# A Quality-Preserving Increase in Four-Year College Attendance: Evidence from NLS-72 and ELS:2002

Robert B. Archibald, David H. Feldman, and Peter McHenry

College of William and Mary

## Abstract

A common criticism of attempts to increase college going in the United States is that marginal students will be ill-prepared and likely to fail without much gain. This argument assumes that the process that matches students to colleges is efficient in the sense that the most capable high school graduates go to college and the least capable do not. We argue using the NLS-72 and ELS:2002 data sets that the matching process is not efficient, although it has improved over time. As college attendance rates increased substantially between 1972 and 2004, average preparedness of four-year college-goers did not decrease. Most of the increase in four-year college attendance over time came from high-achieving (well-prepared) high school students attending at higher rates. Attendance rate increases at two-year colleges were more evenly spread across the spectrum of high school achievement, which is appropriate since two-year colleges offer the highest returns for lower-achieving students. We use multinomial logit models to demonstrate that a measure of likely success (GPA) became more predictive of college attendance over time, while other student characteristics such as race and parents' education became less predictive. These findings are consistent with a movement toward a more efficient matching process between students and colleges, and we note that further improvements remain possible.

**Keywords**: higher education, human capital, college attendance

**JEL codes**: I21, I23, J1, J24

## I. Introduction

There are many who argue that the U.S. needs to expand college going and college completion to reverse our slide down the international rankings of college completion rates. At the same time there are those who argue that too many people are already going to college. They argue that sending more people to college will wind up sending students who are ill prepared for the rigors of higher education. By this argument diminishing returns have set in. Expansions of the student body will lead to meager increases in degree attainment. Baum, Kurose, and McPherson (2013) put it this way: "As a larger share of young people completes high school and a larger share of those completers participates in postsecondary education, the academic preparation of college students who are enrolling at the margin will be lower than average."

The diminishing returns argument presumes that the process that links students to colleges and universities is based on a strict meritocracy. In such a system those currently going to college would be those most able to benefit, and expanding the national student body would require admitting students who are less meritorious. This argument falls apart if the process linking students to college is not a strict meritocracy. In fact, it is not. The process is not a simple one. Students have to choose to start at a two-year school or a four-year school. Seventeen- and eighteen-year-olds play a significant role in these decisions, and they cannot all be counted on to behave sensibly. Attending college requires a considerable up-front expenditure as well as significant foregone income, which may make it difficult for some students with considerable ability. Liquidity constraints combined with fear of borrowing may hold back other high-ability potential college students. Also, the process of applying to college and financial aid is complex, and many talented students have parents who do not know how to help them cut through the complexity.

Recent research has highlighted the haphazard nature of matches made between students and colleges. Some of this literature focuses on the students at the top of the ability distribution and finds that many of them do not link up with the selective colleges to which they could easily gain admittance (e.g., Hoxby and Avery, 2012; Dillon and Smith, 2013). The mismatch problem concerns more than just selection among colleges: some very talented high school students don't attend college at all (Bowen, Chingos, and McPherson, 2009; Roderick, Coca, and Nagaoka, 2001). In addition, some students do not do a very good job choosing between two-year and four-year institutions (Brand, Pfeffer, and Goldrick-Rab, 2012). All of these margins are potentially important. To have good matches we need to have the right students going to the college that best suits them. If, as this research suggests, we do not always have good matches, it is possible for the number of students attending colleges to expand without a reduction in student quality. The trick is to do a better job of sorting students into their best option.

In this paper we investigate the efficacy of the system that matches college students with colleges. Our analysis uses two nationally representative surveys to study the changes in college matching over the 32 years between two high school classes. The data sets are the National Longitudinal Study of the High School Class of 1972 (NLS-72) and the Education Longitudinal Study of 2002 (ELS:2002), which describes the high school class of 2004. We find that the matching process has improved considerably. First, over the time span covered by these two surveys, attendance at four-year colleges has expanded and there has been no decrease in the average quality of the student body in these schools. In other words, there has been no evidence of diminishing returns in four-year institutions. Second, attendance at two-year colleges has also expanded, but in this case the average quality of the student body has decreased. The findings of Brand, Pfeffer, and Goldrick-Rab (2012) provide some justification for our claim that this

change was efficiency-enhancing; they note that "while enrolling in community college expands the likelihood of bachelor's degree completion for students *who otherwise would not have attended college at all*, it penalizes students *who would otherwise have attended four-year schools*, especially highly selective four-year schools" (page 27). In terms of matching efficiency, it appears beneficial to send higher-achieving students to four-year colleges and lower-achieving students to two-year colleges. This is precisely what happened between 1972 and 2004.

Third, we use multinomial logit models of college choice to demonstrate that an important measure of likely college success (GPA) became more predictive of college attendance over time. At the same time, other student characteristics such as race and parents' education became less predictive of college attendance. These findings are consistent with a movement toward a more efficient matching process between students and colleges.

The paper follows in 6 additional sections. Section II provides the evidence for our claim that the college matching process has improved over time. The remainder of the paper presents our explanation for this finding. Section III describes the data we will use in more detail. Section IV presents the analysis of four-year and two-year college attendance rates from the two surveys. Section V presents the results of a multinomial logit analysis of the choice of attending a two-year college, a four-year college, or not attending college. Section VI analyzes the changes in the marginal effects between the two surveys. Section VII presents our summary and conclusions.

## II. Increasing Classes and Improving Matches

Over time the United States has been rather successful in linking more students to colleges and universities. In the NLS-72 data 52.3 percent of the high school seniors in 1972

attended a college or university within the first three semesters following high school graduation, 18.9 percent at a two-year college and 33.4 at a four-year college. The comparable figures in the ELS:2002 data for the high school seniors in 2004 are 73.4 percent overall, 26.4 percent at a two-year college and 47 percent at a four-year college. The increases, 7.5 percentage points at two-year colleges and 13.6 percentage points at four-year colleges, are quite large. In what follows we will demonstrate that these increases were not uniformly spread across the ability distribution.

We will focus on how college going has changed across the distribution of ability measured by high school grade point averages (GPA). We chose to use GPA as opposed to test scores because many studies have found that high school grades are much better predictors of college graduation than our test scores. Bowen, Chingos and McPherson (2009) present convincing evidence on this point. They also present the likely reason for the result: "they [grades] reveal qualities of motivation, and perseverance – as well as the presence of good study habits and time management skills—that tell us a great deal about the chances that a student will complete a college program" (page 124).

To describe the matching process between students and colleges, we divided students in each data set into 10 deciles based on their high school grade point average. Figure 1 presents these data for four-year colleges. NLS-72 attendance rates are dark bars, and ELS:2002 attendance rates are gray bars. For both surveys, the four-year college going rate was clearly related to high school grade point average: students with higher GPAs go to college more frequently than students with lower GPAs. Looking first at the NLS-72 results, there is a lot of room in the top half of the GPA distribution for the four-year college going rates to increase. And the ELS:2002 results (the gray bars) fill in a lot of this room. The vast majority, 81 percent,

of the increase in participation between the two surveys resulted from increased participation from students in the top half of the high school GPA distribution. Half of the increase in four-year college participation was the result of increases in the college going rates of the top three deciles, and only 2.2 percent was a result of increases for the bottom three deciles. The increase in college going rates between 1972 and 2004 resulted from both increasing participation of students who should have been well prepared for college (high grades) and from those whose preparation is questionable (low grades), but the first group was dramatically bigger than the second.[1]

The increases in college going rates for the students from the top of the GPA distribution make it more difficult to improve four-year college attendance in this same way in the future, but the figure reveals that we are far from the limiting case. The college going rate for students in the top ten percent of high school GPAs is 86.7 percent, and it is 79.2 percent, 68.1 percent, and 60.1 percent for students in the next three deciles. This shows that there is still room for increases in overall college going rates without having to slip down the GPA distribution. For example, if we could achieve 90 percent attendance rates for the top four GPA deciles, then the overall rate would increase from 47 percent to 52.1 percent.

Figure 2 presents analogous results for two-year college attendance. The bars for two-year college attendance exhibit a much lower variance and do not follow the same pattern as the bars in the figure for four-year schools. Most importantly, the bigger increases in two-year college attendance are concentrated in the lower GPA deciles. Increases in the bottom half of the GPA distribution amounted to 82.6 percent of the increase in two-year attendance between high

[1] Comparing students based on their scores on standardized tests in math and reading yields similar results: most of the college going increases occurred for students in the top of the test score distribution. Results are available upon request.

school classes of 1972 and 2004, and a little more than half of this increase, 51 percent, came from students in the bottom three deciles of the high school GPA distribution.

These two figures demonstrate that college attendance has shifted considerably in the 32 years between the high school class of 1972 and the high school class of 2004. Higher ability students are more likely to go to four-year colleges and students of lesser ability are more likely to go to two-year colleges. This suggests that the college matching process has improved. In what follows we expand the measures of ability to include scores on mathematics tests and reading tests that were given to the respondents in the two samples and look at what has happened to average scores over the 32 years between entry of the two groups into higher education.

Table 1 gives the average percentiles of high school GPA, scores on a mathematics test, and scores on a reading test given to the respondents to the two surveys. Survey participants were given different tests, but since we are interested in the relative position of the students rather than their absolute scores, this is not a concern. Similarly, because we measure grades using percentiles, the increase in grades over time between the NLS-72 and the ELS:2002 (grade inflation) is not a concern.

Focus first on columns 1 and 2 of Table 1. Each student has a percentile (1 to 100) of the GPA distribution across seniors in his or her cohort (NLS-72 or ELS:2002). In the full sample (top panel), the averages of these percentiles are near 50, as expected. This is true for math and reading test scores as well. The averages are not closer to 50 because of sampling error and attrition in the sample (college attendance is only known for respondents to the follow-up survey, so we drop from our sample people who left prior to that survey).

The middle panel shows averages of GPA and test score percentiles for students who did not attend four-year college. They are significantly lower than the overall averages and similar between the NLS-72 and ELS:2002 cohorts. The third panel shows the same calculations for students who attended two-year colleges. The levels of high school success of those who attended a two-year college after high school graduation in 1972 are very close to the overall averages, but this is not so for those who attended a two-year college after high school graduation in 2004. The 2004 students are considerably less well prepared for college than the average student.

The final panel gives the results for students who attended a four-year college. Their levels of high school success are higher than average, as expected, since the highest-performing high school students probably have the highest returns from college (and have revealed through their high school performance a preference for attending college). Importantly, the GPAs and test scores of college-goers in the NLS-72 and ELS:2002 are very similar. The average GPA and reading test score percentiles are slightly higher in the ELS:2002 (contrary to the diminishing returns hypothesis), while the average math test score percentile was slightly lower. Even though four-year college attendance increased substantially between the two time periods (by 13.6 percentage points), the average preparation of four-year college attenders did not fall. This implies that the marginal college-goer was probably *not* less prepared in 2004 than in 1972.

Columns 3 and 4 of Table 2 repeat the exercise with a different calculation of GPA and test score percentiles. It appears that attrition between the survey of seniors and the survey two years later induced positive selection on GPA and test scores, in particular in the ELS:2002 (see the top panel of column 2, where average percentiles in the two-years-later sample are somewhat higher than 50). This would mechanically increase the measured college preparation of college-

goers over time (and support our result mistakenly). So for columns 3 and 4, we calculate GPA and test score distributions in the sample of respondents to the follow-up survey two years after expected high school graduation (dropping attriters). Essentially, this nets out the attrition effect on average GPA percentiles. The averages of percentiles in the ELS:2002 are closer to 50 (top panel of column 4). Nevertheless, the bottom panel (columns 3 and 4) shows again that average college preparation of four-year college attenders did *not* decline between 1972 and 2004, despite the large increases in attendance rates.

In an analysis of 1972 and 1992 high school senior cohorts (NLS-72 and NELS:88), Bound, Lovenheim, and Turner (2010) find that "Math percentiles remained constant or increased in all four year [college] sectors across the two surveys" (page 137).[2] Our finding is similar to theirs in that a measure of college preparation did not decrease over time, despite large attendance gains. Our end point is 12 years later than theirs, and we actually find small reductions in math test score percentiles, the measure used in Bound et al. (2010). However, as we argued above, we think that the results for GPA are more telling, and the results for reading test scores also show increases. In any event, the result that four-year college going increased and student quality did not drop seems to hold.[3]

## III. Data

In this section we describe the data in more detail. Our data come from the NLS-72 and ELS:2002 data sets at the U.S. Department of Education. Both include longitudinal data on cohorts of high school students. We use them to analyze a representative sample of high school

[2] Four-year college sectors are top 50 public, non-top 50 public, less selective private, and highly selective top fifty.

[3] Carneiro and Lee (2011) argue that college student quality has fallen over time in the U.S. Their direct evidence focuses on attenders of any college, without distinguishing between two-year and four-year institutions. Our results imply that the quality decline in Carneiro and Lee (2011) is a feature of two-year college students rather than four-year college students.

seniors in 1972 (NLS-72) and a representative sample of high school seniors in 2004 (ELS:2002). Our intent is to investigate changes in the determinants of college going, so we focused on variables that were available in both surveys.

Our choice of variables is guided by similar studies that utilized multiple data sets to study college attendance. Ellwood and Kane (2000) use data from the High School and Beyond high school class of 1980 and the NELS:88 high school class of 1992. Belley and Lochner (2007) and Lovenheim and Reynolds (2011) use the National Longitudinal Survey of Youth (NLSY79 and NLSY97). The sets of variables in these studies were similar in several respects. All of these studies included variables for sex, race, parents' education, income, and student ability. Interestingly, Ellwood and Kane (2000) is the only other study that uses data on GPA as an indicator of ability.

With three exceptions, the variables we utilize are measured in the same fashion in our two surveys. The first exception is family income, which in each survey is measured using a series of indicator variables for nominal income at the time of the survey. To obtain comparable data we first adjusted boundaries for the family income variables using the Consumer Price Index so that the boundaries for the NLS-72 variables were expressed in 2002 dollars. Unfortunately, the results yielded only a few cases in which the NLS-72 boundaries lined up with the ELS:2002 boundaries. We are left with only four income categories: family income less than $25,000, family income between $25,000 and $50,000, family income between $50,000 and $75,000, and family income greater than $75,000 (all in 2002 dollars). These are fewer categories than we would have preferred.[4] The second exception is test scores. The two surveys

[4] In the NLS-72, our lowest family income category includes those reporting less than $6,000 of family income in 1972 nominal dollars. The subsequent categories include those reporting between $6,000 and $12,000, between $12,000 and $18,000, and more than $18,000. Family

gave students different aptitude tests. In both surveys students were given tests of mathematics and reading. We express test scores as percentiles in the sample's distribution (separately for NLS-72 and ELS:2002 samples) so that the units would be the same for both surveys. NLS-72 respondents took the aptitude tests in the 12th grade, while ELS:2002 respondents took the aptitude tests in the 10th grade.

The third variable measured somewhat differently between surveys is the high school grade point average. Both come from surveys of high school administrators, rather than the students themselves, and they refer to grades as of 12th grade. The raw NLS-72 data are in different formats including letter grades (e.g., B+), numbers on a 100-point scale, and numbers on a 4-point scale. We convert reported GPAs to a consistent 4-point scale. The ELS:2002 GPA variable is also on a 4-point scale and measures the student's average grade across all academic courses.

Table 2 presents summary statistics for the variables included in our analysis. We present data for the full sample and subsamples divided by high school GPA. Throughout the paper, our calculations use sampling weights that make sample statistics representative of high school senior classes in 1972 and 2004. Use of weights explains why the averages of subsample averages are not exactly equal to corresponding full sample averages. Focusing first on the full sample, we see some clear differences between the two nationally representative samples. The racial composition of the NLS-72 is very different from the racial composition of the ELS:2002 sample. There are fewer whites, more blacks, dramatically more Hispanics, and more Asians as

---

income data in the NLS-72 come from the student questionnaire. When the data are missing, we impute family income based on parents' occupations and annual salaries by occupation in the 1970 U.S. Census.

shares of the ELS:2002 sample. These changes reflect changes in the racial and ethnic composition of the United States.

Parents in the ELS:2002 survey have higher levels of education.[5] There are fewer high school dropouts, and more parents with a BA or an MA or higher. This reflects increases in college going nationally. If the NLS-72 seniors in high school were 18 years old, they were born in 1954. That means their parents attended college in the 1940s and early 1950s when college going rates were very low. On the other hand, in the ELS:2002 data, 18 year old seniors in high school in 2004 were born in 1986, so their parents went to college in the late 1970s or early1980s when college going was much more common. The two samples also reflect the migration patterns one would expect. Larger fractions of the ELS:2002 are from the South and West than in the NLS-72. Also, there appears to have been a movement from schools located in towns to schools located in the suburbs.

The data in the two halves of the GPA distribution are interesting as well (columns 3 through 6 of Table 2). Females are overrepresented in the top half of the GPA distribution in both surveys, and this overrepresentation increases over time. While females represent about half of the full sample, the female share in the top-half of the GPA distribution is .578 in the NLS-72 and .598 in the ELS:2002. Whites and Asians are overrepresented in the top half of the GPA distribution in both surveys. Parents with college degrees are overrepresented in the top half of the GPA distribution, particularly parents with graduate degrees in the ELS:2002 sample, whose share increases from .174 of the full sample to .241 in the top half of the GPA sample. Students from families in the highest income group in the ELS:2002 survey are overrepresented

[5] The parents' education variable takes the value of the higher level of education between the respondent's mother and father.

in the top-half of the GPA distribution, increasing from a .289 share in the full sample to a share of .364 in the top half of the GPA distribution.

## IV. College Attendance Over Time

In this section, we compare college attendance rates from our two surveys by characteristics of the sample. We measure college attendance from survey questions about whether the respondent was enrolled in school in October after normal high school graduation and October of the following year (October 1972 and October 1973 in NLS-72; October 2004 and October 2005 in ELS:2002). We categorize a respondent as attending a two-year college if she did so in the first year, or if she did not attend college in the first year but attended a two-year college in the following year. We categorize a respondent as attending a four-year college if she did so in the first year, or if she did not attend college in the first year but attended a four-year college in the following year. We categorize a respondent as not attending college if she reported attending neither a two-year nor a four-year college in both survey questions. Table 3 presents the two-year and four-year college attendance rates for the variables summarized in Table 2 for the two samples and the splits of the samples by high school GPA.

These data demonstrate that the findings from Figure 1 and Figure 2 are the results of changes across many types of students. For all types of students in the top half of the GPA distribution (columns 9 and 10), the four-year attendance rates for the ELS:2002 are larger than the four-year attendance rates for NLS-72. On the other hand, two-year college attendance increased for all categories of students in the bottom half of the GPA distribution (columns 5 and 6). In the other two cases, four-year attendance for the bottom half of the GPA distribution and two-year attendance for the top-half of the distribution, the results are mixed. While it increased for most categories, four-year attendance for students in the bottom half of the GPA distribution

decreased for Asian students, students whose parents had some postsecondary education, students whose parents' highest degree is a BA, students from the suburbs, and students from families with incomes below $25,000 (columns 11 and 12). Two-year attendance for students in the top half of the GPA distribution also increased for most categories, but it decreased for males, Asians, Other Race, students whose parents' highest degree is an MA or higher, students from the West, and students from families with incomes above $75,000 (columns 3 and 4).

The results for males and females are striking. In the NLS-72, males were more likely to go to a two-year college by 4.8 percentage points (21.3 percent for males and 16.5 percent for females) and more likely to go to a four-year college than females by 2.7 percentage points (34.8 percent for males compared to 32.1 percent for females). This is reversed in the ELS:2002, where females are .9 percentage points more likely to go to a two-year college (26.9 percent for females compared to 26 percent for males) and 6.1 percentage points more likely to go to a four-year college than males (50.3 percent for females compared to 44.2 percent for males). These results are accentuated when we look at the top half of the GPA distribution, where attendance at two-year colleges increased 5.3 percentage points for females compared to a decrease of .3 percentage points for males, and attendance at four-year colleges increased 25.1 percentage points for females compared to 14.3 percentage points for males. The large increases in female college attendance are consistent with findings in Goldin, Katz, and Kuziemko (2006), who use Census data to show that female college graduation rates surpassed male college graduation rates during the time period between our two surveys.

Attendance at two-year colleges increased for all of the racial categories except Asians. The decrease in Asians' attendance at two-year colleges is driven by Asian students in the top half of the GPA distribution. Attendance at four-year colleges increased for all of the racial

categories, and with one exception, the increase is a result of increases in both halves of the GPA distribution. The exception is the fall in the four-year college participation rate of Asian students in the bottom half of the GPA distribution (from 28 percent in the NLS-72 to 21 percent in the ELS:2002). One interesting finding in the table is that the four-year attendance rate for blacks in the bottom half of the GPA distribution in the ELS:2002 survey, 31 percent, is higher than the attendance rate for any other racial/ethnic category in the bottom half of the GPA distribution.

Students with more-educated parents are generally less likely to attend two-year colleges and more likely to attend four-year colleges. However, there have been increases in first-generation college students at both types of colleges between the two surveys. The differences in college attendance rates, both two-year and four-year, in the two surveys are larger for students whose parents were high school dropouts or high school graduates than for students whose parents had some postsecondary education or a bachelor's degree. Breaking the pattern, the biggest increase in four-year college attendance came from students whose parents had an MA or higher.

The results for the regions are consistent with what we know about college attendance patterns. Students from the West in both data sets are more likely to attend two-year colleges and much less likely to attend a four-year college than students in the other three regions. The prevalence of two-year colleges in the West particularly in the earlier survey is responsible for this result. Also, attendance at two-year schools expanded more in rural areas than in urban areas. The increase in two-year attendance was 4.9 percentage points in cities, 7.1 percentage points in the suburbs, 9.9 percentage points in towns, and 12.2 percentage points in rural areas. There is a dramatic increase in four-year attendance rates for students who went to a high school in a city: 17.8 percentage points (from 33.2 percent to 51.1 percent) compared to an overall

increase of 13.6 percentage points. As is the case for all of the changes in four-year attendance rates, the majority of the increase comes from the upper half of the GPA distribution, but the increase in the bottom half of the GPA distribution is unusually large for students who attended high school in a city.

The results for income show that the students from the poorest families had large increases in two-year college attendance rates and only small increases in four-year college attendance rates. The overall increase for the poorest families was 14.7 percentage points for two-year attendance and only 4.0 percentage points for four-year attendance. The increases in two-year attendance decline as income increases, and the increases in four-year attendance generally increase as income increases.

## V. Multinomial Logit Estimates

Tables 4, 5, and 6 present average marginal effects from multinomial logit models of college attendance where the choice is among three options: attend no college, attend a two-year college, or attend a four-year college. The tables give separate estimates for the full sample, the top half of the GPA distribution, and the bottom half of the GPA distribution. We computed the standard errors for the changes in the marginal effects under the assumption that the estimates using the NLS-72 data and the ELS:2002 data are independently distributed.

**Two-year college attendance –** Columns 1 and 3 in the tables present the association between a one-unit increase in each independent variable on the predicted probability of two-year college attendance. In the NLS-72, the marginal effect for Female is negative and statistically significant, and the estimated marginal effects are very similar in the two halves of the GPA distribution (from Tables 5 and 6). In the ELS:2002, the marginal effect for female is again statistically significant in the overall results (Table 4), but the sign has changed from

negative to positive. Also, the marginal effects of female are quite different in the two halves of the GPA distribution. The marginal effect of being female is very small and not statistically significant in the top half of the GPA distribution (Table 5), but it is much larger and statistically significant in the bottom half of the GPA distribution (Table 6).

The indicator variable for Black is the only statistically significant racial category (relative to non-Hispanic whites). In the overall results and the results for the top half of the GPA distribution, the marginal effect of being Black is negative and statistically significant. In the bottom half of the GPA distribution the marginal effect for Black is still negative, but it is smaller in absolute value and not statistically significant.

There are only two statistically reliable results for the variables measuring parental education. First, in the NLS-72 the marginal effects for parents with some postsecondary education are positive and statistically significant for the overall sample and the top half of the GPA distribution. Second, again in the NLS-72 the marginal effect for parents with an MA or higher is negative and statistically significant for students in the top half of the GPA distribution.

West is the omitted region. All of the regional variables are negative and statistically significant in all six estimates. Two-year schools were much more prevalent in the West. U.S. Department of Education data show that fall 1972 enrollment in 2-year schools was 20.55 percent of total enrollment in the Northeast, 23.31 percent in the Midwest, 24.24 percent in the South, and 47.16 percent in the West. Fall 2004 enrollment saw a continued but diminished advantage for two-year enrollment in the West. Two-year enrollment in 2004 was 33.87 percent

in the Northeast, 24.67 percent in the Midwest, 37.78 percent in the South, and 48.59 percent in the West.[6]

The omitted category for the urbanization variables is City school. These variables include few statistically significant marginal effects for the NLS-72 data. The ELS:2002 results include positive and statistically significant results for Suburban schools, town schools, and rural schools. These results are stronger for the top half of the GPA distribution than they are for the bottom half of the GPA distribution.

The results for GPA and test scores are interesting. The marginal effect of GPA percentile is negative in the overall sample and for the top half of the GPA distribution. The effects are not strong. The marginal effect of a 10 point increase in GPA percentile reduces the probability of going to a two-year school by less than a percentage point in the NLS-72 results and by 1.5 percentage point in the ELS:2002 results. The results are much stronger for students in the top half of the GPA distribution, but they change sign and lose statistical significance for the bottom half of the GPA distribution. The marginal effects of the percentile on the Math test vary considerably. In the overall results the measured effect is positive and quite small in the NLS-72 and negative and again quite small in the ELS:2002. These effects are negative in both surveys for students in the top half of the GPA distribution and positive for students in the bottom half of the GPA distribution. The results for percentiles on the Reading test follow the same pattern, negative overall and in the top half of the GPA distribution and positive in the bottom half of the GPA distribution. These results are not statistically significant as often as are the results for the Math percentile.

[6] The more recent data come from the *2005 Digest of Education Statistics,* and the earlier data come from *Historical Trends: State Education Facts 1969 to 1989*, both published by the National Center for Education Statistics.

The lowest income families are the omitted category for the family income variables. Income does not seem to have a strong effect on two-year attendance. In the overall results for the NLS-72, students from families in the two top income categories are only roughly 3.5 percentage points more likely to go to a two-year college than are students from the lowest income category. This effect disappears in the ELS:2002 results. There are stronger income effects in the NLS-72 results for the bottom half of the GPA distribution, but again they disappear in the ELS:2002 results.

**Four-year college attendance –** Columns 2 and 4 give the results for four-year college attendance. The results for sex are similar to the results for two-year attendance. In the full sample, the marginal effect of being Female is negative in both data sets, but the effect is very small and statistically insignificant in the ELS:2002. The change is much greater in the top-half of the GPA distribution, where being Female goes from being a 5.94 percentage point disadvantage to being a 2.76 percentage point advantage. There is no such change in the bottom half of the GPA distribution, where being female has a stronger negative impact in the ELS:2002.

There are two statistically reliable results for racial categories, ones for Black and Asian. The marginal effect for Black is 25.38 percentage points in the NLS-72 and 15.35 percentage points in the ELS:2002. The results for the two halves of the GPA distribution are very close to these findings for the full sample. The very large estimated marginal effects for Black are consistent with previous findings using similar surveys; for example, Belley and Lochner (2007) in a regression determining college attendance at a two-year or a four-year college find a coefficient for Black of .2236 using the NLSY79 data and a coefficient of .1445 using the NLSY97 data. The results for Asians also show positive marginal effects, a 14.51 percentage

point advantage in the NLS-72 and a 6.58 percentage point advantage in the ELS:2002. In this case the advantage holds for the top half of the GPA distribution but not for the bottom half of the GPA distribution.

The effect of parental education levels on four-year college attendance in the NLS data is very strong, but it is either nonexistent or much weaker in the ELS:2002. This pattern follows in the results for the two halves of the GPA distribution. The only statistically significant marginal effects in the ELS:2002 are for parents with an MA or higher in the full sample and the top half of the GPA distribution.

With West as the omitted category, the regional effects are a mirror image of the results for two-year attendance. All of the two-year effects are negative and significant, and all of the four-year effects are positive and significant. The results for urbanization again reflect the two-year results. Particularly in the ELS:2002, students from more rural origins (Town school and Rural school) are more likely to go to two-year schools and less likely to go to four-year schools. These results can be found in the results for both halves of the GPA distribution.

The results for GPA suggest that the marginal effect of GPA on the likelihood of attending four-year college is very large, much larger than the effect on two-year enrollment. A 10 percentile increase in GPA increases four-year attendance by 3.3 percentage points in the NLS-72 full sample and 5.7 percentage points in the ELS:2002. In the top half of the GPA distribution (Table 5), these effects are 4.8 percentage points for the NLS-72 and 6.1 percentage points for the ELS:2002. In the bottom half of the GPA distribution, GPA effects are 2.9 percentage points for the NLS-72 and 6.8 percentage points for the ELS:2002. Grades matter a great deal. Math scores also matter, but the effects for math test scores are slightly smaller, particularly in the ELS:2002. As an example, in the overall sample a 10 percentile increase in

the math test score has a marginal effect of 3.2 percentage points for the NLS-72 and only 2.1 percentage points in the ELS:2002. Reading scores also have significant effects, but the effects are much smaller than they are for GPA and math scores.

Our results for student ability are comparable with findings in other studies. Unlike our findings, Ellwood and Kane (2000) find that the coefficient for GPA in the regression explaining four-year college going for 1980/82 and 1992 are very similar. Belley and Lochner (2007) show statistically significant increases in college attendance for students in the second and third quartile of the Armed Forces Qualification Test (AFQT) between 1979 and 1997. Finally, Lovenheim and Reynolds (2011) summarize their estimates by saying, "These estimates suggest that ability has become more important over time in the decision between choosing a 2-year instead of a 4-year college and in choosing between attending a 4-year college and not enrolling at all …" (page 85).

Family income has the effects one would expect. Students from families with higher income are more likely to attend four-year colleges, and given the higher prices of four-year schools the effects are much larger than the comparable effects for two-year schools. Also, the effects for the full sample and the top half of the GPA distribution are quite similar in the NLS-72 and the ELS:2002. The effects of income are more pronounced in the ELS:2002 for the bottom half of the GPA distribution.

## VI. Differences in Enrollment Patterns

Between 1972 and 2004 the United States increased its college going rate from 52.3 percent of the high school class to 73.4 percent. The majority of this increase, 13.6 percent points out of a total increase of 21.1 percentage points, was in four-year college attendance. With such a large increase in four-year college attendance, it is not unreasonable to expect a decrease in the

average preparedness of the four-year college class. Our results show that there has been no such decrease. The average preparedness of students entering four-year institutions measured by high school GPA and percentiles on reading test scores has actually increased. There were decreases in these students' average percentiles on math tests, but all of the changes are small. Our conclusion is that the United States has been able to increase the four-year college going rate significantly without decreasing the average ability of the student body. The same cannot be said for the two-year student body. The average percentiles of GPAs, reading tests, and math tests for students attending two-year colleges all decreased. These results are consistent with the broad shift in college attendance that in most cases sent better-prepared students to four-year colleges and less-well-prepared students to two-year colleges.

Our results have uncovered two types of findings that explain the preservation of student ability with an increased college going rate. First, the influences of some student characteristics on college going changed differentially between the two halves of the GPA distribution, and the result was improved sorting. Second, some shifts in college going that occurred throughout the GPA distribution also improved sorting.

The first example of different shifts comes from the dramatic increase in college attendance by women. Women increased college attendance by 28.6 percentage points compared to the overall increase of 21.1 percentage points. In addition, our results show that the marginal effect of being female on four-year attendance increased 8.71 percentage points between 1972 and 2004 for women in the top half of the GPA distribution (column 6 in Table 5), but it did not change for women in the bottom half of the GPA distribution (column 6 in Table 6). The results were reversed for two-year attendance. The marginal effect of being female on two-year attendance increased 7.85 percentage points for women in the bottom half of the GPA

distribution (column 5 in Table 6), and it did not change for women in the top half of the GPA distribution (column 5 in Table 5). In sum, more women went to college, and other things equal the impact of being a woman increased four-year attendance for well-prepared women and two-year attendance for less well-prepared women. This improvement in sorting for women plays a large role in explaining the improved sorting overall.

Other studies using multiple data sets have found similar shifts. In regressions comparing two data sets, Ellwood and Kane (2000) found an increase in the coefficient for female, and Belley and Lochner (2007) found a decrease in a coefficient for male. Neither of these studies performed separate regressions for different parts of the ability distribution. Goldin, Katz, and Kuziemko (2006) discuss the reasons for the changes in college going and college completion for women and suggest several possible explanations.

The results for family income provide our second example. These results suggest that the effects of income are very similar in both surveys for students in the top half of the GPA distribution for both two-year and four-year attendance (all of the differences are small and none of them is statistically significant). However, for four-year attendance income has become more important for students in the bottom half of the GPA distribution (the differences in marginal effects are all larger and two of the three changes are statistically significant). Given the significant increases in college tuition between our two surveys, this result is consistent with financial aid doing a better job of mitigating the effects of high tuition for high ability students than it does for low ability students. The increased targeting of financial aid to high ability students documented in McPherson and Shapiro (2006) would lead to this result. Also, our two surveys straddle the introduction of the Georgia HOPE Scholarship in 1993 and scholarship programs in other states that base scholarships on students' grades rather than financial need.

However, we should be careful not to place too much emphasis on these results for income for two reasons. First, family income is probably the least accurately measured of our independent variables, especially in the NLS-72, where data come from students. Second, as we detailed above, our results for family income are dissimilar to the results of other studies that have investigated changes in the effects of family income over time.

The results for the effects of GPA on two-year attendance provide a third example of a change that differentially affected the top and bottom of the ability distribution. There was a statistically significant decrease in the marginal effect of GPA on two-year attendance in the results for the top half of the GPA distribution, and a much smaller statistically insignificant decline in the results for the bottom half of the GPA distribution. This suggests that good students in the top half of the GPA distribution were less likely to go to a two-year school than were good students in the bottom half of the GPA distribution. This shifts the sorting of students in the appropriate direction.

The second possible explanation for our finding of improved sorting by ability is that the entire distribution shifted in a way that emphasized quality for four-year attendance and decreased the emphasis on quality for two-year attendance. The obvious place to start to look for this kind of a change is with the effects of the direct measures of student quality. For four-year attendance, the coefficient on GPA percentile increased in the regressions for both halves of the GPA distribution, and the magnitude of the increase was larger for the bottom half. This means that in the ELS:2002 a very low GPA was a larger hurdle toward attending a four-year college than it was in the NLS-72 data. The results for the percentile on the math score are not consistent with our main finding. The cross-sample differences in the coefficients for this variable are all negative. This indicates that performance on the math test became less important

for determining which students go to a four-year college over time. However, in alternative specifications that drop the GPA and reading test score variables, the effects of math test scores on four-year college attendance are very similar between NLS-72 and ELS:2002 samples. Finally, there is little if any difference in the importance of percentiles on the reading test.

The results for the race indicator variables also help explain the improved sorting in the student body. The importance of the coefficients for Black and Hispanic declined in the overall results and in the results for the top half of the GPA distribution, though the results for Hispanic are not statistically significant. This means that, controlling for student quality, the race of the student mattered less for four-year attendance in the ELS:2002 than in the NLS-72, particularly for students with high GPAs. As we mentioned above, other studies that have compared college going across time have made this same finding: the effect of race on college going has gotten smaller over time.

The results for parents' education level also show that, controlling for student quality, this factor had less influence on four-year attendance in the ELS:2002 than it had in the NLS-72. In this case, the differences in the coefficients are very similar in the two halves of the GPA distribution. Students of parents who have higher levels of educational attainment are likely to be better students for both genetic and environmental reasons. If the advantages these students have are reflected in grades and test scores, we would not expect the marginal effects for these variables to be large. The fact that they are large, 21.69 percentage points for parents with a BA and 23.78 for parents with a graduate degree in the NLS-72 regressions, shows that they are a factor even after controlling for grades and test scores. The marginal effects decrease significantly in the ELS:2002, to 5.31 percentage points and 9.11 percentage points respectively, but they still are statistically significant. While these variables are still important, the decrease in

the magnitudes of the marginal effects is consistent with students and colleges paying more attention to indicators of student success and less to student background.

Our results from the indicator variables for race and parental education point in the same direction. If student quality were the only factor affecting college attendance, then racial categories and parental education would not affect which students go to college. The results for the ELS:2002 show that these factors do matter, but their effects are smaller than they were in the NLS-72. This decrease in the importance of these kinds of factors suggests that student quality has gotten relatively more important over time.

## VII. Conclusions

Between 1972 and 2004 the percentage of high school students attending college right after high school increased significantly at both two-year and four-year institutions. At the same time the average academic qualification of students attending four-year colleges changed very little. This change was caused by a larger share of the academically well-prepared students going to four-year colleges and a larger share of the less-well-prepared students going to two-year colleges. This represents an improvement in the mechanism that sorts students into various colleges. Since the sorting has improved, there is less room for further improvement, so it will be more difficult for us to continue to increase four-year enrollment without experiencing a decline in the qualifications of the student body. Still, the fact that we accomplished this change in the past demonstrates that there is nothing automatic about the relationship between the size of the four-year college class and the average qualification of students in the class.

To determine the reasons why the United States was able to accomplish this, we estimated multinomial logit models for four-year attendance, two-year attendance, and not going to college for two data sets, the NLS-72 and the ELS:2002. Our results for changes in the marginal effects

of variables between these two estimates fill in some of the reasons why there was a shift in college attendance that tended to send well-prepared students to four-year colleges and not so well-prepared students to two-year colleges. Here are three results we want to highlight.

First, the shift in college going behavior by sex has been an important part of the explanation. Women perform better in high school. They do in the 1972 data and they do in the 2004 data. In the earlier data, despite their better high school performance, they attended four-year colleges less often than men. That difference has been erased, and in the early 2000s women, particularly women in the top half of the high school GPA distribution, go to four-year colleges more often than men. This shift has significantly increased the fraction of the top half of the high school class going to college.

Our second finding, from multinomial logit estimation, is that the estimated marginal effects on four-year attendance of family background measures such as race and parents' education levels have fallen over time. To the extent that these characteristics are less important, measures of student ability such as grade point averages and test scores, the importance of which is fairly steady if not increasing, became relatively more important. This led to a college-going group that was drawn more heavily from those with good indicators of academic preparedness.

The third finding is that indicators of parental income did not seem to change much as predictors of four-year college attendance in the results for the top half of the GPA distribution, but parental income did increase in importance in the results for the bottom half of the GPA distribution. This result is consistent with more financial aid being directed toward high-ability students than is directed toward low-ability students. To the extent that states have shifted toward merit-based grants and colleges have given more generous financial aid packages to high-

ability students, this change is partially responsible for the shift in the student body to more high-ability students.

## Bibliography


Baum, Sandra, Charles Kurose, and Michael S. McPherson, "An Overview of American Higher Education," *The Future of Children* (Spring 2013): 23(1) pages 17-39.

Belley, Philippe and Lance Lochner, "The Changing Role of Family Income and Ability in Determining Educational Achievement" NBER Working Paper 13527, October 2007.

Bound, John, Michael Lovenheim and Sarah Turner, "Why Have College Completion Rates Declined? An Analysis of Changing Student Preparation and Collegiate Resources," *American Economic Journal: Applied Economics* (July 2010) pages 129-157.

Bowen, William G., Matthew M. Chingos, and Michael S. McPherson*, Crossing the Finish Line: Completing College at America's Public Universities,* (Princeton University Press, 2009).

Brand, Jennie E., Fabian T. Pfeffer, and Sara Goldrick-Rab, "Interpreting Community College Effects in the Presence of Heterogeneity and Complex Counterfactuals," California Center for Population Research, PWP-CCPR-2012-004, March 27, 2012.

Carneiro, Pedro and Sokbae Lee, "Trends in Quality-Adjusted Skill Premia in the United States, 1960-2000," *American Economic Review* (October 2011): 101(6) pages 2309-2349.

Dillon, Elanor W and Jeffrey A. Smith, "The Determinants of Mismatch Between Students and Colleges," NBER Working Paper 19286, August 2013.

Ellwood, David T. and Thomas J. Kane, "Who is Getting a College Educaiton? Family Background and Growing Gaps in Enrollment," in Sheldon Danziger and Jane Waldfogel (ed.) *Securing the Future: Investing in Children from Birth to College*, (Russel Sage Foundation, 2000), pages 283- 324.

Hoxby, Caroline M. and Christopher Avery, "The Missing 'One-offs': The Hidden Supply of High-Achieving, Low Income Students," NBER Working Paper 18586, December 2012.

Goldin, Claudia, Lawrence F. Katz and Ilyana Kuziemko, "The Homecoming of American College Women: The Reversal fo the College Gender Gap," *Journal of Economic Perspectives,* Vol. 20, No. 4 (Fall 2006) pages 133-156.

Lovenheim, Michael F. and C. Lockwood Reynolds, "Changes in Postsecondary Choices by Ability and Income: Evidence from the National Longitudinal Surveys of Youth, *Journal of Human Capital,* Vol. 5, No. 1 (Spring 2011) pages 70-109.

McPherson, Michael S. and Morton Owen Schapiro. "Watch What We Do Not What We Say: How Student Aid Awards Vary with Financial Need and Academic Merit," in *College Access: Opportunity or Privilege,* Michael S. McPherson and Morton Owen Schapiro (ed), College Board, 2006, pages 49-73.

Roderick, Melissa, Jenny Nagaoka, Vanessa Coca and Eliza Moeller, *From High School to the Future: Making Hard Work Payoff, The Road to College for Students In CPS's Academically Advanced Programs*, Research Report, April 2009, Consortium on Chicago School Research.

**Figure 1. Four-Year College Attendance by High School GPA: NLS-72 and ELS:2002**

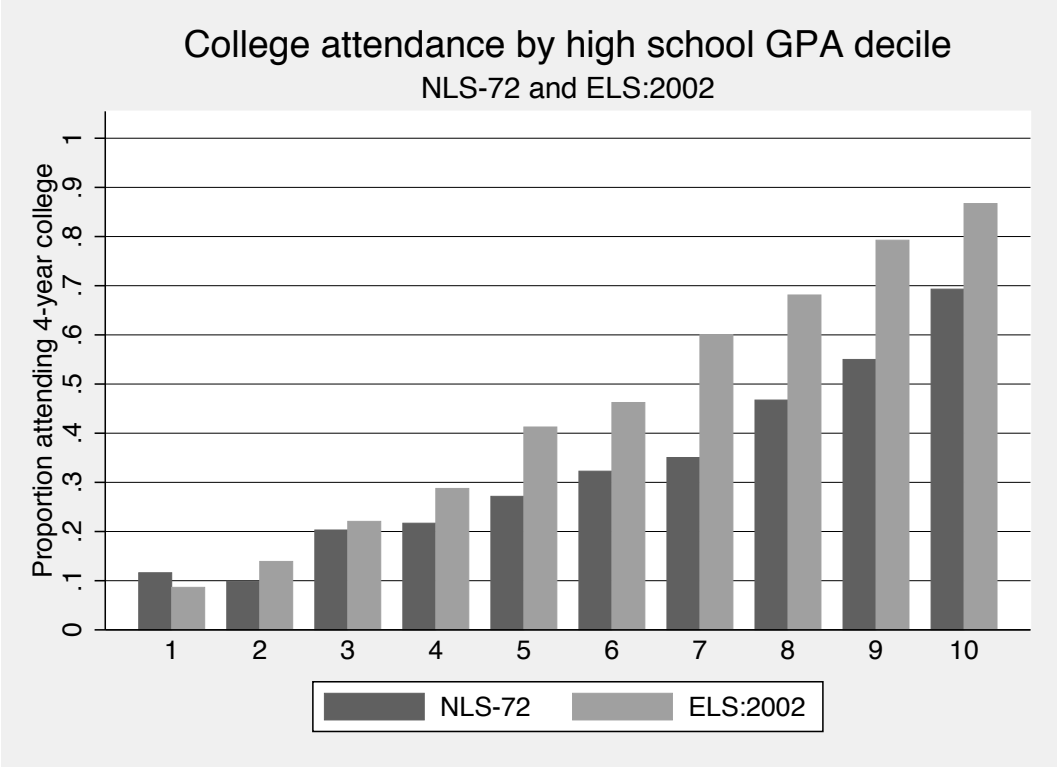


**Figure 2. Two-Year College Attendance by High School GPA: NLS-72 and ELS:2002**

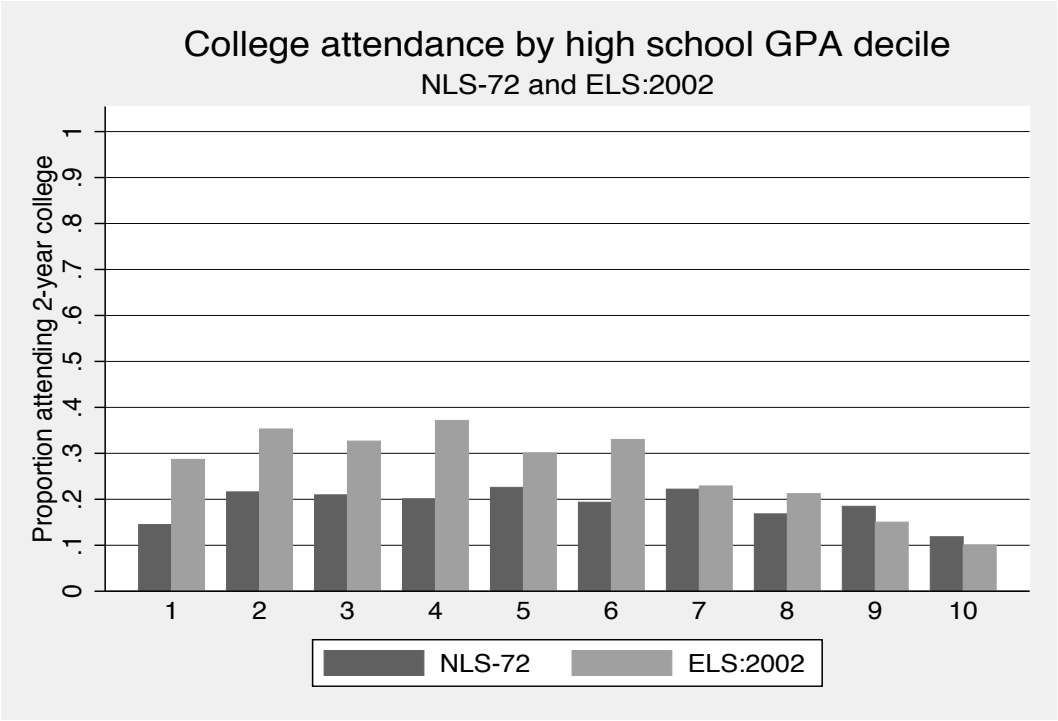

**Table 1. Summary Statistics by College Attendance, NLS-72 and ELS:2002 senior cohorts**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | 12th grade distribution | | HS+2 distribution | |
| | NLS-72 | ELS:2002 | NLS-72 | ELS:2002 |
| | Full Sample | | | |
| GPA percentile | 50.2 | 52.1 | 49 | 50.3 |
| Math test percentile | 50.2 | 51.9 | 48.7 | 50.5 |
| Reading test percentile | 49.8 | 51.9 | 48.5 | 50.5 |
| | Did not attend college | | | |
| GPA percentile | 40.1 | 32.5 | 38.9 | 30.3 |
| Math test percentile | 37.9 | 34.5 | 36.4 | 33.1 |
| Reading test percentile | 39.6 | 35.3 | 38.2 | 33.8 |
| | Attended 2-year college | | | |
| GPA percentile | 48.3 | 44.3 | 47.1 | 42.2 |
| Math test percentile | 51.3 | 44.2 | 49.7 | 42.6 |
| Reading test percentile | 49.9 | 44.6 | 48.5 | 43 |
| | Attended 4-year college | | | |
| GPA percentile | 65.7 | 67.7 | 64.6 | 66.2 |
| Math test percentile | 67.3 | 65.8 | 65.9 | 64.5 |
| Reading test percentile | 64.4 | 65.2 | 63.3 | 63.8 |

NOTES: Each cell presents a sample average. Percentiles refer to the distribution across the sample, separately for the NLS-72 or ELS:2002. Sampling weights used. "$12^{th}$ grade distribution" refers to percentiles in the sample of $12^{th}$ graders. "HS+2 distribution" refers to percentiles in the sample of respondents to the follow-up survey two years later.

**Table 2. Summary Statistics by HS GPA, NLS-72 and ELS:2002**

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| | Full sample | | Top-Half GPA | | Bottom-Half GPA | |
| | NLS-72 | ELS:2002 | NLS-72 | ELS:2002 | NLS-72 | ELS:2002 |
| Male | .493 | .484 | .422 | .402 | .568 | .575 |
| Female | .507 | .516 | .578 | .598 | .432 | .425 |
| White | .861 | .646 | .906 | .745 | .814 | .535 |
| Black | .077 | .126 | .043 | .064 | .112 | .195 |
| Hispanic | .03 | .136 | .025 | .096 | .036 | .181 |
| Asian | .0089 | .042 | .011 | .052 | .0062 | .031 |
| Other race | .023 | .05 | .015 | .043 | .031 | .058 |
| Parent dropout | .165 | .045 | .128 | .026 | .205 | .066 |
| Parent HS | .371 | .198 | .336 | .155 | .408 | .245 |
| Parent some PSE | .262 | .346 | .275 | .306 | .248 | .39 |
| Parent BA | .12 | .234 | .153 | .27 | .085 | .194 |
| Parent MA plus | .082 | .174 | .108 | .241 | .055 | .099 |
| Northeast | .261 | .179 | .241 | .172 | .282 | .187 |
| South | .246 | .341 | .224 | .317 | .27 | .367 |
| Midwest | .314 | .25 | .323 | .271 | .305 | .227 |
| West | .179 | .23 | .212 | .24 | .143 | .219 |
| City school | .27 | .266 | .238 | .244 | .303 | .29 |
| Suburban school | .274 | .398 | .28 | .403 | .268 | .393 |
| Town school | .281 | .114 | .283 | .124 | .278 | .104 |
| Rural school | .172 | .217 | .194 | .225 | .148 | .209 |
| GPA percentile | 50.2 | 52.1 | 74.9 | 75.4 | 23.9 | 26.6 |
| Math test percentile | 50.2 | 51.9 | 62.1 | 64.8 | 37.6 | 37.7 |
| Reading test percentile | 49.8 | 51.9 | 60.2 | 64.1 | 38.7 | 38.6 |
| Fam. inc. <25K (2002$) | .145 | .184 | .121 | .128 | .171 | .245 |
| Fam. inc. 25-50K (2002$) | .459 | .309 | .435 | .276 | .484 | .344 |
| Fam. inc. 50-75K (2002$) | .268 | .219 | .295 | .232 | .24 | .205 |
| Fam. inc. >75K (2002$) | .127 | .288 | .148 | .364 | .105 | .206 |

NOTES: Each cell presents a sample average. Sampling weights used. GPA distribution is across the sample, either NLS-72 or ELS:2002. HS means high school, PSE means post-secondary education, BA means bachelor's degree, MA means master's degree.

## Table 3. College Attendance Rates, NLS-72 and ELS:2002

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Two-year college attendance rates | | | | | | Four-year college attendance rates | | | | | |
| | Full sample | | Top-Half GPA | | Bottom-Half GPA | | Full sample | | Top-Half GPA | | Bottom-Half GPA | |
| | NLS-72 | ELS:02 | NLS-72 | ELS:02 | NLS-72 | ELS:02 | NLS-72 | ELS:02 | NLS-72 | ELS:02 | NLS-72 | ELS:02 |
| All | 18.9 | 26.4 | 17.7 | 20.5 | 20.2 | 32.9 | 33.4 | 47 | 47.8 | 68.2 | 18.2 | 23.8 |
| Male | 21.3 | 26 | 19.7 | 19.4 | 22.6 | 31.1 | 34.8 | 44.2 | 53.8 | 68.1 | 19.8 | 25.7 |
| Female | 16.5 | 26.9 | 16.2 | 21.5 | 17 | 35.4 | 32.1 | 50.3 | 43.4 | 68.5 | 16.2 | 21.8 |
| White | 19 | 25.4 | 17.4 | 20.3 | 20.9 | 33.3 | 34.4 | 52.5 | 48.4 | 70.3 | 17.9 | 24.9 |
| Black | 14.2 | 24.9 | 13.4 | 16.6 | 14.5 | 27.9 | 31.9 | 41.4 | 48.5 | 70.2 | 25.2 | 31 |
| Hispanic | 24.9 | 34.6 | 26 | 29.6 | 24.1 | 37.5 | 17.1 | 27.6 | 28.7 | 49.9 | 8.6 | 14.4 |
| Asian | 30.6 | 26 | 29.1 | 14.4 | 33.6 | 47.8 | 45.2 | 59.4 | 54.1 | 79.8 | 28 | 21.2 |
| Other race | 16.7 | 22.5 | 24.2 | 20 | 12.9 | 24.6 | 18.8 | 42.4 | 36 | 62.8 | 10.2 | 25.7 |
| Parent dropout | 14.7 | 30.7 | 16.7 | 30.5 | 13.3 | 30.8 | 15.7 | 23 | 26.8 | 44.4 | 8.4 | 13.7 |
| Parent HS | 18.1 | 30.3 | 18.1 | 31 | 18.1 | 29.8 | 23.5 | 28.7 | 35.8 | 48.1 | 12.8 | 15.1 |
| Parent some PSE | 24.1 | 30.6 | 21.6 | 26 | 27.1 | 34.6 | 37.7 | 39.6 | 49 | 59.5 | 24.3 | 22.2 |
| Parent BA | 17.4 | 24 | 14.6 | 16.2 | 22.8 | 36.1 | 58.5 | 60.4 | 70.3 | 77.4 | 35.9 | 34 |
| Parent MA plus | 16.6 | 16.1 | 12.1 | 11.1 | 25.8 | 29.4 | 63.8 | 73.6 | 75.1 | 85.5 | 40.4 | 41.5 |
| Northeast | 15.9 | 22 | 8.8 | 12.3 | 22.3 | 31.8 | 37.5 | 55.4 | 58.2 | 81.1 | 18.8 | 29.5 |
| South | 16.9 | 24.4 | 16.2 | 20.3 | 17.5 | 28.3 | 34 | 46.6 | 49.9 | 67 | 20 | 27.3 |
| Midwest | 15 | 26 | 15.5 | 20.4 | 14.4 | 33.3 | 33.9 | 50 | 46.5 | 69.3 | 19.9 | 24.8 |
| West | 32.9 | 33.3 | 32.7 | 26.6 | 33.3 | 41.4 | 25.8 | 37.8 | 35.7 | 59.4 | 10.4 | 12 |
| City school | 18.5 | 23.4 | 15.6 | 16.7 | 20.8 | 29.6 | 33.2 | 51.1 | 49.5 | 72.9 | 19.7 | 30.9 |
| Suburban school | 20.1 | 27.2 | 17.3 | 19.4 | 23.1 | 36 | 40.3 | 49.5 | 56.5 | 71.9 | 22.6 | 24.2 |
| Town school | 19.1 | 29 | 19.5 | 28.1 | 18.6 | 30.2 | 32.1 | 38.2 | 46.2 | 55.3 | 17.1 | 15.9 |
| Rural school | 15.7 | 27.9 | 16.6 | 22.9 | 14.4 | 33.8 | 25.1 | 42.9 | 36.2 | 64.4 | 9.7 | 17.5 |
| Fam. inc. <25K (2002$) | 15 | 29.7 | 18.9 | 29 | 12.2 | 30.1 | 22.7 | 26.7 | 33.6 | 49 | 14.5 | 14 |
| Fam. inc. 25-50K (2002$) | 18.3 | 29.5 | 17.4 | 25.1 | 19.1 | 33.4 | 27 | 38.7 | 40.6 | 59.1 | 14 | 20.6 |
| Fam. inc. 50-75K (2002$) | 21.6 | 27.6 | 19.7 | 22.4 | 24.1 | 34.2 | 39.7 | 49.4 | 52.5 | 68.2 | 23 | 26 |
| Fam. inc. >75K (2002$) | 19.8 | 20.1 | 13.8 | 12.7 | 29 | 34.3 | 55.7 | 67.1 | 71 | 81.9 | 32.8 | 38.4 |

NOTES: Each cell presents a college attendance rate in a different subsample indicated by the row and columns titles. Sampling weights used. GPA distribution is across the sample, either NLS-72 or ELS:2002. HS means high school, PSE means post-secondary education, BA means bachelor's degree, MA means master's degree.

**Table 4. Multinomial logit models of college attendance in the NLS-72 and ELS:2002**
**Marginal effects**

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| | NLS-72 | | ELS:2002 | | Comparison of NLS-72 and ELS:2002 | |
| | 2-year attendance | 4-year attendance | 2-year attendance | 4-year attendance | 2-year attendance | 4-year attendance |
| Female | -.0267*** | -.0344*** | .0275** | -5.0e-04 | .0542*** | .0339** |
| | ( .0095) | ( .0099) | ( .0108) | ( .0099) | ( .0143) | ( .014) |
| Black | -.0537*** | .2538*** | -.0757*** | .1535*** | -.0221 | -.1003*** |
| | ( .0135) | ( .0181) | ( .0143) | ( .0148) | ( .0197) | ( .0234) |
| Hispanic | .0357 | .0452 | .0185 | -.0032 | -.0172 | -.0484 |
| | ( .0236) | ( .0313) | ( .0176) | ( .0175) | ( .0294) | ( .0359) |
| Asian | .0481 | .1451*** | .0146 | .0658*** | -.0335 | -.0794* |
| | ( .0365) | ( .0433) | ( .019) | ( .0165) | ( .0412) | ( .0464) |
| Other race | -.0043 | .0043 | -.0709*** | .0467* | -.0667* | .0423 |
| | ( .0284) | ( .0353) | ( .0217) | ( .0257) | ( .0357) | ( .0437) |
| Parent HS | .0148 | .037** | .0392 | -.0577** | .0244 | -.0947*** |
| | ( .0142) | ( .0152) | ( .0276) | ( .0271) | ( .0311) | ( .031) |
| Parent some PSE | .0554*** | .1174*** | .0393 | -.014 | -.0161 | -.1314*** |
| | ( .0161) | ( .0167) | ( .0257) | ( .0259) | ( .0303) | ( .0308) |
| Parent BA | -.0017 | .2169*** | .0241 | .0531* | .0258 | -.1637*** |
| | ( .0187) | ( .0211) | ( .0277) | ( .0274) | ( .0334) | ( .0346) |
| Parent MA plus | -.0042 | .2378*** | -.0116 | .0911*** | -.0074 | -.1466*** |
| | ( .0207) | ( .0235) | ( .0288) | ( .0289) | ( .0355) | ( .0372) |
| Northeast | -.1213*** | .1068*** | -.1*** | .1501*** | .0213 | .0433** |
| | ( .0099) | ( .015) | ( .0146) | ( .0156) | ( .0177) | ( .0216) |
| South | -.1038*** | .1003*** | -.0821*** | .0976*** | .0217 | -.0028 |
| | ( .0098) | ( .0144) | ( .0139) | ( .0137) | ( .017) | ( .0199) |
| Midwest | -.1369*** | .0968*** | -.0648*** | .095*** | .0721*** | -.0019 |
| | ( .01) | ( .0141) | ( .0146) | ( .0149) | ( .0177) | ( .0205) |
| Suburban school | .0141 | -.002 | .0488*** | -.0677*** | .0347* | -.0657*** |
| | ( .0126) | ( .0127) | ( .014) | ( .012) | ( .0189) | ( .0175) |
| Town school | 4.6e-04 | -.0223* | .0666*** | -.1398*** | .0662*** | -.1175*** |
| | ( .0118) | ( .0124) | ( .0201) | ( .0165) | ( .0233) | ( .0207) |
| Rural school | -.0121 | -.0735*** | .0556*** | -.101*** | .0677*** | -.0275 |
| | ( .0138) | ( .0136) | ( .0166) | ( .0139) | ( .0216) | ( .0194) |
| GPA percentile | -7.8e-04*** | .0033*** | -.0015*** | .0057*** | -7.4e-04*** | .0024*** |
| | ( 1.8e-04) | ( 1.8e-04) | ( 2.2e-04) | ( 1.8e-04) | ( 2.8e-04) | ( 2.6e-04) |
| Math test percentile | 5.0e-04** | .0032*** | -6.6e-04** | .0021*** | -.0012*** | -.0011*** |
| | ( 2.2e-04) | ( 2.3e-04) | ( 2.8e-04) | ( 2.5e-04) | ( 3.6e-04) | ( 3.4e-04) |
| Reading test percentile | -1.2e-04 | .0016*** | -5.4e-04** | .0016*** | -4.2e-04 | 3.5e-05 |
| | ( 2.1e-04) | ( 2.2e-04) | ( 2.7e-04) | ( 2.4e-04) | ( 3.4e-04) | ( 3.2e-04) |
| Fam. inc. 25-50K (2002$) | .0159 | -.0025 | -9.7e-04 | .0303** | -.0169 | .0328 |
| | ( .0138) | ( .015) | ( .0144) | ( .0146) | ( .02) | ( .0209) |
| Fam. inc. 50-75K (2002$) | .0362** | .0354** | .0052 | .0577*** | -.031 | .0223 |
| | ( .0166) | ( .0171) | ( .0164) | ( .0159) | ( .0233) | ( .0233) |
| Fam. inc. >75K (2002$) | .0371* | .0998*** | -.0196 | .1179*** | -.0566** | .0181 |
| | ( .0211) | ( .0213) | ( .0177) | ( .0175) | ( .0276) | ( .0275) |

NOTES: Marginal effects from multinomial logit models of college choice: options are attend no college, attend a two-year college, or attend a four-year college. Separate estimation for NLS-72 and ELS:2002 data sets. Column 5 is the difference between columns 3 and 1. Column 6 is the difference between columns 4 and 2. Standard errors of differences assume estimates across data sets are uncorrelated.

**Table 5. Multinomial logit models of college attendance in the NLS-72 and ELS:2002**
**Top half of the GPA distribution, Marginal effects**

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| | NLS-72 | | ELS:2002 | | Comparison of NLS-72 and ELS:2002 | |
| | 2-year attendance | 4-year attendance | 2-year attendance | 4-year attendance | 2-year attendance | 4-year attendance |
| Female | -.0255** | -.0594*** | .0092 | .0276** | .0348* | .0871*** |
| | ( .0127) | ( .0148) | ( .0136) | ( .014) | ( .0186) | ( .0204) |
| Black | -.0709*** | .2514*** | -.1033*** | .1404*** | -.0323 | -.111*** |
| | ( .0203) | ( .0277) | ( .0176) | ( .0208) | ( .0269) | ( .0346) |
| Hispanic | .0055 | .0699 | -.0046 | -.016 | -.0101 | -.0859 |
| | ( .031) | ( .0468) | ( .0235) | ( .0258) | ( .0389) | ( .0534) |
| Asian | .0248 | .1462*** | -.0498** | .0883*** | -.0746* | -.058 |
| | ( .0396) | ( .0457) | ( .0198) | ( .0211) | ( .0443) | ( .0503) |
| Other race | .0876* | -.0532 | -.0407 | -.0039 | -.1283** | .0493 |
| | ( .0517) | ( .0529) | ( .0296) | ( .0339) | ( .0596) | ( .0628) |
| Parent HS | .0016 | .0353 | .0511 | -.0691* | .0495 | -.1045** |
| | ( .0193) | ( .0229) | ( .0395) | ( .0418) | ( .044) | ( .0477) |
| Parent some PSE | .0242 | .1137*** | .0188 | -.0139 | -.0054 | -.1276*** |
| | ( .0206) | ( .0237) | ( .0353) | ( .038) | ( .0409) | ( .0448) |
| Parent BA | -.0288 | .2305*** | -.0247 | .0678* | .0041 | -.1627*** |
| | ( .0221) | ( .0269) | ( .0348) | ( .0378) | ( .0412) | ( .0464) |
| Parent MA plus | -.0476** | .2578*** | -.0465 | .1069*** | .001 | -.151*** |
| | ( .0229) | ( .029) | ( .0347) | ( .0376) | ( .0416) | ( .0475) |
| Northeast | -.1643*** | .1323*** | -.1178*** | .1559*** | .0465** | .0236 |
| | ( .0118) | ( .0212) | ( .0168) | ( .0193) | ( .0205) | ( .0287) |
| South | -.1007*** | .1027*** | -.0564*** | .0651*** | .0443** | -.0376 |
| | ( .0123) | ( .02) | ( .0167) | ( .018) | ( .0207) | ( .0269) |
| Midwest | -.12*** | .0955*** | -.0532*** | .0665*** | .0668*** | -.0291 |
| | ( .0131) | ( .0194) | ( .0177) | ( .0191) | ( .022) | ( .0272) |
| Suburban school | .033* | -4.3e-04 | .0451** | -.046** | .0121 | -.0456* |
| | ( .018) | ( .0196) | ( .0184) | ( .0179) | ( .0257) | ( .0266) |
| Town school | .0398** | -.0399** | .0983*** | -.1498*** | .0585* | -.1099*** |
| | ( .0172) | ( .019) | ( .0271) | ( .0259) | ( .0321) | ( .0321) |
| Rural school | .012 | -.081*** | .0588*** | -.0784*** | .0468 | .0026 |
| | ( .0189) | ( .0211) | ( .0216) | ( .0208) | ( .0287) | ( .0296) |
| GPA percentile | -.0013*** | .0048*** | -.0035*** | .0061*** | -.0023*** | .0013* |
| | ( 4.3e-04) | ( 4.9e-04) | ( 4.9e-04) | ( 4.8e-04) | ( 6.5e-04) | ( 6.9e-04) |
| Math test percentile | -2.4e-04 | .0039*** | -.0017*** | .0025*** | -.0014*** | -.0014*** |
| | ( 2.8e-04) | ( 3.4e-04) | ( 3.5e-04) | ( 3.5e-04) | ( 4.5e-04) | ( 4.9e-04) |
| Reading test percentile | -2.7e-04 | .0017*** | -4.7e-04 | .0016*** | -1.9e-04 | -1.1e-04 |
| | ( 2.7e-04) | ( 3.3e-04) | ( 3.4e-04) | ( 3.4e-04) | ( 4.4e-04) | ( 4.8e-04) |
| Fam. inc. 25-50K (2002$) | -.011 | .0088 | -.0068 | .0075 | .0042 | -.0013 |
| | ( .0182) | ( .0227) | ( .0187) | ( .02) | ( .0261) | ( .0303) |
| Fam. inc. 50-75K (2002$) | .0049 | .0482* | -.0134 | .0462** | -.0183 | -.002 |
| | ( .0206) | ( .0249) | ( .02) | ( .0207) | ( .0287) | ( .0324) |
| Fam. inc. >75K (2002$) | -.0263 | .1355*** | -.0577*** | .1024*** | -.0314 | -.0331 |
| | ( .0234) | ( .0298) | ( .0211) | ( .0225) | ( .0315) | ( .0374) |

NOTES: Marginal effects from multinomial logit models of college choice: options are attend no college, attend a two-year college, or attend a four-year college. Sample includes only respondents with GPAs in the top half of the sample distribution. Separate estimation for NLS-72 and ELS:2002 data sets. Column 5 is the difference between columns 3 and 1. Column 6 is the difference between columns 4 and 2. Standard errors of differences assume estimates across data sets are uncorrelated.

**Table 6. Multinomial logit models of college attendance in the NLS-72 and ELS:2002**
**Bottom half of the GPA distribution, Marginal effects**

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| | NLS-72 | | ELS:2002 | | Comparison of NLS-72 and ELS:2002 | |
| | 2-year attendance | 4-year attendance | 2-year attendance | 4-year attendance | 2-year attendance | 4-year attendance |
| Female | -.0292** | -.0079 | .0492*** | -.033** | .0785*** | -.0251 |
| | ( .0138) | ( .0128) | ( .0171) | ( .0136) | ( .022) | ( .0187) |
| Black | -.0209 | .237*** | -.0353 | .1534*** | -.0144 | -.0836** |
| | ( .0204) | ( .025) | ( .0234) | ( .0219) | ( .031) | ( .0332) |
| Hispanic | .075** | .0115 | .0344 | .0185 | -.0406 | .007 |
| | ( .0339) | ( .0383) | ( .0263) | ( .0238) | ( .0429) | ( .0451) |
| Asian | .0518 | .1504 | .1163*** | .0084 | .0645 | -.142 |
| | ( .0662) | ( .0923) | ( .0352) | ( .0268) | ( .0749) | ( .0961) |
| Other race | -.0504 | .0374 | -.1051*** | .1023*** | -.0547 | .0649 |
| | ( .0316) | ( .0421) | ( .0317) | ( .0373) | ( .0447) | ( .0563) |
| Parent HS | .0266 | .0411** | .0206 | -.0419 | -.006 | -.083** |
| | ( .0204) | ( .0196) | ( .0388) | ( .0333) | ( .0438) | ( .0386) |
| Parent some PSE | .0848*** | .127*** | .0528 | -.0084 | -.0319 | -.1354*** |
| | ( .0242) | ( .0236) | ( .0372) | ( .0335) | ( .0443) | ( .0409) |
| Parent BA | .0216 | .2047*** | .0751* | .038 | .0535 | -.1666*** |
| | ( .0305) | ( .0344) | ( .0429) | ( .0378) | ( .0527) | ( .0511) |
| Parent MA plus | .0454 | .2144*** | .0289 | .0661 | -.0165 | -.1483*** |
| | ( .0357) | ( .0392) | ( .0466) | ( .0416) | ( .0587) | ( .0572) |
| Northeast | -.0772*** | .0822*** | -.0918*** | .171*** | -.0146 | .0888** |
| | ( .0166) | ( .0237) | ( .0246) | ( .0291) | ( .0297) | ( .0375) |
| South | -.1004*** | .1043*** | -.1165*** | .1526*** | -.0161 | .0483 |
| | ( .0158) | ( .0238) | ( .0229) | ( .0236) | ( .0278) | ( .0335) |
| Midwest | -.1532*** | .1086*** | -.0841*** | .1423*** | .0691** | .0337 |
| | ( .0152) | ( .0236) | ( .0239) | ( .0267) | ( .0283) | ( .0356) |
| Suburban school | 5.7e-04 | -.0069 | .0473** | -.0813*** | .0467* | -.0745*** |
| | ( .0178) | ( .0159) | ( .0212) | ( .0152) | ( .0277) | ( .022) |
| Town school | -.0342** | -.0052 | .0211 | -.107*** | .0553* | -.1018*** |
| | ( .0163) | ( .0155) | ( .0294) | ( .0183) | ( .0336) | ( .024) |
| Rural school | -.0321 | -.0683*** | .0438* | -.1109*** | .076** | -.0426* |
| | ( .0205) | ( .0172) | ( .025) | ( .0165) | ( .0323) | ( .0238) |
| GPA percentile | 4.1e-04 | .0029*** | 8.4e-05 | .0068*** | -3.3e-04 | .0039*** |
| | ( 4.6e-04) | ( 4.6e-04) | ( 5.9e-04) | ( 5.0e-04) | ( 7.5e-04) | ( 6.7e-04) |
| Math test percentile | .0014*** | .0024*** | 5.2e-04 | .0016*** | -8.3e-04 | -7.8e-04* |
| | ( 3.2e-04) | ( 3.0e-04) | ( 4.5e-04) | ( 3.6e-04) | ( 5.6e-04) | ( 4.7e-04) |
| Reading test percentile | 9.3e-05 | .0014*** | -5.7e-04 | .0015*** | -6.6e-04 | 1.2e-04 |
| | ( 3.0e-04) | ( 2.7e-04) | ( 4.3e-04) | ( 3.4e-04) | ( 5.2e-04) | ( 4.3e-04) |
| Fam. inc. 25-50K (2002$) | .0485** | -.0118 | .0087 | .0472** | -.0398 | .059** |
| | ( .0204) | ( .0184) | ( .022) | ( .0211) | ( .0301) | ( .028) |
| Fam. inc. 50-75K (2002$) | .0728*** | .0253 | .0251 | .0665*** | -.0477 | .0411 |
| | ( .0258) | ( .0222) | ( .0264) | ( .0247) | ( .0369) | ( .0332) |
| Fam. inc. >75K (2002$) | .1217*** | .0595** | .0223 | .1363*** | -.0994** | .0768* |
| | ( .036) | ( .0285) | ( .029) | ( .0276) | ( .0463) | ( .0397) |

NOTES: Marginal effects from multinomial logit models of college choice: options are attend no college, attend a two-year college, or attend a four-year college. Sample includes only respondents with GPAs in the bottom half of the sample distribution. Separate estimation for NLS-72 and ELS:2002 data sets. Column 5 is the difference between columns 3 and 1. Column 6 is the difference between columns 4 and 2. Standard errors of differences assume estimates across data sets are uncorrelated.