior, Pregnancy, and Schooling among Young People in Urban South Africa," *Studies in Family Planning* 39(4): 351–368.

Nnko, Soori, J Ties Boerma, Mark Urassa, Gabriel Mwaluko, and Basia Zaba. 2004. “Secretive females or swaggering males?: An assessment of the quality of sexual partnership reporting in rural Tanzania,” *Social Science & Medicine* 59(2): 299-310.

Sieving, Renee E. Marla E. Eisenberg, Sandra Pettingell, and Carol Skay. 2006. “Friends’ Influence on Adolescents’ First Sexual Intercourse,” *Perspectives on Sexual and Reproductive Health* 38(1): 13-19.

Sacerdote, Bruce. 2001. “Peer Effects with Random Assignment: Results for Dartmouth Roommates.” *The Quarterly Journal of Economics*, May 2001, 681-704.

Swart, Dehran, Priscilla Reddy, Blanche Pitt and Saadhna Panday. 2001. “The prevalence and determinants of tobacco use among grade 8-10 learners in South Africa.” Medical Research Council: Cape Town.

Table 1. Sample size by population group and attrition between waves, Cape Area Panel Study Waves 1-4

Population group	Wave 1	Unweighted percent	Weighted percent	Interviewed in Wave 3	Interviewed in Wave 3 or Wave 4	Attrition
Full sample aged 14-22 in 2002						
Black/African	2,151	45.27	28.22	1,515	1,724	19.9%
Coloured	2,005	42.2	53.16	1,679	1,801	10.2%
White	595	12.52	18.62	337	391	34.3%
Total	4,751	100	100	3,531	3,916	17.6%
Sample aged 14-17 in 2002						
Black/African	902	41.09	25.7	696	772	14.4%
Coloured	999	45.51	55.3	881	935	6.4%
White	294	13.39	19	205	228	22.4%
Total	2195	100	100	1,782	1,935	11.8%

Table 2. OLS regressions of outcomes on age, CAPS respondents aged 14-20 in wave 1

	African		Coloured		White	
Outcome	**Male**	**Female**	**Male**	**Female**	**Male**	**Female**
Had sex by wave 1						
Age coefficient	0.122***	0.145***	0.108***	0.0951***	0.100***	0.0979***
Age std. error	[0.0079]	[0.0069]	[0.0075]	[0.0068]	[0.012]	[0.012]
Observations	726	970	779	862	239	236
Smoked in last month in wave 1						
Age coefficient	0.0500***	0.00245	0.0678***	0.0533***	0.0655***	0.0314**
Age std. error	[0.0069]	[0.0022]	[0.0087]	[0.0083]	[0.015]	[0.014]
Observations	726	965	786	869	243	242
Consumed alcohol in last month in wave 1						
Age coefficient	0.0362***	0.00511	0.0784***	0.0449***	0.116***	0.114***
Age std. error	[0.0063]	[0.0035]	[0.0077]	[0.0069]	[0.015]	[0.015]
Observations	726	967	786	870	243	242

Standard errors in brackets, *** p<0.01, ** p<0.05, * p<0.1

Table 3. Percentage 2+ and 3+ years older than normative age for grade, CAPS wave 1 respondents enrolled in grades 9-11 in 2002

		2+ years older		3+ years older	
Population Group	**N**	**Males**	**Females**	**Males**	**Females**
African	797	57%	43%	33%	22%
Coloured	680	14%	8%	4%	2%
White	222	5%	0%	1%	0%
Total	1699	24%	18%	12%	8%

Note: Normative age defined as 15 for grade 9, 16 for grade 10, and 17 for grade 11.

Table 4. Characteristics of Africans Enrolled in Grade 9 by Age, CAPS Wave 1, 2002

		Age in 2002						
	Variable	**14**	**15**	**16**	**17**	**18**	**19-22**	**Total**
1	Female	0.62	0.71	0.63	0.51	0.49	0.38	0.56
2	Number of grades failed	0.14	0.22	0.64	0.99	1.31	1.89	0.86
3	Age started school	5.22	6.06	6.61	7.22	7.76	8.55	6.94
4	Had sex by 2002	0.03	0.28	0.38	0.47	0.69	0.87	0.46
5	Smoking by 2002	0.00	0.04	0.07	0.11	0.14	0.24	0.10
6	Drinking by 2002	0.00	0.02	0.09	0.08	0.10	0.22	0.09
7	Standardized LNE score	-0.12	-0.36	-0.63	-0.63	-0.78	-1.10	-0.62
8	Mother's education	9.72	8.43	8.32	7.64	8.02	6.11	7.99
9	Father's education	8.80	7.26	7.50	6.95	6.94	5.53	7.11
10	Log household income	5.85	5.52	5.47	5.43	5.51	5.20	5.48
11	Mother co-resident	0.74	0.75	0.68	0.71	0.57	0.60	0.68
12	Father co-resident	0.48	0.31	0.39	0.40	0.48	0.40	0.40
13	Sample size	33	54	80	66	44	46	323
14	Percentage	10.2%	16.7%	24.8%	20.4%	13.6%	14.2%	100%

Table 5. Means and standard deviations of key variables, Cape Area Panel Study respondents aged 14-17 in 2002

	Female			Male		
Variable	**African**	**Coloured**	**White**	**African**	**Coloured**	**White**
Overall sample size	440	492	117	340	443	111
Had sex by 2002	0.30	0.08	0.05	0.34	0.14	0.04
	(0.46)	(0.28)	(0.23)	(0.47)	(0.35)	(0.20)
Smoking in 2002	0.01	0.28	0.17	0.11	0.36	0.16
	(0.12)	(0.45)	(0.38)	(0.31)	(0.48)	(0.37)
Drinking in 2002	0.03	0.12	0.32	0.09	0.16	0.44
	(0.18)	(0.33)	(0.47)	(0.29)	(0.37)	(0.50)
Conditional on no sex by 2002:						
Conditional sample size	265	404	90	192	347	90
Sexual debut 2002-2005	0.68	0.37	0.30	0.61	0.40	0.36
	(0.47)	(0.48)	(0.46)	(0.49)	(0.49)	(0.48)
Age difference of first sexual partner	2.65	2.86	2.21	0.12	0.10	-0.46
	(2.19)	(2.92)	(2.08)	(2.35)	(1.86)	(1.59)
Enrolled in school in 2002	0.98	0.94	1.00	0.98	0.93	0.99
	(0.13)	(0.23)	(0.00)	(0.15)	(0.25)	(0.09)
Grades completed in 2002	7.68	8.36	8.63	6.80	8.14	8.64
	(1.48)	(1.40)	(1.18)	(1.61)	(1.46)	(1.31)
Grades completed in 2005	9.84	10.23	11.12	9.21	9.79	11.1
	(1.44)	(1.67)	(0.90)	(1.58)	(1.77)	(0.96)
Passed grade in 2002	0.88	0.92	0.99	0.88	0.90	0.99
	(0.32)	(0.27)	(0.10)	(0.33)	(0.30)	(0.10)
Exposure to peers 2+ years older	0.96	0.46	0.09	0.68	0.40	0.07
	(0.73)	(0.36)	(0.13)	(0.62)	(0.36)	(0.10)
Literacy and numeracy score (standardized)	-0.47	0.08	1.17	-0.59	0.12	1.31
	(0.82)	(0.78)	(0.55)	(0.84)	(0.86)	(0.55)
Household income per capita	441	944	4270	453	999	4081
	556	914	2968	733	1060	2784
Log household income per capita	5.59	6.49	8.10	5.59	6.53	8.09
	(1.00)	(0.87)	(0.80)	(0.95)	(0.88)	(0.70)
Mother's highest grade	8.49	8.68	12.65	8.63	8.98	12.83
	(3.02)	(3.00)	(1.65)	(2.77)	(2.77)	(1.91)
Father's highest grade	7.83	9.12	13.30	7.81	8.99	13.03
	(3.80)	(3.22)	(2.20)	(3.78)	(3.19)	(1.89)
Mother co-resident in 2002	0.72	0.81	0.94	0.79	0.83	0.97
	(0.45)	(0.39)	(0.24)	(0.41)	(0.38)	(0.18)
Father co-resident in 2002	0.42	0.55	0.77	0.44	0.61	0.80
	(0.50)	(0.50)	(0.42)	(0.50)	(0.49)	(0.40)
Mother's grade missing	0.09	0.09	0.04	0.08	0.10	0.00
	(0.29)	(0.28)	(0.19)	(0.27)	(0.30)	(0.00)
Father's grade missing	0.40	0.32	0.10	0.39	0.28	0.08
	(0.49)	(0.47)	(0.30)	(0.49)	(0.45)	(0.27)

Note: Standard deviation in parentheses. Variable for exposure to older students is cumulative from age 12. Household income per capita in rands per month in 2002.

Table 6. Marginal effects from probit regressions for sexual debut between 2002 and 2005, CAPS respondents aged 14-17 in 2002

	Female			Male		
Variable	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Enrolled 2002	-0.169	-0.192*	-0.173*	-0.304***	-0.289***	-0.280***
	[0.106]	[0.106]	[0.105]	[0.0948]	[0.0967]	[0.0992]
Grades completed in 2002	0.0355	0.0576**	0.0152	0.0435**	0.0628***	0.0468*
	[0.023]	[0.024]	[0.032]	[0.018]	[0.020]	[0.028]
Exposure to peers 2+ years older			0.138**			0.0714
			[0.069]			[0.084]
Literacy/numeracy score		-0.083***	-0.083***		-0.060*	-0.062*
		[0.031]	[0.031]		[0.032]	[0.032]
Coloured	-0.356***	-0.343***	-0.252***	-0.328***	-0.317***	-0.281***
	[0.047]	[0.047]	[0.068]	[0.052]	[0.053]	[0.069]
White	-0.282***	-0.238***	-0.0934	-0.169*	-0.122	-0.0513
	[0.0700]	[0.0756]	[0.116]	[0.0901]	[0.100]	[0.141]
Log household income per capita	-0.0182	-0.0084	-0.0103	-0.0397	-0.0311	-0.0324
	[0.0291]	[0.0291]	[0.0292]	[0.0304]	[0.0303]	[0.0307]
Mother's highest grade	0.0078	0.0121	0.0118	-0.0188*	-0.0172*	-0.0166*
	[0.009]	[0.009]	[0.009]	[0.010]	[0.010]	[0.010]
Father's highest grade	-0.0283***	-0.0270***	-0.0266***	-0.0153	-0.0148	-0.0153
	[0.009]	[0.009]	[0.009]	[0.010]	[0.010]	[0.010]
Mother co-resident in 2002	-0.0279	-0.0188	-0.0237	-0.0614	-0.0547	-0.0532
	[0.0633]	[0.0635]	[0.0641]	[0.0663]	[0.0659]	[0.0660]
Father co-resident in 2002	-0.105*	-0.109*	-0.113**	-0.0767	-0.0817	-0.0770
	[0.0554]	[0.0554]	[0.0554]	[0.0664]	[0.0667]	[0.0671]
Age in months since age 14	0.0169***	0.0161***	0.0176***	0.0172**	0.0169**	0.0170**
	[0.006]	[0.006]	[0.006]	[0.007]	[0.007]	[0.007]
Age in months squared (*1000)	-0.210*	-0.216*	-0.210*	-0.163	-0.173	-0.160
	[0.114]	[0.113]	[0.114]	[0.142]	[0.143]	[0.143]
Number months between waves	0.00302	0.00304	0.00350	-0.0196*	-0.0195*	-0.0199*
	[0.0107]	[0.0106]	[0.0107]	[0.0110]	[0.0112]	[0.0112]
Mother's grade missing	0.0377	0.0830	0.0700	-0.208**	-0.187*	-0.180*
	[0.123]	[0.124]	[0.125]	[0.102]	[0.108]	[0.109]
Father's grade missing	-0.276***	-0.275***	-0.274***	-0.134	-0.133	-0.135
	[0.0882]	[0.0886]	[0.0888]	[0.104]	[0.105]	[0.105]
Observations	808	808	808	683	683	683

Robust standard errors in brackets, adjusted for sample clustering;
* significant at 10%; ** significant at 5%; *** significant at 1%
Marginal effects evaluated at sample mean for all variables.

Table 7. OLS regressions for age difference of first sexual partner, CAPS respondents aged 14-17 in 2002

	Female	Male
Variable	**(1)**	**(2)**
Enrolled 2002	0.859	-0.137
	[0.68]	[0.529]
Grades completed in 2002	-0.475*	-0.109
	[0.28]	[0.223]
Exposure to peers 2+ years older	0.870*	0.0571
	[0.46]	[0.433]
Literacy/numeracy score	0.588**	0.214
	[0.25]	[0.207]
Coloured	1.226**	0.0763
	[0.53]	[0.424]
White	1.749*	-0.482
	[1.05]	[0.861]
Log household income per capita	-0.488**	0.00166
	[0.19]	[0.139]
Mother's highest grade	-0.0185	0.0265
	[0.070]	[0.0481]
Father's highest grade	-0.0474	-0.0222
	[0.073]	[0.0455]
Mother co-resident in 2002	-0.141	-0.233
	[0.45]	[0.480]
Father co-resident in 2002	0.136	0.215
	[0.39]	[0.336]
Age in months since age 14	0.00599	0.000670
	[0.050]	[0.0295]
Age in months squared (*1000)	0.308	-0.122
	[0.90]	[0.668]
Number months between waves	-0.007	0.0478
	[0.079]	[0.0727]
Mother's grade missing	-0.591	-0.171
	[0.84]	[0.648]
Father's grade missing	0.223	0.464
	[0.71]	[0.556]
Observations	348	280

Robust standard errors in brackets, adjusted for sample clustering;
* Significant at 10%; ** significant at 5%; *** significant at 1%
Sample is restricted to respondents who had not had sex in Wave 1.

Table 8. Marginal effects from probit regressions for smoking and drinking in Wave 1, CAPS respondents aged 14-17 in 2002

	Female		Male	
Variable	**Smoking**	**Drinking**	**Smoking**	**Drinking**
Enrolled 2002	-0.252***	-0.0672	-0.280***	-0.191***
	[0.082]	[0.059]	[0.075]	[0.072]
Grades completed in 2002	0.0194	0.0253	-0.0141	-0.00423
	[0.017]	[0.017]	[0.024]	[0.018]
Exposure to peers 2+ years older	0.0127	0.00336	0.0317	0.0204
	[0.039]	[0.029]	[0.055]	[0.044]
Literacy/numeracy score	-0.0316*	-0.0242	-0.0388*	-0.0229
	[0.019]	[0.017]	[0.023]	[0.019]
Coloured	0.345***	0.109***	0.298***	0.0743*
	[0.043]	[0.036]	[0.048]	[0.044]
White	0.568***	0.451***	0.329**	0.419***
	[0.12]	[0.12]	[0.13]	[0.14]
Log household income per capita	-0.0129	0.0133	-0.0226	0.027
	[0.016]	[0.014]	[0.023]	[0.021]
Mother's highest grade	-0.0060	-0.0073	0.0028	0.0018
	[0.0054]	[0.0045]	[0.0076]	[0.0063]
Father's highest grade	0.0050	0.0006	-0.0044	-0.0017
	[0.0053]	[0.0043]	[0.0072]	[0.0061]
Mother co-resident in 2002	0.0098	0.0243	-0.134**	-0.0100
	[0.034]	[0.029]	[0.060]	[0.040]
Father co-resident in 2002	-0.015	-0.0492*	-0.0241	-0.0625
	[0.031]	[0.027]	[0.045]	[0.041]
Age in months since age 14	0.0064*	0.00335	0.0137***	0.00788*
	[0.0038]	[0.0042]	[0.0051]	[0.0046]
Age in months squared (*1000)	-0.105	-0.0399	-0.125	-0.0144
	[0.070]	[0.072]	[0.097]	[0.085]
Number months between waves	0.00324	-0.00205	-0.00597	0.000312
	[0.0060]	[0.0051]	[0.0089]	[0.0068]
Mother's grade missing	-0.0194	-0.0623*	-0.0751	-0.0213
	[0.065]	[0.032]	[0.076]	[0.073]
Father's grade missing	0.0312	0.0116	-0.0922	-0.0408
	[0.062]	[0.046]	[0.067]	[0.060]
Observations	1020	1023	872	873

Robust standard errors in brackets, adjusted for sample clustering.
* significant at 10%; ** significant at 5%; *** significant at 1%
Marginal effects evaluated at sample mean for all variables.

Table 9. Sexual debut between 2002 and 2005 by grade advancement 2002-03, CAPS respondents enrolled in grades 9-11 in 2002

	Female		Male	
	African	Coloured	African	Coloured
Variable	(1)	(2)	(3)	(4)
Advanced grade from 2002 to 2003	0.207**	0.137*	-0.105	-0.007
	[0.105]	[0.082]	[0.0711]	[0.133]
Exposure to peers 2+ years older	0.139**	0.083	0.135	0.163
	[0.059]	[0.134]	[0.115]	[0.175]
Literacy/numeracy score	-0.113***	-0.085*	-0.109**	0.007
	[0.038]	[0.049]	[0.049]	[0.062]
Log household income per capita	-0.0448	-0.0667	-0.0032	-0.0331
	[0.0451]	[0.0471]	[0.0392]	[0.0615]
Mother's highest grade	0.0181	0.0226*	0.0268**	-0.0456**
	[0.0138]	[0.0128]	[0.0131]	[0.0185]
Father's highest grade	-0.0007	-0.0360**	-0.0029	-0.0044
	[0.0104]	[0.0150]	[0.0147]	[0.0200]
Mother co-resident in 2002	-0.096	0.0601	-0.154**	-0.504***
	[0.082]	[0.087]	[0.072]	[0.116]
Father co-resident in 2002	-0.041	0.018	-0.101	0.135
	[0.077]	[0.081]	[0.102]	[0.107]
Grade in 2002	-0.007	0.162*	0.0430	-0.167**
	[0.055]	[0.085]	[0.060]	[0.071]
Age in months since age 14	-0.008	-0.009	0.0089*	0.0005
	[0.006]	[0.0096]	[0.0052]	[0.0122]
Age in months squared (*1000)	0.205**	0.199	-0.0750	0.307
	[0.103]	[0.142]	[0.0541]	[0.192]
Number months between waves	0.0068	-0.0018	0.0077	-0.0790***
	[0.0138]	[0.0171]	[0.0193]	[0.0272]
Mother's grade missing	0.144**	0.385**	0.154***	-0.421***
	[0.069]	[0.157]	[0.049]	[0.050]
Father's grade missing	0.0592	-0.189	-0.296	0.196
	[0.115]	[0.134]	[0.199]	[0.246]
Observations	179	264	97	181

Table shows marginal effects from probit regression, evaluated at sample mean for all variables.
Robust standard errors in brackets, adjusted for sample clustering.
* significant at 10%; ** significant at 5%; *** significant at 1%
Advanced grade=1 if respondent was enrolled in 2003 at a grade higher than their 2002 grade.
Advanced grade=0 if respondent was enrolled in 2003 in a grade at or below their 2002 grade.

Figure 1. Age profiles of sex, smoking, and drinking
CAPS respondents aged 14-22, 2002

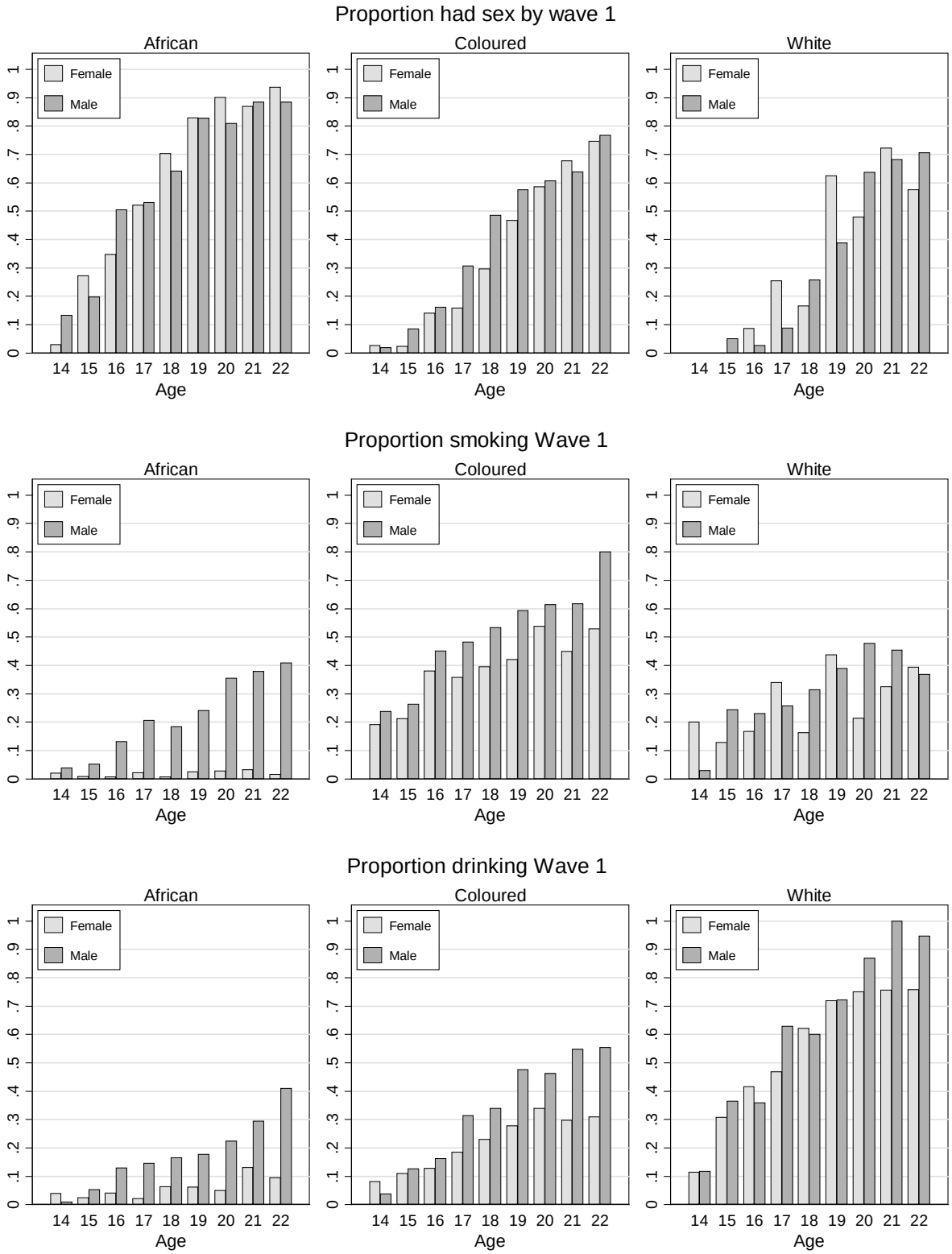
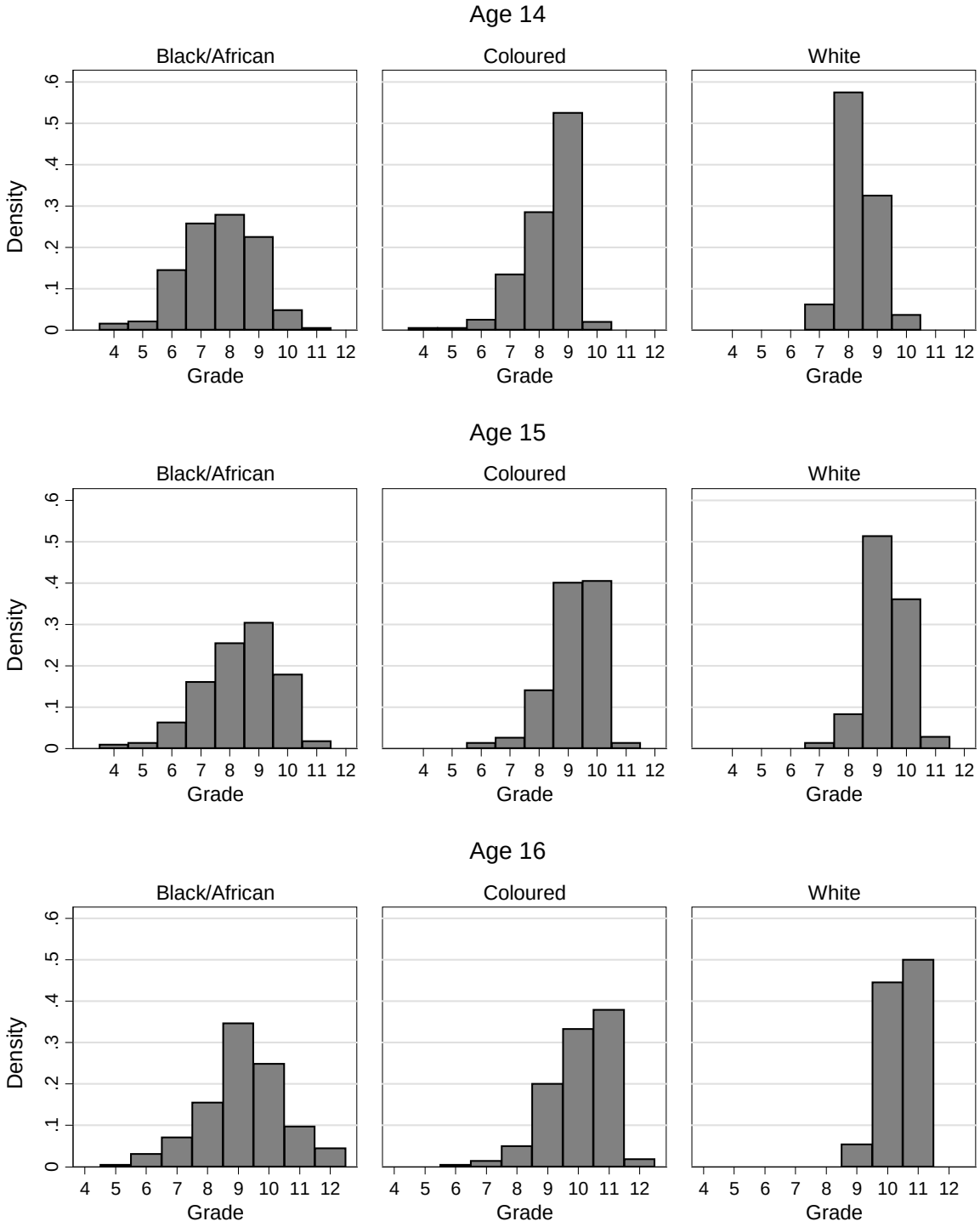


Figure 2. Grade distribution by age

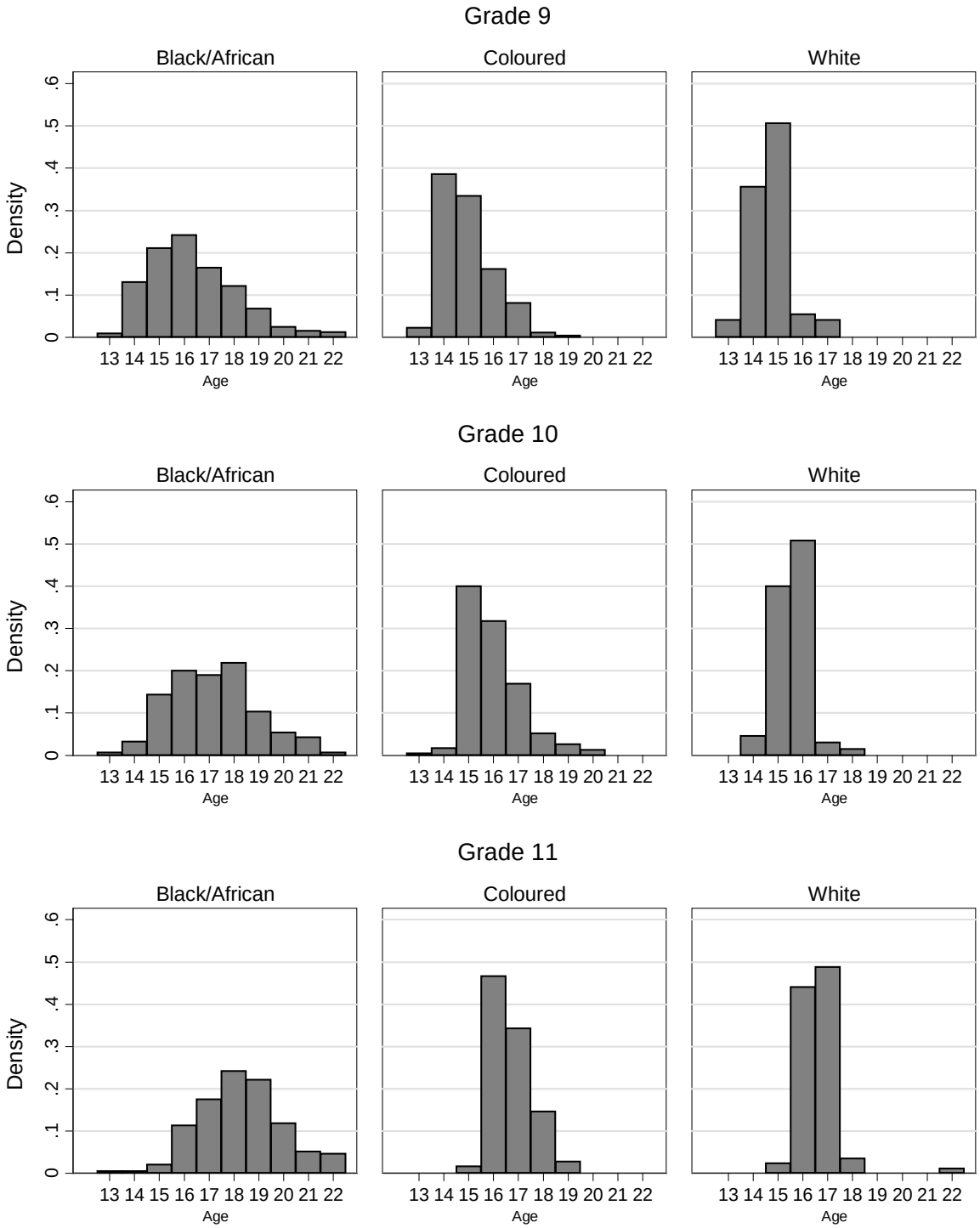
CAPS Wave 1, 2002



Note: Age as of 1 July 2002

Figure 3. Age distribution by grade

CAPS Wave 1, 2002



Note: Age as of 1 July 2002