Do High School Peers Have Persistent Effects on College Attainment and Other Life Outcomes?

Robert Bifulco
Jason Fletcher
Sun Jung Oh
Stephen L. Ross

March, 2014

Do High School Peers Have Persistent Effects on College Attainment and Other Life Outcomes?[1]

Robert Bifulco
Syracuse University

Jason M. Fletcher
University of Wisconsin-Madison

Sun Jung Oh
Syracuse University

Stephen L. Ross
University of Connecticut

Abstract

Using data from the National Longitudinal Study of Adolescent Health, this study examines the impact of high school cohort composition on the educational and labor market outcomes of individuals during their early 20s and again during their late 20s and early 30s. We find that the positive effects of having more high school classmates with a college educated mother on college attendance in the years immediately following high school decline as students reach their later 20s and early 30s, and are not followed by comparable effects on college completion and labor market outcomes. The results suggest that factors that increase college attendance are not always sufficient to improve college graduation rates and longer term outcomes.

Key Words: Education, Peer Effects, Cohort Study

JEL Codes: I21, I24, J15

I. Introduction

Existing research indicates that college enrollment decisions can be influenced by policy relevant factors. For instance, researchers have consistently found that programs that reduce the cost of college increase college enrollments (Abraham and Clark 2006; Cornwell, Mustard and Sridhar 2006; Dynarski 2000; 2003; 2004; Kane 2003; 2007). Recent evidence suggests that the composition of one's high school classmates can also influence college enrollment. Bifulco, Fletcher, and Ross (2011) finds that students whose school cohorts have higher percentages of students with college educated mothers are more likely to attend college; using data from Norway, Black, Devereux, and Salvanes (2010) presents evidence that the gender composition of a school cohort can influence years of schooling; and Billings, Deming and Rockoff (2013) finds that changes in peers arising from the end of school desegregation policies in Charlotte, NC have affected college attendance.[2]

Over the last several decades, however, as college going rates have increased significantly, completion rates among those who enter college have fallen (Belley and Lochner 2007). Thus, it is not clear that increasing college enrollment is sufficient for increasing degree attainment. Descriptive evidence suggests that returns to completing college are substantially higher than the return to merely attending college (Baum, Ma, & Payea 2010), and thus, effects on college completion may be more policy relevant than effects on college enrollment.

Research on whether policy relevant factors can have the same impact on college completion and degree attainment as they have on college enrollment is relatively scarce. Recent studies have examined the effect of scholarship and grant programs on college completion with somewhat mixed results. Dynarski (2008) and Scott-Clayton (2011) find that merit-based scholarship programs increase the likelihood of earning a bachelor's degree, and Castleman and

---

[2] See Ross (2011) for a recent survey on the peer effects literature.

Long (2012) find that a need-based grant program in Florida has increased the likelihood of earning a bachelor's degree. Sjoquist and Winters (2012), however, raise some questions about the robustness of the findings reported in Dynaraki (2008), and in a separate study, Bruce and Carruthers (2011) find little evidence that a merit based aid program in Tennessee has had positive impacts on degree attainment.

Similar studies of the impact of high school cohort or peer group composition on degree completion have not been conducted. The lack of research on this question is in part explained by the difficulty of isolating causal impacts with existing data. Students are not typically assigned to schools randomly, which makes it difficult to isolate the causal effect of high school composition. Following Hoxby (2000), several recent studies have addressed this challenge by using arguably idiosyncratic variation in student composition across cohorts within the same schools to identify the effects of classmate characteristics. These studies have found that variation in the race, gender, ability, exposure to family violence, and home language of classmates can have short term effects on individual test scores (Angist and Lang 2004; Carell and Hoekstra 2010; Friesen and Krauth 2011; Hanushek, Kain, and Rivkin 2002; Hoxby 2000; Lavy, Paserman, and Schlosser 2008; Lavy and Schlosser 2007). This approach, however, requires data on multiple cohorts of students from the same schools, and data sources that have sufficient information on school context for multiple cohorts of students usually do not track students longitudinally following the completion of high school.

In this study, we turn to data that was used in one of the three previous studies that have demonstrated a relationship between high school peers and college attendance. Bifulco, Fletcher, and Ross (2011) use Waves 1 and 3 of the National Longitudinal Study of Adolescent Health (Add Health) to demonstrate that students between one and five years out of high school

are more likely to have attended college if they were exposed to cohorts in high school that had higher than the school average share of students whose mothers are college graduates. The present study further explores the Wave 3 data and uses the recently released Wave 4 of the Add Health, which measures the outcomes of respondents between the ages of 27 and 32.

First we re-examine the responses in Wave 3 in order to see if high school peers lead to greater persistence in college and so higher rates of completing multiple years of college. While this analysis is imperfect because many students in the Wave 3 sample are too young to have completed multiple years of college, we find that students exposed to a high share of students whose mothers are college graduates are more likely to complete one, two, three and even four years of college education by Wave 3 and that these effects are concentrated among males. The magnitude of this effect falls for more years of education, declining by 37 percent between one and four years completed, but so does the baseline percentage of students completing a given number of years of education. As a result, the effect of classmates as a percentage of average completion rates actually increases with the number of years of education considered.

We then turn to Wave 4 to examine college attendance. By Wave 4, a substantially larger fraction of students have attended college than was observed in Wave 3, and we find that respondents ages 27-32 exposed to a large share of classmates with college educated mothers are no more likely to have attended some college as of Wave 4. We demonstrate this finding both descriptively by comparing means for the cohorts with the highest and lowest share of students whose mothers completed college and using regression analyses that control for school fixed effects and trends. We also show that the estimated effect of peers on college attendance in Wave 3 (complete at least one year) and Wave 4 (some college) statistically differ from each other. The effects on Wave 3 and Wave 4 outcomes are estimated using the same sample and so

4

the apparent differences in effects across the waves cannot be attributed to differential attrition between waves. The set of findings that (1) high school classmates influence years of college completed among 19-23 year olds, (2) high school classmates do not influence college attendance among 27-32 year olds, and (3) the number of students completing some college increases substantially between ages 18-23 and ages 27-32, suggest that differences in college attendance are eliminated over time as individuals who might have been discouraged from attending college by their high school classmates have time to re-evaluate their economic opportunities and return to school.

Next, we examine whether or not the composition of one's high school cohort has any longer term effects on college completion and labor market outcomes in Wave 4. The simple differences in means comparison of cohorts with the highest and the lowest share of students whose mothers graduated from college shows no evidence of any positive effect of classmates with college educated mothers on college completion. Regression analysis that controls for school fixed effects and trends identifies positive point estimates of the effect of classmates with college educated mothers for males on completion of two year degrees, four year degrees and employment that are sizable in magnitude, but these point estimates are always insignificant and notably smaller than our estimated effects on college attendance in Wave 3. While we cannot rule out an effect, the data provides no statistical evidence for a long run effect of high school classmates with college educated mothers.

The rest of the paper is organized as follows. Section II provides a brief discussion of why we might not expect short-run effects of cohort composition on college enrollment to persist over time. Section III describes the data and sample used in the analyses. Section IV explains the within-school, across cohort identification strategy that we use and how we implemented it.

It also presents the results of balancing tests that provide empirical support for this identification strategy. Section V presents our empirical results for Waves 3 and 4, and a concluding section summarizes and discusses the implications of our findings.

II. Why Might Short-Run Peer Effects not Persist?

There are several reasons that we might expect the composition of one's high school cohort to influence post-secondary educational outcomes. First, attending school and classes with more educationally advantaged students, e.g. more students with college educated parents, might influence the development of academic skills, attitudes towards school, and college aspirations. Changes in skills, attitudes, and aspirations are likely to change the perceived and real benefits of a college education to an individual. Following Becker (1964), if we view the decision to attend college as an investment choice in which students choose to enroll if the perceived discounted value of the benefits exceeds the costs, increasing the real or perceived benefits of college will increase enrollments. Of course, improved skills and attitudes toward education can also be expected to increase the likelihood of persisting and succeeding in college. Increases in skills and attitudes would also be expected to translate into improved labor market outcomes. Thus, if cohort composition influences college enrollment choices through intermediate impacts on skills and attitudes, we would expect that cohort composition will also influence college completion and other longer term outcomes.

However, there are reasons that cohort composition might influence college enrollment decisions even if it does not influence skills and attitudes. An important strand in the literature on post-secondary attainment has emphasized the importance of the information provided through social networks (Granovetter 1995). Information about opportunities can influence decisions to attend college even if we hold expectations, grades, and cognitive achievement

6

constant. If improved information about post-secondary options both makes an individual more likely to choose to attend college and also improves a student's match with the college he or she chooses, then we might also expect effects on college completion.

Alternatively, if a student has incomplete information about the returns to attending college, one might use the decisions of one's peers as a guide (Bikhchandani, Hirshleifer, and Welch 1992; Manski 1993). As students on the margin of enrolling or not enrolling in college see other students applying for and preparing to attend college, they might be tipped toward enrolling themselves (Fletcher 2012). Thus, peers can influence behavioral choices without influencing attitudes, skills, or information. If cohort composition influences the college enrollment decision in this way, without substantially changing skills, attitudes, or information, then we might expect to see the effect of cohort composition on college enrollment decline as individuals age. As individuals enter their mid-twenties and make decisions about whether or not to return to school, the influence of high school peers is likely to wane, and underlying skills and attitudes are likely to be more telling. Also, if those on the margin of enrolling, whose decisions are influenced by what their peers are doing, are among those least likely to succeed in college, then effects on the enrollment decisions will often not translate into effects on degree attainment. Unlike in the case of policies to reduce the cost of attending college, which might have continuing effects on persistence in college by either reducing students' need to work while they are in enrolled or limiting students accumulation of debt, there is little reason to expect cohort composition to increase persistence in the absence of changes in individual skills and attitudes, or improved matches with the college they choose.

Bifulco, Fletcher, and Ross (2011) present evidence that although the percent of classmates with college educated parents increases the likelihood of having attending college by

one's early 20s, it does not have any effect on individual skills and attitudes during school. That study also documents that children of college educated parents are substantially more likely to attend college themselves, even when compared to other students who have selected into the same high school. This evidence suggests that the effect of the percent of classmates with college educated parents on college attendance is due largely to imitative behavior during or shortly after high school, rather than any change in underlying skills or attitudes. Thus, short-term effects on college attendance may not persist as individuals make decisions about returning to school during their 20s, and any positive effects on degree attainment will be substantially smaller than the short-term effects on college attendance.

III. Data and Samples

The Add Health is a school-based, longitudinal study of the health-related behaviors of adolescents and their outcomes in young adulthood. The study used a clustered sampling design in which first a nationally representative set of high schools, and then a random sample of students from each grade in each school were selected. The survey consists of four waves. Wave 1 was conducted between April 1995 and December 1995. In addition to an extensive in-home survey for the students selected for the longitudinal study, an in-school survey was administered to all of the students attending each sampled school. Wave 2, 3, and 4 consist of follow-up in-home surveys of individuals in the longitudinal sample conducted 1, 7, and 13 years after the Wave 1 survey.

Approximately 20,000 individuals completed the full Wave 1 survey and of these approximately 12,300 responded to the Wave 3 and Wave 4 surveys. Following Bifulco, Fletcher, and Ross (2011), the analyses presented here are conducted using subsamples of these students who, during Wave 1, were in grades 9-12; attended a school serving each of grades 10,

11, and 12; reported themselves as either white, black, Hispanic, or Asian; and had 10 or more students in their school cohort.

The independent variables of interest in the analyses are the percent of students in an individual's school cohort who have a college educated mother and the percentage who are either black or Hispanic. These variables are computed using information from the in-school survey administered to students in the sample schools, and thus are based on a census of students in each cohort. The Wave 3 outcomes examined include whether or not the individual is a high school dropout, whether or not the student has completed one, two or three years of college, and whether or not the individual was idle. High school dropouts are defined as individuals who had not completed 12th grade and were not in high school at the time of Wave 3 survey, and include students who had GEDs. Being idle is defined as neither working nor attending school at the time of the survey. Finally, completing n years of college is set to one if the student completed n years, to zero if the student has not completed n years and is not still in school and to missing if the student is still in school, but has not yet completed n years of college. The reason for setting completion to missing for students who are still in school is that they are still on a path to complete additional years of schooling, but simply might not have had sufficient time to do so. This correction is especially important given that in Wave 3 many students have only been out of high school for one or two years.

Wave 4 outcomes examined include information on high school graduation, idleness plus indicators of whether or not the individual has completed some college, received an associate degree or higher, has received a bachelor degree, and is currently employed at the time of Wave 4. In addition, we examine the log of household income and the log of earnings reported at the time of Wave 4. Respondents were asked to report total household income before taxes in

$5,000 to $50,000 increments up to $150,000 or more. Values were coded at the midpoint of the reported range and at $150,000 for those in the top category.[3] Respondents also reported the dollar amount of income from personal earnings during the most recent year.[4]

Notably, because of differences in the questions asked, the measures of college attendance from Waves 3 and 4 differ. College attendance is measured as having completed one or more years of college in Wave 3 and by having completed "some" college in Wave 4. Later in the paper, we use data from the National Longitudinal Survey of Youth 1997 to examine the number of people in a nationally representative sample who started college and so report having completed some college, but never report having completed one year or more of college. Using this information, we carefully examine the likelihood that this difference in how college attendance is measured between Waves 3 and 4 can account for the fade-out of the effect of cohort composition that we observe.

Table 1 provides the descriptive statistics for each of the cohort composition variables, each of the outcome variables just described, and an additional set of variables that are used in either the regression models that we estimated or the balancing tests conducted (both of which are described below). Means and standard deviations are provided for individuals who meet the primary sampling criteria described above. In Wave 3, the fraction of students completing at least a given number of years of college naturally declines as the number of years considered increases. Our first analysis will be to examine whether the effect of having classmates with college educated mothers persists for completing two, three or four years of college. Comparing Wave 3 and Wave 4 outcome measures shows that the number of dropouts changes little from

---

[3] We also estimated interval regression models coding each interval as its own category and assuming that conditional log income has a normal distribution. The results of these regressions were substantially similar.
[4] In the analysis below we use the log of household income and the log of earnings. When income or earnings were zero the log was coded as 0.

Wave 3 to Wave 4. Because we count GED recipients as dropouts in the analysis, there is little opportunity for students to change their dropout status after Wave 3.[5] The percent of individuals who have ever attended college, however, increases from 58.5 percent in Wave 3 to 70.7 percent in Wave 4, which suggests that a substantial number of people first attended college sometime after Wave 3. One of our questions is whether these delayed choices to attend college lead to a decline of the effect of high school cohort composition. Note also that many who have attended college had not earned a degree—41.8 percent earned an associate degree or higher and 33.3 percent earned a bachelor degree by Wave 4. Another question we address is whether the effect of cohort composition on the decision to delay college has any lasting impact on degree attainment.

IV. Method of Estimating Cohort Composition Effects

The primary challenge in estimating cohort composition effects is the potential correlation between unobserved student characteristics and cohort attributes created by self-selection of individuals into schools (Moffitt 2001). Under the assumption that families do not select schools based on the differences between their child's school specific cohort and the average school composition, a cross-cohort within-school design breaks the correlation between unobserved student characteristics and cohort attributes.

To implement this strategy, we estimate the following regression equation:

$$Y_{ics} = \beta_0 + \beta_1 \%CollEd_{cs} + \beta_2 \%Minority_{cs} + \beta_3 MomEd_{ics} + \beta_4 Ethnicity_{ics}$$
$$+ \delta_c + \phi_s + \gamma_s C + \varepsilon_{ics}$$

---

[5] In coding the Wave 4 dropout and college attendance measures we imposed the restrictions that an individual's education level cannot be less in Wave 4 than in Wave 3 and anyone who reported a GED in Wave 3 is recorded as a dropout in Wave 4.

Here, Y_{ics} is an outcome for student i, who was a member of cohort c in school s; $\%CollEd_{cs}$ is the proportion of students in the individual's school-specific cohort who have a college educated mother; $\%Minority_{cs}$ is the proportion of students in the individual's school cohort who are either black or Hispanic; β_1 and β_2 are the parameters of interest and represent the effect of cohort composition on individual student outcomes; $MomEd_{ics}$ is a set of dummy variables indicating the education level of the student's own mother; $Ethnicity_{ics}$ is an indicator of the student's own race/ethnicity; δ_c is a non-school specific, cohort fixed-effect; ϕ_s is a school fixed effect; C is the cohort variable indicating the student's grade during Wave 1, the effect of which is allowed to vary by school; and ε_{ics} is a random error term.

The inclusion of school fixed effects ensures that the estimation of cohort composition effects is based on comparisons across cohorts within a school. Inclusion of $\gamma_s C$ controls for school specific trends as well. In models that only include school fixed effects, schools that show a systematic trend in cohort attributes are a concern. For instance, parents might observe when the share minority in a school is increasing over time. If the preferences for racial composition differ across families, students from older cohorts who select into a school might differ in systematic, unobserved ways from students in younger cohorts. Also, minorities and students whose parents have less education are more likely to drop out between grades 9 and 12, and also tend to be less motivated to continue their schooling. Thus, compared to younger cohorts the older cohorts in the Add Health will have higher percentages of classmates with a college educated mother and lower percentages of minorities and also students who are generally more motivated, creating a potential correlation between cohort composition and unobserved student characteristics. If we assume, however, that families do not choose schools based on

unanticipated changes in cohort composition, and that the effects of dropouts on cohort composition and student unobservables are approximately linear in grades within schools, then controlling for school trends will eliminate any correlation between cohort composition and unobserved student characteristics.

The results that we report here are from models that include controls for the individual student characteristics related to our cohort variables. Race and parents' education are correlated with several factors that influence outcomes. Thus, even if deviations from school trends in the cohort composition measures are randomly distributed, students in cohorts with higher than predicted percentages of minority students or college educated parents will systematically differ from students in other cohorts. Including individual controls for race and parent education is necessary to prevent these systematic differences from biasing the estimates of cohort composition effects. We also estimated models that include a more extensive set of individual level controls including those listed in Table 1. In principle, if the identification strategy employed here successfully isolates idiosyncratic variation in cohort composition, additional individual level controls are not necessary. And in fact, models with additional controls provide results substantively similar to those reported below.[6]

Because students are nested within schools, we compute Huber/White standard errors that are robust to clustering within schools for all our regressions. Various analyses use different subsamples including those that completed the Waves 1 and 3 surveys, the Wave 1, 3 and 4 surveys, and the Waves 1 and 4 surveys. In each case, the longitudinal sampling weights that correspond to the sample used are provided by the Add Health and are used in the analyses presented.

---

[6] These results are provided in an appendix.

If the identifying assumptions of the model specified above are met, deviations from school specific trends in the cohort composition variables should be uncorrelated with deviations from school specific trends in student background characteristics. To test that this condition is met, we regressed a series of student background characteristics on the cohort composition variables controlling for cohort fixed effects, school fixed effects and trends, the student's race/ethnicity, and the education level of the student's mother. The results of these "balancing tests" are provided in Table 2, which includes the results of 10 separate regressions and a total of 20 separate coefficient estimates. Only one, marginally significant t-statistic is obtained—the estimated relationship between percent black or Hispanic and whether or not the student's parent was born in the U.S is significant at the 0.10 level. One significant t-statistic among 20 coefficient estimates is less than we would expect to emerge by chance. Also, none of the regressions has statistically significant F-statistics. These results provide support for the assumption that the variation in cohort composition used to identify effect estimates is not systematically related to student characteristics.[7]

V. Results

In this section, we present the results from three sets of analyses. The first set examines the effect of cohort composition on college attendance in Wave 3 of the Add Health, when individuals are ages 19-24. These analyses extend those presented in Bifulco, Fletcher and Ross (2011) by looking at effects on completing multiple years of college and estimating effects by gender. The second set of analyses try to determine whether the effects estimated on college attendance observed in Wave 3 persist through Wave 4, when individuals are 27-32 years old.

---

[7] Bifulco, Fletcher and Ross (2011) show balancing test results for the Wave 1 and 3 samples.

The third set of analyses examines impacts on degree completion and labor market outcomes in Wave 4.

a. Effects on Completion of Years of College in Wave 3

Table 3 presents our analysis of college attendance using Wave 3 data. The model estimates capture the likelihood of completing one, two, three, or four years of college as compared to having completed less years of college and not being in school. Positive effects of classmates with college educated mothers arise for all years of college completion, and these effects are concentrated among male students. For instance, a one percentage point increase in share of classmates with college educated mothers implies a 0.962 and a 0.916 percentage point increase in the likelihood of male students completing one or two years of college, respectively. These effects decline considerably, 19 and 36 percent compared to the male one year effect, for completion of three and four years of college, respectively. However, the base completion rates also decline, 35 and 54 percent for three and four years (see Table 1). Thus, as a percentage of the average, the effect of classmate characteristics on completing three and four years of college is actually larger than the effect of completing one year. An alternative model was estimated where college completion is one if the individual has either completed n years of college or is still in school, and zero if the individual is not in school and has not completed the specified number of years of college. The estimates for that specification are very similar to the estimates presented in Table 3.

b. Persistence of Effects from Wave 3 into Wave 4

Prior to presenting the regression analyses for Wave 4, we conduct a simple descriptive analysis that provides a differences-in-differences style assessment of the changes in the effects of peers between Waves 3 and 4. Specifically, for every school, we create two subsamples

where one subsample contains the cohorts of students with the lowest share of college educated mothers in each school and the other subsample contains the cohorts of students with the highest share of college educated mothers. Table 4 presents for each sample the means of the three variables that are observed in both Waves 3 (panel 1) and 4 (panel 2): drop-out of high school, attend college (complete one year or some), and idleness. The last three columns show the differences between the means for each wave, and panel 3 shows the differences in the differences.

The first panel confirms earlier results from Bifulco, Fletcher and Ross (2011) that a higher share of classmates with college educated mothers decreases drop-out and increases college attendance rates (the likelihood that the student completed one or more years of college) and has no effect on idleness in Wave 3. The second panel shows the effects of peers as of Wave 4. The drop-out result is persistent across the two waves, which is expected given that individuals cannot change whether they dropped out of high school between Waves 3 and 4. There is no effect of share of classmates with college educated mothers on college attendance. Although the exact measures of college attendance in the Wave 3 and Wave 4 survey are somewhat different (completing one-year of college v. completing some college), it is noteworthy that the difference between the Wave 3 and 4 estimates for these two very similar outcomes is statistically significant (see panel 3 of Table 4). Peer composition appears to have no effect on idleness in Wave 3 or 4. The second panel also presents differences between the two subsamples for two Wave 4 college attainment variables that are not observed in Wave 3. The last three columns suggest a negative, statistically insignificant effect of share of classmates with college educated mothers on the likelihood that students have completed either two or four year college degrees by Wave 4. Together, the results in Table 3 and 4 are striking. Cohort

16

composition does not show any effect on either attending college or obtaining degrees in Wave 4, despite the fact that the percent of classmates with college educated mothers has positive effects on completing as many as four years of college in Wave 3.

Next, we repeat the regression analyses from Table 3 with Wave 4 data. Table 5 presents the results for the Wave 1, 3 and 4 samples for the three outcomes that are approximately common for both Waves 3 and 4. The estimates presented in the first three columns of the top panel are for Wave 3 outcomes. A one point higher percentage of college educated mothers in one's high school cohort is associated with 0.325 percentage point lower high school dropout rate and a 0.343 percentage point higher rate of college attendance.

The first three columns of the middle and bottom panels of Table 5 indicate that these effects of cohort composition on Wave 3 outcomes are driven almost entirely by effects on males. Among males, a one percentage point increase in the percent of students whose parents are college educated is associated with a 0.552 percentage point decrease in the likelihood of dropping out and a 0.752 percentage point increase in the likelihood of attending college. These effects on males are significant considering the growing gap between males and females in educational attainment, particularly among less advantage groups (Buchman and Diprete 2006; Goldin, Katz and Kuziemko 2006; Heckman and La Fontaine 2007). We also see that a higher percentage of black and Hispanic students in one's high school cohort is associated with substantially lower rates of college attendance and higher rates of idleness among males, and the estimate on idleness is statistically significant. These effects of percent black and Hispanic, however, are offset by effects in the opposite direction for females.

Counter to expectations, the percent of high school classmates with a college educated mother is associated with higher rates of idleness and the percent black or Hispanic is associated

with higher rates of college attendance among females. The counter intuitive result on idleness might be explained by childrearing. Because of household responsibilities, women with young children are especially likely to be neither attending school nor employed outside the home. When we recode the idleness variable such that women with children under the age of 5 are not counted as idle, the effect of the percent of college educated mothers in the high school cohort becomes smaller and statistically insignificant.[8] The positive effect of the share of minority students on college attendance among females is perhaps more puzzling, but is not the focus of this paper.[9]

The next three columns of Table 5 present the estimated effects of cohort composition on the same three outcomes, only now measured at the time of Wave 4. The sample used to compute the estimated effects on outcomes measured in Wave 4 is exactly the same as that used to compute the estimated effect on outcomes measured in Wave 3. Thus, any difference between Wave 3 and Wave 4 cannot be attributed to sample changes. The last three columns of Table 5 present the differences between the effect estimates in Wave 3 and those estimated in Wave 4. We conducted F-tests of the hypothesis that the differences between the effect estimates in Wave 3 and Wave 4 equal zero, and the probabilities of obtaining the observed differences if the true difference were zero are reported in brackets.

Again, the estimated effect of the percent college educated mothers on the dropout for Wave 4 is very close to that obtained from Wave 3. Thus, as expected, we see little change in the effect of percent with a college educated mother on the dropout variable. For the college attendance variable, however, there is strong evidence that the effects of the percent of

---

[8] Results available from the authors upon request.

[9] We hypothesized that this effect was for community college attendance, but using the only information available, completion of two and four year degrees, we found no evidence that the effect of share minority students on college attendance is driven by students who were more likely to obtain a two-year degree than a four year, as indicated by either maternal education or a predictive model using a broad set of demographic variables.

classmates with a college educated mother does not persist into Wave 4. The estimated effect of percent with a college educated mother on college attendance by Wave 4 not only is statistically insignificant, but the point estimate is about one-third the size of the estimated effect on college attendance by Wave 3. We see a similar difference between the estimated effect on college attendance measured in Wave 3 and Wave 4 among males.

Similarly, the effect of the percent black or Hispanic in the high school cohort on idleness among males declines in Wave 4—a statistically significant effect estimate of 0.446 in Wave 3 is very close to zero in Wave 4. Among females the statistically significant effect of percent of classmates with a college educated mother on idleness during Wave 3 also becomes much smaller and statistically insignificant in Wave 4. The one result that persists from Wave 3 to Wave 4 is the positive association between the share of students in the high school cohort who are black or Hispanic and college attendance among females.

c. *Is the Fadeout of College Attendance Effects due to Differences in Measures?*

We interpret the results in Table 5 as evidence that the cohort composition effects observed in Wave 3 have been undone overtime as individuals who did not attend college immediately after high school due to more disadvantaged peers enter college at a later date. One threat to this interpretation is that our Wave 3 college attendance variable captures completing one year of college while our Wave 4 college attendance variable captures completing some college. Some individuals with some college might not have completed one year, and it is possible that the effect of high school composition in Wave 3 is caused by cohort composition influencing the likelihood of completing a year of college among students who had started college. If so, this effect would not be captured by the Wave 4 question, even if the effect on completing one year persisted.

To assess the likelihood that differences between the Wave 3 and 4 effect estimates are due to differences in the measures of college attendance used, we need to two pieces of information. First, we need an estimate of how many people who report some college have not completed one year. Since information on who completed one year of college is not available in Wave 4 of the Add-Health, we turn to another nationally representative sample of youth from a similar time period, the National Longitudinal Survey of Youth 1997 (NLSY97). Specifically, using the NLSY97, we select a sample of individuals who had completed between 8 and 11 years of education in 1997 (high school students) from the full sample of 12 to 16 year olds as of December 31, 1996, who also report educational outcomes in 2004 and in 2010 (which is the analogue of sample of individuals that appear in Wave 1, 3 and 4 of the Add-Health). Out of this sample of 3435 students, 359 (10.4 percent) reported having attended some college, but did not report completing one year of college in 2010 (the analogue of Wave 4). These 359 students represent 28.3 percent of students who report some college, but did not complete a two or four year college degree..

Next, we must determine which students in the Add Health who report some college are most likely to not have completed one year of college, and these resulting probabilities must reflect the correlation between not completing one year and cohort peer composition that arises from the peer effects. Therefore, we need an estimate of how much of the effect of cohort composition on the likelihood of completing one year of college by Wave 3 is due to increasing the likelihood of attending college at all, and how much is due to increasing the likelihood of completing a year of college conditional on having enrolled in college. One extreme assumption is that the entire estimated peer effect on completing one year of college in Wave 3 is due to

increasing the likelihood of completing a year of college conditional on having enrolled and that none of it is due to increasing the likelihood of attending college.

To determine how much of the difference between estimated effects on college attendance in Wave 3 and Wave 4 might be due differences in the measures of college attendance under this extreme assumption, we did the following. (1) We took the subset of individuals in Waves 1, 3, and 4 of the Add-Health sample who report some college but did not complete a two or four year degree by Wave 4. This subsample represents the set of people who might have some college but less than one year. (2) Based on the percentage computed using the NLSY97, we determined that 28.29 percent of the individuals in this subset of our Add-Health sample, 561 individuals, need to be a assigned a value indicating that they had not completed one year of college by Wave 4. (3) Using the subset of our Add-Health sample with some college but no degree, we reestimated the model of completing one year of college attendance in Wave 3 as a function of our cohort composition variables, individual race and maternal education variables, school fixed effects, school trends, and cohort fixed effects.[10] We then used the estimates from this model to predict for each individual the likelihood of not completing one year of college conditional on having completed some college, and rescaled these probabilities so that they summed to 561. (4) Using these probabilities, we randomly assigned individuals to have the value of zero on the Wave 4 college attendance variable. In essence, we imputed new values of the Wave 4 college attendance variable for this those individuals who had some college but no degree in order to eliminate the influence of those unknown individuals who did not complete one year. (5) Using the entire Wave 1,3, and 4 sample, we reestimated the effect of cohort composition on college attendance, using the new, modified measure of college

---

[10] This model is identical to the college attendance model used throughout the paper except estimated on a different sample.

attendance as the dependent variable. (6) We repeated the random assignment and estimation, steps (4) and (5), 1000 times.

The results of these reestimations indicate how much of the difference in the estimated effect of cohort composition between Waves 3 and 4 could be due to difference in measurement if the entire peer effect operates through a greater likelihood of completion with no affect on attendance. Recall, from Table 5 that the estimated effect of the percent of college educated mothers in the high school cohort on college attendance in Wave 3 is 0.343 and the estimated effect on college attendance in Wave 4 is 0.121. If the results of the reestimations tend to be closer to the Wave 3 effect, then that would indicate that the apparent change in effect between Wave 3 and 4 could be due to differences in the Wave 3 and Wave 4 measures of college attendance. If, on the other hand, the results of the reestimations tend to be closer to the Wave 4 estimates, then that suggests that the effect of cohort composition on college attendance really did attenuate between Waves 3 and 4. Because the reassignment of the college attendance variable is based on the extreme assumption concerning the effect of cohort composition, the simulation maximizes the correlation between cohort composition and failing to complete one year, and the results of the reestimation maximize the difference between the reestimated effects of cohort composition and estimated effect of cohort composition on Wave 4 college attendance reported in Table 5 (0.121). In that sense, the results of these reestimations indicate something like an upper bound on how much of the difference between the Wave 3 and Wave 4 estimates might be due to the differences in measures of college attendance used.

Across the 1000 reestimations, the average estimated effect of the percent of students with college educated mothers in the high school cohort was 0.198 and the average standard error was 0.230. Thus, even under the extreme assumption about how cohort composition

22

influences college attendance, the estimated effect of cohort composition on college attendance in Wave 4 is still, on average, 42 percent smaller than the estimated effect in Wave 3, statistically indistinguishable from zero, and relatively close to the actual wave 4 estimate. These results indicate that even under the extreme assumption, only 1/3 of the difference in the estimated effect of cohort composition between Wave 3 and Wave 4 is due to the difference in the measure of college attendance used. In addition, in only 13 percent of the 1000 reestimations was the difference between the Wave 3 and Wave 4 estimates 0 or less. Given these results, along with the results reported in Table 3, we conclude that most of the difference between the Wave 3 and Wave 4 estimates is due to a real attenuation of the effect of cohort composition on college attendance, and that it is unlikely that the entire difference is due to difference in measures.

d. Effects on degree attainment and labor market outcomes in Wave 4

Even though the effect of classmates with a college educated mother on college attendance falls over time, particularly among males, differences in cohort composition might have longer lasting effects on college attainment and labor market outcomes. Delaying college, for instance, might make it less likely that one will obtain a degree (Bozick and DeLuca 2005; Niu and Tienda 2011; Rowan-Kenyon 2007), and might also influence income and earnings by an individual in his or her late 20s and early 30s. Table 6 presents results that address this issue.

Considering the results for the sample as a whole, males and females together, we do not find significant relationships between the high school cohort composition variables and either degree attainment or labor market outcomes. Nor do we find any statistically significant relationships when we focus solely on males. In the sample of males, several of the point estimates suggest substantial effects of cohort composition, but the estimates have large standard errors. For instance, the estimated effect of the percent of classmates on male students with a

23

college educated mother on attaining an associate's or bachelor's degree is 0.345 and 0.355, respectively. An effect of this size implies that a one half standard deviation (7.1 percentage point) increase in the percent of classmates with college educated parents increases the likelihood of attaining either degree by 2.5 percentage points. A one half standard deviation increase in the percent of classmates with college educated parents increases the likelihood of being employed for males by 2.2 percentage points. Although arguably important substantively, these effects are substantially smaller than the estimated effects on college attendance in Wave 3 of 5.3 percentage points, and similar to the estimated effects on college attendance in Wave 4. Further, unlike the wave 3 results, there is little difference between the peer effect estimates for males and the estimates for the entire sample.

The estimated effects on income and earnings are generally small and statistically insignificant. For instance, the point estimate of the effect of percent of mothers with a college education on the log of family income is 0.463, one of the larger coefficients in the last two columns of Table 6. This effect implies that a one half standard deviation increase in the percent of college educated mothers is associated with a 3.3 percent increase in income, an effect which is not nearly statistically significant. We also estimated alternatives to the earnings models where we tested whether high school classmates affected the likelihood of earning over a specific threshold, e.g. likelihood of falling in the 1st or 2nd income quartile, and again found no effect of peers.[11]

In an attempt to increase precision, we reestimated the regression presented in Table 6 using the sample of individuals who responded to Waves 1 and 4 of the survey regardless of whether or not they responded to Wave 3. This addition increased the sample size by more than 20 percent. The results are reported in Table 7. The effects of expanding the sample on standard

---

[11] These results are provided in an appendix.

error sizes are modest. The point estimates computed using this expanded sample are generally smaller, particularly among males. Among other results, the marginally significant effect (10 percent level) of the percent black or Hispanic on the likelihood of earning a bachelor's degree among males is substantially smaller than the corresponding estimate in Table 6, and not nearly statistically significant. Given the imprecision of the effect estimates by gender, the sensitivity of the point estimates to choice of sample, and the negative effects in the descriptive analysis above, it is difficult to draw strong conclusions about the impacts of cohort composition on degree attainment and employment outcomes, but the evidence of any lasting impacts is weak.

VI. Conclusion

The purpose of this analysis has been to determine whether the effects of high school cohort composition on college attendance persist as individuals age from their early twenties into their later twenties and early thirties. We also examine if cohort composition variables have any longer lasting effects on degree attainment and labor market outcomes. We find that the cohort composition effects on college attendance and idleness do not persist into later life. The evidence on degree attainment and employment is less conclusive. Point estimates for the effect of classmates with a college educated mother on degree attainment and employment among males are similar in magnitude to the insignificant long run impacts on college attendance, and arguably large, but the estimates are statistically insignificant and become smaller when the sample is expanded. Further, examinations of simple differences across cohorts find no positive effect of classmates with college educated mothers on educational attainment. Thus, we conclude that the Add Health contains little if any evidence to support the existence of long lasting impacts of our measures of peer composition on these specific outcomes.

The decline or disappearance of initial effects on post-secondary outcomes is consistent with the evidence Bifulco, Fletcher and Ross (2011) present that the percent of classmates with college educated parents has no effect on individual skills and attitudes during school. Combined with this earlier evidence, the decline of cohort composition effects suggests that the initial effect on college attendance is due largely to imitative behavior during or shortly after high school. As contact with the school cohort becomes less frequent, and individual attitudes and skills become more telling, the effect on educational enrollment falls. The effect of peers on completion for up to four years of college among 19-23 year olds and the large increase in the number of students completing some college by ages 27-32 are consistent with differences in attendance being eliminated over time as individuals who might have been discouraged from attending college by their peers have time to re-evaluate their economic opportunities and return to school. Also absent changes in underlying skills and attitudes, or improved college-student matches, short-run impacts of cohort composition on college enrollment do not translate into discernible effects on college completion or labor market outcomes.

It is important to note that the estimates based on within-school variation might underestimate the effect of school composition on student outcomes. Some of the mechanisms through which the composition of a school might influence outcomes are constant across cohorts. For instance, a school's ability to garner resources or teacher expectations for students might be influenced as much by other cohorts in the school as a student's own cohort. Also, the effect of school composition on skill and attitude development might depend on the amount of integration of students from different backgrounds that occurs within schools. Finally, the fact that the effects of having more classmates with college educated parents do not persist as students age does not imply that the effects of other aspects of cohort composition fade over time. For

instance, using data from Norway, Black, Devereaux, and Salvannes (2010) find that the percent female in the cohort influences educational attainment, and those effects might in fact be longer lasting. Nonetheless, the evidence presented here suggests that the initial influence of having more educationally advantaged high school classmates on post-secondary outcomes in the U.S. are short-lived.

REFERENCES

Abraham, Katharine, and Melissa Clark. 2006. Financial aid and students' college decisions: Evidence from the District of Columbia Tuition Assistance Grant Program." *Journal of Human Resources* 41(3): 578-610.

Angrist, Joshua, and Kevin Lang. 2004. "Does school integration generate peer effects? Evidence from Boston's Metco Program." *American Economic Review* 94(5): 1613-34.

Baum, Sandy, Jennifer Ma, and Kathleen Payea. 2010. *Education pays 2010: The benefits of higher education for individuals and society.* New York: The College Record.

Becker, Gary S. 1964. *Human capital: A theoretical and empirical analysis, with special reference to education.* Chicago: University of Chicago Press.

Belley, Phillipe, and Lance Lochner. 2007. The changing role of family income and ability in determining educational achievement. *Journal of Human Capital* 1(1): 37-89.

Bifulco, Robert, Jason Fletcher, and Stephen L. Ross. 2011. The effect of classmate characteristics on post-secondary outcomes: Evidence from the Add Health." *American Economic Journal: Economic Policy* 3(1): 25-53.

Bikchandani, Sushil, David Hirshleifer, and Ivo Welch. 1992. A theory of fads, fashion, custom, and cultural change in informational cascades. *Journal of Political Economy* 100(5): 992-1026.

Black, Sandra E., Paul J. Devereux and Kjell G. Salvanes. 2010. Under pressure? The effects of peers on outcomes of young adults. Working Paper no. 16004, National Bureau of Economic Research, Cambridge, MA.

Bozick, Robert and Stephanie DeLuca. 2005. Better late than never? Delayed enrollment in the high school to college transition. *Social Forces* 84(1): 527-50.

Bruce, Donald J. and Celeste K. Carruthers. 2011. Jackpot? The impact of lottery scholarships on enrollment and earnings in Tennessee. Unpublished manuscript, Department of Economics, University of Tennessee, Knoxville.

Buchmann, Claudia and Thomas DiPrete. 2006. The growing female advantage in college completion: The role of family background and academic achievement. *American Sociology Review* 71(4): 515-41.

Carrell, Scott E. and Mark L. Hoekstra, 2010. Externalities in the classroom: How children exposed to domestic violence affect everyone's kids, *American Economic Journal: Applied Economics* 2: 211-28.

Castleman, Benjamin L. and Bridget T. Long. 2012. Looking beyond enrollment: The causal effect of need-based grants on college access, persistence, and graduation. Paper presented at NBER Education Program Meeting. Accessed at http://conference.nber.org/confer//2012/EDf12/Castleman_Long.pdf .

Cornwell, Christopher, David Mustard, and Deepa Sridhar. 2006. The enrollment effects of merit-based financial aid: Evidence from Georgia's HOPE Scholarship. *Journal of Labor Economics* 24(4): 761-86.

Dynarski, Susan. 2000. Hope for whom? Financial aid for the middle class and its impact on college attendance. *National Tax Journal* 53(3): 629-61.

Dynarski, Susan. 2003. Does aid matter? Measuring the effect of student aid on college attendance and completion. *American Economic Review* 93(1): 279-88.

Dynarksi, Susan. 2004. Who benefits from the college savings incentives? Income, educational expectations and the value of the 529 and Coverdell. *National Tax Journal* 57(2): 359-83.

Dynarski, Susan. 2008. Building the stock of college-educated labor. *Journal of Human Resources* 43(3): 576-610.

Fletcher, Jason F. 2012. "Similarity in peer college preferences: New evidence from Texas." *Social Science Research* 41: 321–30.

Friesen, Jane and Brian Krauth. 2011. Ethnic enclaves in classrooms. *Labour Economics* 18(5): 656-63.

Goldin, Claudia, Lawrence Katz, and Ilyana Kuziemko. 2006. The homecoming of American college women: The reversal of the college gender gap." *Journal of Economic Perspectives* 20(4): 133-56.

Granovetter, Mark. 1995. *Getting a job: A study of contacts and careers.* Chicago: University of Chicago Press.

Hanushek, Eric A., John F. Kain, and Steven G. Rivkin. 2002. New evidence about *Brown v. Board of Education*: The complex effects of racial composition on achievement. Working Paper no. 8741, National Bureau of Economic Research, Cambridge, MA.

Heckman, James J. and Paul A. LaFontaine. 2007. The American high school graduation rate: Trends and levels. Working Paper no. 13670, National Bureau of Economic Research, Cambridge, MA.

Hoxby, Caroline M. 2000. Peer effects in the classroom: Learning from gender and race variation. Working Paper no. 7867, National Bureau of Economic Research, Cambridge, MA.

Kane, Thomas J. 2003. A quasi-experimental estimate of the impact of financial aid on college-going. Working Paper no. 5164, National Bureau of Economic Research, Cambridge, MA.

Kane, Thomas J. 2004. Evaluating the impact of the D.C. Tuition Assistance Grant Program. Working Paper no. 10658, National Bureau of Economic Research, Cambridge, MA.

Killingsworth, Mark R. and James J. Heckman. 1986. Female labor supply: A survey. In *Handbook of labor economics*, ed. Orley Ashenfelter and R. Layard, 103-204. Amsterdam: North Holland.

Kling, Jeffrey R., Jeffrey B. Liebman, Lawrence F. Katz. 2007. Experimental analysis of neighborhood effects. *Econometrica* 75: 83-119.

Lavy, Victor, M. Daniele Paserman, and Analia Schlosser. 2008. Inside the black-box of peer ability effects: Evidence from variation in high and low achievers in the classroom. Working Paper no. 14415, National Bureau of Economic Research, Cambridge, MA.

Lavy, Victor and Analia Schlosser. 2007. Mechanisms and impacts of gender peer effects at school. Working Paper no. 13292, National Bureau of Economic Research, Cambridge, MA.

Manski, Charles. 1993. Adolescent econometricians: How do youth infer the returns to schooling. In *Studies of supply and demand in higher education*, eds. Charles T. Clotfelter and Michael Rothschild, 43-60. Chicago: University of Chicago Press.

McDaniel, Anne C., Thomas A. McDaniel, Claudia B. Buchman, and Uri Shwed. 2011. The black gender gap in educational attainment: Historical trends and racial comparisons." *Demography* 48(3): 889-914.

Moffitt, Robert. 2001. Policy interventions, low-level equilibria and social interactions. In *Social dynamics*, eds. Steven N. Durlauf and H. Peyton Young, 45-82. Cambridge, Mass.: MIT Press.

Niu, Sunny X, and Marta Tienda. 2011. Delayed enrollment and college plans: Is there a postponement penalty. Unpublished paper. Princeton University.

Ross, Stephen L. 2011. Social interactions within cities: Neighborhood environments and peer relationships. In *Handbook of Urban Economics and Planning* (Eds. N. Brooks, K. Donaghy, G. Knapp). Oxford University Press.

Rowan-Kenyon, H. T. 2007. Predictors of delayed college enrollment and the impact of socioeconomic status. *The Journal of Higher Education* 78(2): 188-214.

Scott-Clayton, Judith. 2011. On money and motivation: A quasi-experimental analysis of financial incentives for college achievement." *Journal of Human Resources* 46(3): 614-46.

Sjoquist, David J. and John V. Winters. 2012. Building the stock of college-educated labor revisted. *Journal of Human Resources* 47(1): 270-85.

Table 1: Sample Descriptives, Wave 1, 3 & 4 Sample

	N	Mean	SD
Wave 3 outcome variables			
Drop out of high school	8023	0.133	0.340
Complete one year of college or more	7707	0.585	0.493
Complete two years of college or more	7461	0.495	0.500
Complete three years of college or more	6970	0.378	0.485
Complete four years of college or more	6140	0.270	0.444
Idle	7747	0.131	0.337
Wave 4 outcome variables			
Drop out of high school	8024	0.125	0.331
Attend some college	7706	0.707	0.455
Associate degree	8024	0.418	0.493
Bachelor degree	8024	0.333	0.471
Employed	8023	0.833	0.373
Idle	7746	0.124	0.330
Log of household income	7566	10.83	0.809
Log of earnings	7885	9.59	2.66
Cohort variables			
Proportion with college educated mother	8025	0.288	0.140
Proportion black or Hispanic	8025	0.304	0.294
Control Variables			
Black	8025	0.164	0.370
Hispanic	8025	0.117	0.321
Asian	8025	0.043	0.204
Mother high school drop-out	8025	0.159	0.366
Mother high school graduate	8025	0.362	0.481
Mother some college	8025	0.237	0.425
Mother college graduate	8025	0.241	0.428
Grade 10 indicator	8025	0.257	0.437
Grade 11 indicator	8025	0.235	0.424
Grade 12 indicator	8025	0.257	0.437
Male	8025	0.503	0.500
Age	8025	29.53	1.31
Parent's Age	8025	42.59	5.87
Parent native born	8025	0.877	0.297
Parent years in US	8025	35.96	12.96
Parent information missing	8025	0.328	0.470
Log family income	8025	10.44	1.19
Single parent	8025	0.264	0.406
Live with both parents	8025	0.576	0.457
Number older siblings	8025	0.825	1.15

Notes: Cohort variables are calculated for each grade surveyed in each high school using the full in-school Wave 1 sample. Percent black or Hispanic is based and mother's education variable is based on student report. All parent variables measured using Wave 1. Wave 1, 3 & 4 longitudinal weights are used.

Table 2: Balancing tests for cohort composition measures, Wave 1, 3 & 4 Sample

Dependent Variable	% with college educated mother	% black or Hispanic	F-statistic	p-value
Male	-0.131 (0.229)	0.352 (0.261)	0.997	0.381
Age (in years)	0.111 (0.308)	-0.046 (0.340)	0.076	0.927
Parent's age (in years)	2.845 (2.025)	-4.735 (4.351)	1.338	0.269
Parent born in the U.S.	-0.047 (0.098)	-0.182* (0.103)	1.690	0.192
Missing parent information	-0.136 (0.256)	0.275 (0.310)	0.626	0.537
Log of family income	0.265 (0.597)	0.526 (0.621)	0.626	0.538
Single parent family	-0.069 (0.176)	0.048 (0.216)	0.137	0.872
Live w/both biological parents	0.394 (0.274)	-0.240 (0.246)	2.086	0.131
Number of older siblings	-0.386 (0.493)	-0.173 (0.615)	0.371	0.691
Parent alcoholism reported	-0.208 (0.189)	0.056 (0.149)	0.837	0.437

The figures in each row are coefficients from regressions that include, in addition to the cohort composition measures, controls for cohort fixed effects, school fixed effects, school trends, the student's race, and the student's mother's level of education. All variables are measured using Wave 1 of the Add Health. Figures in parentheses are standard errors robust to clustering at school level. The F-statistics is for the joint effect of percent black or Hispanic and percent with college educated mother. All models estimated using Wave 1, 3 & 4 longitudinal weights. * designates significantly different from zero at 0.10.

Table 3: Continuation of Education, Wave 1 & 3 Sample

Dependent Variable	% with college educated mother	% black or Hispanic	Observations	R-square
		Males & Females		
Complete One Year	0.547**	0.072	9,045	0.219
	(0.231)	(0.295)		
Complete Two Years	0.569***	0.075	8,755	0.224
	(0.181)	(0.289)		
Complete Three Years	0.466**	0.111	8,190	0.233
	(0.198)	(0.302)		
Complete Four Years	0.423*	0.071	7,241	0.275
	(0.232)	(0.329)		
		Males		
Complete One Year	0.962***	-0.392	4,345	0.267
	(0.268)	(0.471)		
Complete Two Years	0.916***	-0.178	4,211	0.264
	(0.309)	(0.464)		
Complete Three Years	0.769**	0.281	3,952	0.262
	(0.293)	(0.439)		
Complete Four Years	0.602*	0.196	3,528	0.302
	(0.307)	(0.401)		
		Females		
Complete One Year	0.103	0.545*	4,700	0.224
	(0.381)	(0.306)		
Complete Two Years	0.191	0.375	4,544	0.237
	(0.235)	(0.261)		
Complete Three Years	0.167	0.116	4,238	0.262
	(0.253)	(0.282)		
Complete Four Years	0.167	0.115	3,713	0.307
	(0.245)	(0.308)		

Table 4: Difference-in-Differences: High versus Low Share Cohorts

Educational Outcome	High Percent Maternal College Cohort			Low Percent Maternal College Cohort			Difference		
	Sample	Mean	Std Err	Sample	Mean	Std Err	Difference	Std Err	T-Stat
Wave 3									
Drop out of high school	2141	0.121	0.007	2087	0.153	0.008	-0.032	0.011	3.03
Attend college (At least 1 yr)	2054	0.599	0.011	1999	0.572	0.011	0.027	0.015	1.74
Idleness	2059	0.139	0.008	2015	0.136	0.008	0.003	0.011	0.28
Wave 4									
Drop out of high school	2141	0.113	0.007	2087	0.142	0.008	-0.029	0.010	2.83
Attend college (Some)	2054	0.707	0.010	1999	0.705	0.010	0.002	0.014	0.14
Idleness	2058	0.126	0.007	2015	0.128	0.007	-0.002	0.010	0.19
Associate degree	2141	0.414	0.011	2087	0.427	0.011	-0.013	0.015	0.86
Bachelor degree	2141	0.324	0.010	2087	0.348	0.010	-0.024	0.015	1.65
Difference									
Drop out of high school		0.008	0.002		0.011	0.002	-0.003	0.003	1.01
Attend college		-0.108	0.007		-0.133	0.008	0.025	0.010	2.44
Idleness		0.013	0.009		0.008	0.010	0.005	0.013	0.37

Table 5: Estimated impact of cohort composition on Wave 3 outcomes

	Wave 1, 3 & 4 Sample Wave 3 Outcomes			Wave 1, 3 & 4 Sample Wave 4 Outcomes			Differences		
Cohort Composition	Dropout	Attend College	Idle	Dropout	Attend College	Idle	Dropout	Attend College	Idle
				Males & Females					
% college educated mother	-0.325**	0.343*	0.218	-0.309**	0.121	-0.060	-0.016	0.222*	0.278*
	(0.146)	(0.203)	(0.146)	(0.145)	(0.258)	(0.156)	[0.734]	[0.079]	[0.089]
% black or Hispanic	0.012	0.228	0.074	0.027	0.370	0.004	-0.015	-0.142	0.070
	(0.187)	(0.278)	(0.155)	(0.195)	(0.223)	(0.170)	[0.936]	[0.320]	[0.700])
Observations	8,023	7,707	7,747	8,024	7,706	7,746			
				Males					
% college educated mother	-0.552**	0.752**	0.017	-0.523**	0.346	-0.248	-0.029	0.406**	0.265
	(0.241)	(0.360)	(0.223)	(0.238)	(0.328)	(0.201)	[0.603]	[0.026]	[0.226]
% black or Hispanic	-0.089	-0.181	0.446**	-0.170	-0.043	-0.028	0.081	-0.138	0.474*
	(0.315)	(0.466)	(0.192)	(0.306)	(0.322)	(0.210)	[0.187]	[0.503]	[0.053]
Observations	3,707	3,575	3,592	3,707	3,574	3,592			
				Females					
% college educated mother	-0.177	-0.149	0.390**	-0.178	-0.095	0.129	0.001	-0.054	0.261
	(0.154)	(0.362)	(0.182)	(0.150)	(0.408)	(0.262)	[0.840]	[0.753]	[0.265]
% black or Hispanic	0.066	0.623*	-0.157	0.176	0.726**	0.141	-0.110*	-0.103	-0.298
	(0.195)	(0.313)	(0.271)	(0.197)	(0.308)	(0.274)	[0.062]	[0.597]	[0.262]
Observations	4,316	4,132	4,155	4,317	4,132	4,154			

All regressions include both cohort composition variables along with controls for cohort fixed effects, school fixed effects, and school trends as well as the individual student covariates related to the cohort variables. Estimates for the Wave 1 & 3 sample are computed using Wave 1 & 3 longitudinal sampling weights, and estimates of Wave 1, 3 & 4 sample are computed using Wave 1, 3 & 4 longitudinal sampling weights. Figures in parentheses are standard errors robust to clustering at the school level. The statistical significance of the difference is evaluated using an F-test and the reported significance is shown in brackets. *designates significantly different from zero at 0.10, ** significantly different than zero at 0.05 level.

Table 6: Estimated impacts of cohort composition on Wave 4 outcomes, Waves 1, 3 & 4 Sample

Cohort Composition	Associate Degree	Bachelor Degree	Employed	Log of Household Income	Log of Earnings
			Males & Females		
% college educated mother	0.312	0.214	0.013	0.463	0.296
	(0.194)	(0.178)	(0.165)	(0.319)	(1.169)
% black or Hispanic	0.175	0.307	0.083	0.068	-0.405
	(0.209)	(0.206)	(0.181)	(0.313)	(1.266)
Observations	8,024	8,024	8,023	7,566	7,885
			Males		
% college educated mother	0.345	0.355	0.305	0.325	-0.177
	(0.287)	(0.285)	(0.212)	(0.533)	(1.226)
% black or Hispanic	0.274	0.575*	0.130	-0.550	-2.222
	(0.267)	(0.323)	(0.240)	(0.518)	(1.398)
Observations	3,707	3,707	3,707	3,490	3,644
			Females		
% college educated mother	0.259	0.085	-0.268	0.370	0.904
	(0.279)	(0.240)	(0.245)	(0.711)	(1.779)
% black or Hispanic	0.157	0.168	-0.006	0.627	0.221
	(0.306)	(0.269)	(0.269)	(0.476)	(2.114)
Observations	4,317	4,317	4,316	4,076	4,241

All regressions include both cohort composition variables along with controls for cohort fixed effects, school fixed effects, and school trends as well as the individual student covariates related to the cohort variables. Dependent variables are measured using Wave 4 of the Add Health. Figures in parentheses are standard errors robust to clustering at the school level. Estimates compute using Wave 1, 3 & 4 longitudinal sampling weights. * designates significantly different from zero at 0.10, ** significantly different than zero at 0.05 level.

**Table 7: Estimated impacts of cohort composition on Wave 4 outcomes,
Waves 1 & 4 Sample**

Cohort Composition	Associate Degree	Bachelor Degree	Employed	Log of Household Income	Log of Earnings
			Males & Females		
% college educated mother	0.140	0.100	0.012	0.528*	0.499
	(0.181)	(0.172)	(0.151)	(0.280)	(1.195)
% black or Hispanic	-0.009	0.152	0.138	-0.008	0.088
	(0.178)	(0.222)	(0.138)	(0.361)	(1.323)
Observations	9,757	9,757	9,755	9,178	9,556
			Males		
% college educated mother	-0.035	0.079	0.199	0.651	-0.527
	(0.297)	(0.259)	(0.184)	(0.472)	(1.148)
% black or Hispanic	0.037	0.225	0.121	-0.172	-0.215
	(0.296)	(0.362)	(0.198)	(0.535)	(1.998)
Observations	4,604	4,604	4,603	4,314	4,502
			Females		
% college educated mother	0.303	0.180	-0.223	0.338	1.555
	(0.263)	(0.227)	(0.239)	(0.606)	(1.907)
% black or Hispanic	-0.038	0.073	0.141	0.284	0.384
	(0.258)	(0.251)	(0.202)	(0.515)	(1.803)
Observations	5,153	5,153	5,152	4,864	5,054

All regressions include both cohort composition variables along with controls for cohort fixed effects, school fixed effects, and school trends as well as the individual student covariates related to the cohort variables. Dependent variables are measured using Wave 4 of the Add Health. Figures in parentheses are standard errors robust to clustering at the school level. Estimates compute using Wave 1 & 4 longitudinal sampling weights. * designates significantly different from zero at 0.10, ** significantly different than zero at 0.05 level.

**Appendix 1: Estimated impacts of cohort composition on Wave 4 Outcomes
with and without individual control variables, Wave 1, 3 & 4 Sample**

Cohort Composition	Baseline controls (Males & Females)	Baseline + extended controls (Males & Females)	Baseline controls (Males)	Baseline + extended controls (Males)	Baseline controls (Females)	Baseline + extended controls (Females)
Dropout						
% college educated mother	-0.309**	-0.282**	-0.523**	-0.473**	-0.178	-0.177
	(0.145)	(0.127)	(0.238)	(0.226)	(0.150)	(0.138)
% black or Hispanic	0.027	0.020	-0.170	-0.144	0.176	0.160
	(0.195)	(0.179)	(0.306)	(0.294)	(0.197)	(0.176)
Attend College						
% college educated mother	0.121	0.021	0.346	0.156	-0.095	-0.150
	(0.258)	(0.228)	(0.328)	(0.311)	(0.408)	(0.393)
% black or Hispanic	0.370	0.443**	-0.043	0.140	0.726**	0.689**
	(0.223)	(0.212)	(0.322)	(0.310)	(0.308)	(0.282)
Idle						
% college educated mother	-0.060	-0.046	-0.248	-0.232	0.129	0.127
	(0.156)	(0.147)	(0.201)	(0.198)	(0.262)	(0.252)
% black or Hispanic	0.004	0.034	-0.028	0.011	0.141	0.159
	(0.170)	(0.172)	(0.210)	(0.208)	(0.274)	(0.275)
Associate Degree						
% college educated mother	0.312	0.225	0.345	0.231	0.259	0.247
	(0.194)	(0.195)	(0.287)	(0.297)	(0.279)	(0.267)
% black or Hispanic	0.175	0.233	0.274	0.401	0.157	0.099
	(0.209)	(0.208)	(0.267)	(0.244)	(0.306)	(0.282)
Bachelor Degree						
% college educated mother	0.214	0.141	0.355	0.240	0.085	0.080
	(0.178)	(0.187)	(0.285)	(0.299)	(0.240)	(0.223)
% black or Hispanic	0.307	0.345*	0.575*	0.714**	0.168	0.080
	(0.206)	(0.205)	(0.323)	(0.290)	(0.269)	(0.255)
Employed						
% college educated mother	0.013	0.001	0.305	0.290	-0.268	-0.252
	(0.165)	(0.163)	(0.212)	(0.217)	(0.245)	(0.239)
% black or Hispanic	0.083	0.054	0.130	0.089	-0.006	-0.015
	(0.181)	(0.189)	(0.240)	(0.237)	(0.269)	(0.274)
Log of Household Income						
% college educated mother	0.463	0.414	0.325	0.132	0.370	0.393
	(0.319)	(0.296)	(0.533)	(0.563)	(0.711)	(0.713)
% black or Hispanic	0.068	-0.053	-0.550	-0.603	0.627	0.549
	(0.313)	(0.302)	(0.518)	(0.513)	(0.476)	(0.484)
Log of Earnings						
% college educated mother	0.296	0.188	-0.177	-0.472	0.904	0.769
	(1.169)	(1.202)	(1.226)	(1.282)	(1.779)	(1.859)
% black or Hispanic	-0.405	-0.820	-2.222	-2.200	0.221	-0.100
	(1.266)	(1.255)	(1.398)	(1.371)	(2.114)	(2.116)

All regressions include both cohort composition variables along with controls for cohort fixed effects, school fixed effects, and school trends as well as the individual student covariates related to the cohort variables. Extended controls include the individual characteristics listed in Table 2. Dependent variables are measured using Wave 4 of the Add Health. Figures in parentheses are standard errors robust to clustering at the school level. Estimates compute using Wave 1, 3 & 4 longitudinal sampling weights. * designates significantly different from zero at 0.10, ** significantly different than zero at 0.05 level.

Appendix 2: Estimated impacts of cohort composition on household income and earnings by individuals above the 25th, 50th, and 75th percentile of each distribution, Wave 1, 3 & 4 Sample

Cohort Composition	Log of Household Income			Log of Earnings		
	25th*	50th*	75th*	25th*	50th*	75th*
Males & Females						
% college educated mother	0.207	0.134	0.116	-0.091	0.151	0.054
	(0.211)	(0.207)	(0.187)	(0.176)	(0.218)	(0.187)
% black or Hispanic	-0.049	0.111	0.086	-0.133	-0.157	0.134
	(0.180)	(0.219)	(0.206)	(0.196)	(0.217)	(0.167)
Observations	7,566	7,566	7,566	7,885	7,885	7,855
Males						
% college educated mother	0.037	0.368	0.528	-0.214	-0.214	-0.141
	(0.288)	(0.450)	(0.351)	(0.413)	(0.420)	(0.288)
% black or Hispanic	-0.185	-0.096	-0.187	-0.360	-0.548*	0.143
	(0.222)	(0.394)	(0.333)	(0.266)	(0.311)	(0.305)
Observations	3,490	3,490	3,490	3,644	3,644	3,644
Females						
% college educated mother	0.204	-0.256	-0.316	0.014	0.456	0.186
	(0.420)	(0.398)	(0.258)	(0.282)	(0.314)	(0.256)
% black or Hispanic	0.019	0.320	0.386	-0.050	0.084	0.009
	(0.274)	(0.288)	(0.299)	(0.348)	(0.295)	(0.197)
Observations	4,076	4,076	4,076	4,241	4,241	4,241

All regressions include both cohort composition variables along with controls for cohort fixed effects, school fixed effects, and school trends as well as the individual student covariates related to the cohort variables. Extended controls include the individual characteristics listed in Table 2. Dependent variables are measured using Wave 4 of the Add Health. Figures in parentheses are standard errors robust to clustering at the school level. Estimates compute using Wave 1, 3 & 4 longitudinal sampling weights. * designates significantly different from zero at 0.10.

*Dependent variables are 1 if an individual's household income or earnings above the 25th, 50th, or 75th percentile in sample, 0 otherwise.