NBER WORKING PAPER SERIES

LEAVING BOYS BEHIND:
GENDER DISPARITIES IN HIGH ACADEMIC ACHIEVEMENT

Nicole M. Fortin
Philip Oreopoulos
Shelley Phipps

Working Paper 19331
http://www.nber.org/papers/w19331

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 2013

We would like to acknowledge Lori Timmins for her outstanding research assistance on this project. We would also like to thank Jerome Adda, Joseph Altonji, Marianne Bertrand, Russell Cooper, David Card, Steve Durlauf, Christian Dustmann, Andrea Ichino, Claudia Goldin, Larry Katz, John Kennan, Magne Mogstad, Mario Small, Uta Schonberg, Chris Taber, Thomas Lemieux, Glen Waddell, Ian Walker, Basif Zafar, and seminar participants at Bocconi University, Einaudi Institute for Economics and Finance, European University Institute, Federal Reserve Bank of New York, Harvard University, Norwegian School of Business and Economics, Paris I, Sciences Po, University College London, University of Oregon, University of Wisconsin-Madison, Yale University, the CEA 2011, the CIFAR SIIWB Workshop, the NBER Summer Institute 2013, and SOLE 2012 for helpful comments on this and earlier versions of the manuscript. We thank ICPSR and MTF for allowing us to use the data, and the usual disclaimer applies. The authors are grateful for CIFAR's financial support. Fortin also acknowledges funding from SSHRC Grants #410-2011-0567. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2013 by Nicole M. Fortin, Philip Oreopoulos, and Shelley Phipps. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Leaving Boys Behind: Gender Disparities in High Academic Achievement
Nicole M. Fortin, Philip Oreopoulos, and Shelley Phipps
NBER Working Paper No. 19331
August 2013
JEL No. I20,J16,J24

ABSTRACT

Using three decades of data from the "Monitoring the Future" cross-sectional surveys, this paper shows that, from the 1980s to the 2000s, the mode of girls' high school GPA distribution has shifted from "B" to "A", essentially "leaving boys behind" as the mode of boys' GPA distribution stayed at "B". In a reweighted Oaxaca-Blinder decomposition of achievement at each GPA level, we find that gender differences in post-secondary expectations, controlling for school ability, and as early as 8th grade are the most important factor accounting for this trend. Increases in the growing proportion of girls who aim for a post-graduate degree are sufficient to account for the increase over time in the proportion of girls earning "A's". The larger relative share of boys obtaining "C" and C+" can be accounted for by a higher frequency of school misbehavior and a higher proportion of boys aiming for a two-year college degree.

Nicole M. Fortin
Vancouver School of Economics
University of British Columbia
#997-1873 East Mall
Vancouver, BC V6T 1Z1
Canada
nifortin@interchange.ubc.ca

Philip Oreopoulos
Department of Economics
University of Toronto
150 St. George Street
Toronto, ON M5S 3G7
Canada
and NBER
philip.oreopoulos@utoronto.ca

Shelley Phipps
Department of Economics
Dalhousie University
6214 University Avenue
Halifax, NS B3H 4R2
Shelley.Phipps@Dal.Ca

1. Introduction

Women now far outnumber men among recent college graduates in most industrialized countries (OECD, 2008). As Goldin, Katz, and Kuziemko (2006) observe, the puzzle is: "Why have women overtaken men in terms of college completion instead of simply catching up to them?" The growing female dominance in educational attainment raises new questions about gender disparities arising throughout school-ages.[1] This paper is asking two questions: 1) Are boys and girls equally well-prepared for college? and 2) What accounts for the growing gender disparity in favor of girls in obtaining high grades in secondary school?

Girls have long obtained better grades, on average, in high school than boys.[2] As shown in Figure 1a, the average gender gap in GPA among high school seniors (scaled out of 4 points) hovers steadily around 0.2 between 1976 and 2009.[3] Because historically these achievements never translated into higher levels of educational attainment or better labor market outcomes for women relative to men, much research has concentrated on explaining the remaining gaps in women's performance, particularly in mathematics (e.g. Guiso et al., 2008; Bedard and Cho, 2010). Conversely, the relative underperformance of males, especially in reading, has attracted little attention until recently (LoGerfo, Nichols, and Chaplin, 2006; Cornwell, Mustard, and Van Parys, 2013). Interest in the academic performance gap favoring women is changing for a number of reasons.[4]

The first goal of this paper is to document changes in gender disparities in the academic performance of high school students (12^{th}, 10^{th}, and 8^{th} graders) over the last three decades using survey data from the "Monitoring the Future" (MTF) project.[5] We find that an increasing proportion of students are earning A grades, arguably allowed by the progressive disaffection

[1] According to OECD (2008), the average share of the student population in tertiary education in OECD countries accounted for by women reached 55% in 2005. Only four countries are likely not to achieve at least parity between men and women by 2015: Korea, Turkey, Japan and Switzerland.

[2] This is observed in other countries as well. See Machin and McNally (2005) for Britain, Lai (2010) for China.

[3] The gender gap in GPA from the MTF match (within standard errors) the numbers from the National Assessment of Educational Progress High School Transcript Study for 1990, 2000, 2005 and 2009, also reported in National Center for Education Statistics (2004), as well as the numbers reported in Cho (2007) for 1984 from the High School and Beyond survey.

[4] The more difficult job prospects of men without a post-secondary education and feared labor shortages in some professional specialties that attract few women, such as orthopedic surgeons, are mentioned, as well as repercussions for the marriage prospects of college-educated women and concerns among boys' parents about a "failure to launch" (Bell, Burtless, Gornick and Smeeding, 2007).

[5] To the best of our knowledge, Jacob and Wilder (2012) is the only other contemporaneous paper using the MTF to study educational expectations. They study on the impact of these expectations on college going.

with "grading on the curve".[6] As shown in Figure 1b, the percentage of 12th grade students reporting in the MTF that they earn A's (93-100%) almost doubled, from 8.5% in the 1980s to 16.6 % in the 2000s, and the difference between the proportion female and the proportion male in this category also doubled from 3.2% to 5.4%.[7] From the 1990s to the 2000s, the female advantage in the proportion of 10th and 8th graders earning A's also increased, from 3.6% to 4.8% and from 4.9% to 5.5%, respectively.

The second goal of the paper is to identify the relative importance of four sets of factors that changed differently by gender over time and that could account for this growing gender disparity in academic achievement. These include plans for the future, — likely driven by changes in the labor market—, non-cognitive skills, the family environment, and working while in school. The post-secondary aspirations and expectations of high school students, as well as their choice of high school program (from vocational to academic) to enact these career plans, are the set of factors that changed the most over the last three decades. While returns to college have increased more for men than for women over the last three decades[8], Figure 2a shows that just the opposite happened to expectations about "definitively" attending a graduate or professional school after college: They have risen faster for women than for men. Among seniors, boys' expectations about attending graduate school were slightly higher than girls' from 1976 to 1983, but thereafter a gap in favor of girls began to emerge, widening in the 1990s, and reaching 9 percentage points before the Great Recession. Figure 2b presents the gender ratio among students who say that they "will definitively go to a four year college", a question asked at the three grade levels. Among seniors, the gender ratio (female share) was around 50 percent up to the early 1980s, overshot the gender ratio in actual enrollment rates by a few percentage points in the 1990s to stabilize around 57 percent in the 2000s.[9] Interestingly, the gender ratio in

[6] To be clear, the erosion of grading on the curve is not seen as "causing" an increasing proportion of girls to earn A's, rather the absence of constraints on the proportion of students earnings A's implies that we do not have to be preoccupied by potential general equilibrium effects that such constraints would imply.

[7] In the MTF, an A grade corresponds to a percentile grade in the 93-100% range. The exact years are 1976 to 1988 for the 1980s, and 2000 to 2009 for the 2000s for 12th graders, and 1991-1999 for the 1990s for 10th and 8th graders.

[8] This is a well-known stylized fact (see Fortin, 2006, among others) illustrated in Appendix Figure A1a.

[9] Note that given the higher percentage of boys who drop out of school, the gender ratio in the sample of 12th graders ranges from 51% in earlier years to 52% in later years. The gender ratio in expectations about college-going has thus moved from a 1% deficit to 5% surplus. This also shown in Appendix Figure A2a and A2b, which illustrates how the see-saw pattern of the 1980s is linked to age differences by gender, as seen in staggered gender birth ratios, themselves arising from the changes in family planning methods eighteen years earlier.

expectations about attending a four year college emerges as early as grade 8, when it hovers around 55 percent.

Goldin and Katz (2002) have argued that the 1970s "Pill Revolution" was crucial in allowing young women to formulate plans for higher education without the fear of interruptions for family reasons. We argue that in subsequent decades, the ongoing progress of women in the professions has continued to fuel young women's career plans involving graduate and professional schools, while pink collar jobs have totally lost their appeal. Table 1 shows some dramatic changes in the vocational expectations of high school seniors (available only for a sub-sample). The percentage of girls thinking that, at age 30, they will be working in a professional job requiring a post-graduate degree (doctoral or equivalent) had climbed from 15.3 percent in the 1980s to 27.1 percent in the 2000s, meanwhile that percentage among boys went from 13.5 to 16.4 percent.[10] With the advent of computerization and other office technologies, there has been a substantial decline in labor market demand for clerical work matched by the decline in vocational expectations: the percentage of girls expecting to work in clerical job at age 30 has plummeted from 21 percent in the 1980s to less 3 percent in the 2000s.[11] Interestingly, this sharp decline is not matched by as great a decline in skilled and semi-skilled work, craftsmanship, and protective services as expected occupations for boys. For our complete sample, our educational expectation variables include a full range of career plans for life after high school, such as serving in the army, attending a vocational college, a two-year college, a four-year college and even aiming for graduate or professional school.

As with most studies of changes in gender differentials, we construct counterfactual states of the world based on the observed responses and respective endowments of males and females. We then apply a reweighted decomposition methodology (Fortin, Lemieux, and Firpo, 2011) aimed at separating endowment effects from response effects under the assumption that the distribution of unobservables conditioning on observables is independent of gender. We focus on an analysis of changes over time in the distribution of GPA because gender differences in average GPA have not changed over the past thirty years, while the gender ratio of students admitted to college, those with high GPA, has changed substantially.

[10] To facilitate the exposition, we regroup our data for seniors into three time periods of 10-12 years, 1976-1988, 1989-1999 and 2000-2009, rather than the four decades.

[11] See Table A1 which displays the labor market outcomes of young people (25-39 years old) over the 35 years period. It shows that the actual proportion of young women employed in clerical work has dropped significantly, although not as dramatically as desired occupations shown in Table 1.

Our decomposition of the impact of educational expectations on GPA may only be interpreted as a direct effect if the distribution of unobservables conditional on observables is independent of gender. To explore whether changes in other factors, such as ability, returns to college, or financial constraints underlie changes in expectations, the MTF surveys also include data on educational aspirations and subjective assessments of school ability, which allows us to consider indirect effects, and to present bounds on the direct effects of educational expectations with and without these controls. With respect to the possibility of reverse causality, where changes to GPA distribution may affect education aspirations, we note that the sudden 1991 rise in the expectations of girls about pursuing a graduate degree (Figure 2b) preceded and exceeded in size a similar 1993 rise in the proportion of girls obtaining A's (Figure 1b). That being said, the time line does not allow us to completely dispel concerns about potential endogeneity problems, especially with regards to the size of the effect found. But, to the extent that for 8th graders, there is still time to improve on their academic performance to meet their more distant vocational expectations, we regard the estimates for this group as less tainted by endogeneity biases.

The data do not allow us to consider the effect of teaching styles (Algan, Cahuc, and Shleifer, 2010) or of the teachers' gender (Dee, 2005, 2006), which have attracted recent attention. We do however include information on the type of high school program (academic, vocational, general, etc.) attended, which are associated with different GPA distributions.[12] Following the wave of interest in the impact of non-cognitive traits, we account for smoking, alcohol binging, and school misbehavior.[13] The other sets of factors that we consider are the family environment and working during school. Families with girls are, on average, larger in line with Angrist and Evans (1998), have less educated parents, more working mothers, and more fathers not living in the same household, as documented by Dahl and Moretti (2008).[14] These last two gender gaps in family characteristics are increasing over time. Finally, a decline in the labor

[12] Similar information on the type of high school program (academic, general, vocational, etc.) in which students are enrolled is also asked in the NLS72 and NELS-88, for example.

[13] Some psychologists (e.g. Duckworth and Seligman, 2006) have argued that self-control and self-discipline give girls the "edge", we attempt to capture a similar notion with the "alcohol binging" variable. The gender gap in smoking, which had closed in the 1970s and early 1980s, has reopened more recently. The information on the frequency of being sent to the principal or to detention for bad behavior in the last year is only available for 10th and 8th graders. School misbehavior, which has decreased over time for boys, has reduced - the gender gap in reprehensible behavior. See Figures A1b and A1c.

[14] As shown in Figure A1d.

force participation of boys during school, from 85 percent in the 1980s to the 76 percent in the 2000s, has lead the closing of the gender gap in labor force participation during high school.

The paper is organized as follows. Section 2 introduces the MTF surveys and presents some descriptive statistics about gender disparities in academic achievement and in the explanatory factors. Section 3 presents our empirical specification and explains the reweighted decomposition methodology. Section 4 presents the decomposition results and discusses their interpretation. Finally, section 5 concludes.

2. Data and Descriptive Statistics

The data used are from the "Monitoring the Future" surveys, which have been conducted by the Institute for Social Research, University of Michigan mainly to monitor substance abuse every year from 1976 onwards for Grade 12 students and from 1991 onwards for students in Grades 8 and 10.[15] Given higher male drop-out rates, our sample of 12th graders is only 48 to 49 percent male. Thus our sample of seniors likely comprises a positively selected sample of boys, likely leading us to understate any gender gap favorable to girls by comparison to a wider sample of boys. It is thus useful to compare high school seniors with high school sophomores and 8th graders, who remain subject to minimum age school leaving laws. We focus on the core sample, which comprise 10,000 to 16,000 observations per grade per year, which allow us to perform the breakdown by gender and GPA.[16]

Our dependent variable is the self-reported school grade which is elicited from the following question: "Core 20: Which of the following best describes your average grade so far in high school? D (69 or below), C- (70-72), C (73-76) , C+ (77-79), B- (80-82), B (83-86) , B+ (87-89), A- (90-92), A (93-100)."[17] Obviously, grades from administrative data are preferable to self-reported grades because students with different characteristics may misreport their grades differently.[18] But we find that the self-reported grades from the MTF are very reliable.[19] When

[15] Because of the focus on drug use, those who use illicit drugs as seniors are oversampled, we are careful to use the sample weights provided to remove any bias resulting from that oversampling. There exists a practically inaccessible longitudinal component, which surveys a small subset of the students (Bachman et al., 2002).

[16] Many more attitudes and behavioral questions are asked of students answering one of 6 modules, including a host of non-cognitive variables but they are asked only of a subset of students..

[17] Following standard institutional practice, the self-reported grades in the 9 categories are translated in the numbers: A (93-100) 4.0, A- (90-92) 3.7, B+ (87-89) 3.3, B (83-86) 3.0, B- (80-82) 2.7, C+ (77-79) 2.3, C (73-76) 2, C- (70-72), 1.7, D (69 or below) 1, where 2.3 and 2.7 and so on, are the rounded versions of 2.333 and 2.666.

[18] See Balsaa, Giuliano, and French (2011) on grade misreporting by alcohol-binging students.

we compare the average grades of 12th graders from the MTF to those of the NAEP High School Transcript Surveys (HSTS), we find that the gender differences, as well as the grade inflation, do match within standard errors, even though the scales used are somewhat different.[20] Note that this report finds, as Goldin, Katz, and Kuziemko (2006) also reported that girls are increasingly taking more challenging math and science courses.

There are other questions in the MTF survey of seniors asked before this one directed at getting subjective assessments of school ability (Core 16) and intelligence (Core 17) , which would allow students, who are so inclined, to boast about their abilities. The question on subjective school ability asks: "Core 16: Compared with others your age throughout the country how do you rate yourself on school ability? Far below average, below average, slightly below average, average, slightly above average, above average, far above average."[21] On average both genders rate their subjective school ability equally high, but boys rate themselves more favorably on intelligence than girls do.[22] We note that the raw correlation between subjective school ability and self-reported grades is only 58% among seniors.

Table 2 begins by reporting a simple difference-in-difference analysis of the changes over time and by gender in self-reported grades and in expectations about attending graduate or professional school of 12th graders. Like Figure 1, Panel A of Table 2 shows little change over time in the significant female advantage of about 0.2 (on a 4 point scale) in average grades, if anything boys have made small gains (about 0.01) in relative grades. Panel B shows that the stability in average grades masks a significant increase in the female advantage in the proportion of students with the highest grades (A (93-100) students), which represents the pool of students who can be confident of being admitted to graduate school if they continue to succeed in their undergraduate studies. Our focus on the gender gap in top grades follows from the findings of previous studies (Jacob, 2002; Goldin, Katz, and Kuziemko, 2006; Cho, 2007; Conger and Long, 2010) showing that the lower college admission rates of men can in large part be accounted for

[19] The wording of the question on self-reported grades in terms of an upward scale is similar to commonly used questions about self-reported income where individuals are asked to declare in which income bracket their income falls and may be less prone to error than simple declarative questions.

[20] The HSTS scale has 5 categories, which include a zero: A(90–100) 4.0, B(80–89) 3.0 , C(70–79) 2.0, D(60–69) 1.0, F (less than 60) 0.0.

[21] As with the other categorical variables, we rescale this variable to be between 0 and 1 using the following formula: Category k=1-(n-k+1)/(n+1), when k=n is highest category to be recoded into 1. This recoding presumes equal distance between the categories.

[22] The question on intelligence asks on the same seven points scale: "Core 17: How intelligent do you think you are compared with others your age?" See Figures A3a and A3b.

by their lower high school performance.[23] However, better high school performance explains "how" more girls are admitted to college but not "why".

As in Figure 2a, Panel C of Table 1 shows an even greater and significant increase of the female advantage in expectations of attending graduate school. Indeed from the 1980s to 1990s, the proportion of women expecting to attend graduate school more than doubled from 10% to 21%, while the proportion of men increased only by half, from 10% to 15%. The fact that the increase in the gender differential in expectations to attend graduate school was more sizeable (5.3 percentage points) from the 1980s to the 1990s, when women' progress in the labor market was sharpest, than from the 1990s to the 2000s (2.6 percentage points) are in line with our conjecture that gender differences in plans for the future fuel gender differences in high academic achievement.

A more complete picture of changes in academic achievement is presented in Figure 3 which displays histograms, corresponding to the actual data, overlaid with a kernel density of the self-reported grades of girls and boys in 12th grade. The figures clearly show a progressive disaffection over the past thirty-five years with "grading on a curve" with the alternative "competency grading" gaining in importance.[24] In the 1980s, the mode and median of the grades distribution roughly coincided in the B range. By the 2000s, the mode of the girls' grade distribution had moved from B to A, while the mode of the boys' grade distribution stayed at B.[25] This is what we call "leaving boys behind"; although the proportion of boys in the A range has increased over time, the gender gap in the proportion of students at the very top of the GPA distribution has increased. Figures 5a and 5b report the same data for 10th and 8th graders for two time periods, 1991-1999 and 2000-2009. Here the girls' advantage appears even more dramatic.

One may wonder whether these distributional changes arise from increases in the mean grade pushing the upper tail against the upper boundary or from increases in the upper tail pulling the mean. With the first hypothesis, the explanations behind the increases in mean grade remain unspecified under the heading "grade inflation". We test this hypothesis, by first estimating an ordered probit of GPA levels for the three time periods and then using the

[23] The higher average grades of girls are at times equated with their higher average non-cognitive abilities (Jacob, 2002; Becker, Hubbard, and Murphy, 2010).
[24] "Grading on a curve" means grading relatively to classmates, whereas "competency grading" means that if a student's work deserves an A for example, the student should get an A irrespective of the number of classmates getting A's.
[25] Similar gender differences can be found in the administrative grades available in the Add Health data for example.

estimated cut-offs of the second and third period to inflate the predictions from a similar model estimated only on the first period. The resulting predictions for the A and B+ levels are found to be below the observed ones, which tell us that this type of grade inflation is not sufficient to lead to the observed increases in the proportion of students getting the high grades.

The means of selected core variables for seniors are reported in Table 3 for each of the three time periods of interest.[26] The first two rows display the average school grade index and the students' own evaluation of their school ability. It shows that despite having lower grades, boys rate their own school ability higher than girls.[27] Similar male overconfidence has been reported among college students by Stinebricker and Stinebricker (2009) who find that college bound boys are less likely to succeed, because of their overall lower performance.[28]

Selected demographic characteristics are presented next. A high alcohol binging category is representative of the fact that boys are still more likely than girls to report these risky behaviors. Girls tend to live in families that on the surface might be less likely to foster high academic achievement. Four percent more girls than boys report not living in the same household as their father, 3 percent more girls than boys report that their mother works all the time and about 3 percent more boys than girls report than their father or mother has completed college.[29] The next row shows that the gender gap in paid work participation has closed over time, although boys continue to work longer hours and get higher pay (see Table A2). The types of high school programs show that the gap in favor of girls in the proportion of seniors enrolled in an academic program has grown. For example, while about 3 percent more girls than boys were enrolled in an academic program in the 1980s that proportion increased to 7 percent in the 2000s. Among 8th graders, already 4 percent more girls than boys report being enrolled in a college preparatory program, although a large proportion of students (43 percent of both boys and girls) are not clear about their type of high school program.

[26] The statistics are computed on observations with no missing variables. This reduces the sample sizes by comparison with Table 1. Complete descriptive statistics for 12th graders are presented in Table A2. Descriptive statistics for 10th and 8th graders are available upon request.

[27] Girls in 1976-1988 and boys in 2000-2009 having similar average GPA of 3, but the boys' school ability index of 0.664 is significantly greater than the girls 0.651.

[28] Although grades by topic are not reported in the MTF, numerous studies (especially those using the National Education Longitudinal Study) show that boys continue to maintain an advantage in math test scores (but not in math grades), especially at the high end of the distribution. The boys' overconfidence may be built on these scores.

[29] We note that the gender gaps in family characteristics are similar in the sample without Blacks.

Among 12th graders, the first question about post-secondary plans asks about expectations: "Core 21: How likely (definitively won't, probably won't, probably will, definitively will) is it that you will do each of the following things after high school? a) Attend a technical or vocational school, b) Serve in the armed forces, c) graduate from a two-year college, d) graduate from college (four-year program), e) attend graduate or professional school after college?" A second question asks about aspirations: "Core 22: Suppose you could do just what you'd like and nothing stood in your way. How many of the following things would you WANT to do?" with the five options above being supplemented by none of the above. Among 8th and 10th graders, only the expectations questions are asked. Among 12th graders in particular, the expectations question raises issues of endogeneity with respect to GPA. Some high ability students may have low expectations of graduating from a four-year college because of their low GPA, rather that the other way around. The aspirations question attempts to circumvent that problem with the preamble if "nothing stood in your way". Controlling for subjective school ability (Core 16 above) and aspirations (Core 22) is an attempt to alleviate concerns about cognitive dissonance. Among 8th and 10th graders, the issue of endogeneity of educational expectations is presumably less severe as there is more time to adjust one's level of effort. For these students, we control for two retrospective measures of school ability (grade retention and whether school was often hard), as well as school misbehavior.[30]

Table 3 shows that in the 1980s, although seniors of both genders had similar expectations about graduating from college and attending graduate school, girls already had higher aspirations (close to 2 percentage points) than boys. That is, more girls than boys have "things that stand in their way". By the 2000s, the expectations index for both college and graduate school was 8 percentage points higher for girls than boys.[31] Gender differences in aspirations for college and graduate school are respectively 8 percentage points and 11 percentage points higher in favor of girls. Finally, 6 percent of boys vs. 3 percent of girls have declared no post-secondary aspirations, in line with higher drop out rate among boys.

[30] More precisely, responses to the grade retention question "Have you ever had to repeat a grade in school?" are available as a binary variable. The responses to the two questions: "Now thinking back over the past year in school, how often did you…find the school work too hard to understand?" "…get sent to the office, or have to stay after school, because you misbehaved?" were coded on a 5 points scale.

[31] Comparing seniors in 1972 from the NLS72, in 1980 from the H&B, in 1992 from the NELS88, and in 2004 from the ELS2002, Ingels and Dalton (2008) also find that in 2004, more girls than boys expected to pursue graduate studies, whereas it was the opposite in 1972.

3. Empirical Specification and Reweighted Decomposition Methodology

Our empirical specification is based on a behavioral threshold model of academic performance where educational goals, fashioned in elementary school and likely influenced by parental desires, play a prominent role in determining, given a level of aptitude, an individual's choice of optimal GPA.[32] This follows an emerging consensus in the psychology literature that students form reliable perceptions of their academic competency around 5th grade (Herbert and Stipek, 2005) and can already form some expectations about college-going.[33] Indeed, decisions to enroll in a college preparatory high school program, to move to a neighborhood with a better high school, and to apply to a magnet school have to be made early in a student's life. Under the assumption that effort is costly, the student's optimal choice of GPA will be the minimum of the range that opens the door to the education level needed to fulfill her/his vocational goals. Students motivated towards professional or medical careers will come to understand they need to aim for A's. Those thinking about white collar occupations such financial analyst will need a bachelor degree and can aim for B's; those not as career motivated during youth, or expecting jobs that require fewer credentials may instead aim for C's. Parents are likely involved in helping form these career and grade expectations, implicitly or explicitly through actively assisting with homework, helping set goals, and managing children's time. The above threshold model is consistent with the changes over time in the shape in the distribution of GPA levels (shown in Figures 3 and 4) in response to changes in career expectations, especially for girls. The less pronounced change in shape among boys would be consistent with more convex costs of effort, possibly associated with higher psychic or social costs of being seen as working hard.[34] This model helps rationalize the relative underperformance of boys as the consequence of career choices that require lower levels of educational attainment. We do not exclude the possibility that some students revise their plans, but because we do not have access to the MTF longitudinal data, we cannot explore this avenue.[35]

---

[32] The model is exposited in more detail in Appendix B.

[33] This is consistent with the high school tracking taking place in many European countries around the ages of 10 and 11 (Dustmann, 2004; Checchi and Flabbi, 2007).

[34] Bishop (2006) argues that there are different studying and homework cultures by gender, something like "smart boys get high marks without showing effort" or "it is not cool for boys to work hard to get top grades".

[35] Not having access to the longitudinal MTF data, we cannot address directly the issue of expectations formation by contrast with Stange (2008), Zafar (2011) and Jacob and Wilder (2012). Using data from the NELS88, Jacob and Wilder (2012) report that only 35% of high school students update their educational expectations from grade 8 to grade 10; from grade 10 to 12, that percentage is only 25%.

In this study of gender gaps in academic achievement, we seek to identify how student characteristics map into the distribution of GPAs differently by gender. We are primarily interested in how changes over time in these determinants help account for changes over time in gender differentials in academic achievement. For each of the three time periods, we estimate the following academic achievement equation,

$$Prob[G_i = c] = h_g^c(S_i, A_i, L_i; X_i, X_i^p), \qquad c = 1, \ldots, 9, \tag{1}$$

where G_i is the student's GPA, S_i denotes the student's educational goals and A_i denotes the student' academic aptitude. We combine the high school program, the schooling expectations and aspirations to measure S_i. The student's school aptitude, A_i is proxied using the subjective measure of school ability (introduced in Section 2), available for 12th grade students.[36] For 8th and 10th grade students, we measure aptitude by how often he or she found school "too hard" in the last year, in addition to a measure of past grade retention. We include an indirect measure of effort, following the tradition in labor economics of deriving non-market time, here study time, as the difference between total time (T) and labor market time (L_i): $E_i = T - L_i$. To account for the impact of non-cognitive skills, we include measures of cigarette smoking and alcohol binging, which may relate to time impatience, and a measure of school misbehavior for 8th and 10th graders. Exogenous characteristics of student X_i, including race and living in a SMSA as well as an extended set of family characteristics, X_i^p, thought to be pre-determined variables, are included in the specification.[37]

We estimate a different linear probability model by gender for each level of GPA, which carries some advantages and disadvantages. The advantages of using a linear probability model are that we do not have to rely on the assumptions of normality of residuals. By comparison with an ordered probit model, this model allows the educational responses to be different by level of GPA. Given that the detailed decomposition of the gender differentials requires linear educational responses, this estimation procedure gives us coefficients that can readily be used.[38]

---

[36] Educational aspirations and subjective school ability measures are available only for the 12th graders. Clearly, lagged measures would have been preferred.

[37] These family environment characteristics include living in the same household as the father, the mother, and siblings (separate questions), the number of siblings, whether the mother had a paid job while growing up (not at all, some of the time, most of the time, all the time), the level of education (6 levels) of the father and of the mother.

[38] By comparison with a multinomial logit, there is no need to compute the marginal effects at the mean of characteristics, which may not correspond to a representative student for some GPA levels. Among the disadvantages is the fact that the predicted probabilities are not bounded between 0 and 1. In practice, we will find some under-predictions (<0), but the predicted probabilities over GPA levels sum to 1.

We follow the literature on gender wage differentials in applying an Oaxaca-Blinder (OB) type of decomposition, but we extend the decomposition to the overall distribution of grades and follow the approach of Fortin, Lemieux, and Firpo (2011) to analyze the impact of gender differences in the educational response functions. We now give a short summary of the formulas behind this modified decomposition. With the standard OB decomposition, the researcher seeks to determine what portion of the gender gap in grades is attributable or "explained" by differences in the characteristics of boys and girls and what portion remains "unexplained". Here, owing to reweighting we can argue that the "unexplained" part corresponds to gender differences in the structural function $h_g^c(S_i, A_i, L_i; X_i, X_i^p)$ of equation (1). In the detailed decomposition, we can apportion parts of the aggregate decomposition to particular explanatory factors and responses to determine which of these factors are relatively more important.

The classic OB methodology is based on the construction of a counterfactual state of world. Assuming that grades (G) can be modeled as a linear (in the parameters) function of characteristics (X) that is different for girls ($F = 1$) and boys ($F = 0$)

$$\mathbb{E}(G|X, F = 1) = \mathbb{E}(X|F = 1)\beta_1 \text{ and } \mathbb{E}(G|X, F = 0) = \mathbb{E}(X|F = 0)\beta_0,$$

under the zero conditional mean assumption, $\mathbb{E}(\varepsilon|X, F) = 0$. The OB counterfactual, $\mathbb{E}(G^{OB}) = \mathbb{E}(X|F = 1)\beta_0$, asks "What would boys' grades be if they had the same characteristics as girls?" using the coefficients estimated on the sample of boys. As shown in Fortin, Lemieux, and Firpo (2011), with reweighting we can construct a counterfactual that more precisely isolates the educational responses. This counterfactual uses the coefficients estimated using the grades outcomes of boys, but the characteristics of the sample of boys reweighted to be like girls.

More precisely, we reweight the sample of boys so that the distribution of their characteristics (X) is similar to that of girls, using the following reweighting function

$$\Psi(X) = [(Prob(X \mid F = 1))/(Prob(X \mid F = 0))]$$
$$= [(Prob(F = 1 \mid X))/(Prob(F = 0 \mid X))] \cdot [Prob(F = 0)/Prob(F = 1)].^{39}$$

The counterfactual coefficients β_o^1 are estimated on the sample of boys reweighted to look like girls $\{X_0, \Psi(X_0)\}$. The difference $(\beta_1 - \beta_o^1)$ reflects the true gender gap in educational responses, and the counterfactual means are computed as: $\bar{X}_0^1 = \sum\{i : F = 0\} \Psi(X_i) \cdot X_i$. The

[39] The second equation makes use of Bayes' Law to allow the computation $\Psi(X)$ in the case of continuous X variables. See DiNardo, Fortin, and Lemieux (1996) for details.

reweighted decomposition uses the predicted grades, $(X_0|F = 1)\beta_o^1$, from the reweighted sample as counterfactuals,

$$\Delta_{O,R}^{\mu} = \mathbb{E}(X|F = 1)\beta_1 - \mathbb{E}(X_0|F = 1)\beta_o^1 + \mathbb{E}(X_0|F = 1)\beta_o^1 - \mathbb{E}(X|F = 0)\beta_0$$
$$= \Delta_{E,R}^{\mu} + \Delta_{X,R}^{\mu}$$

to obtain an aggregate decomposition as the sum of an educational response effect, $\Delta_{E,R}^{\mu}$, and a composition effect, $\Delta_{X,R}^{\mu}$. Inasmuch as grade dummies can be averaged out, this decomposition relies on the additional assumptions of common support and ignorability $(F \perp \varepsilon|X)$, that is, conditioning of observables, unobservables are assumed to be the same across gender.

Each term of the reweighted decomposition can be further broken down into the "pure" effect and a residual term. The composition effect, $\Delta_{X,R}^{\mu}$, is written as the sum of a pure composition effect, $\Delta_{X,p}^{\mu}$, and a specification error, $\Delta_{X,e}^{\mu}$,

$$\Delta_{X,R}^{\mu} = \mathbb{E}(X_0|F = 1)\beta_o^1 - \mathbb{E}(X|F = 0)\beta_0 + \mathbb{E}(X_0|F = 1)\beta_0 - \mathbb{E}(X_0|F = 1)\beta_0$$
$$= [\mathbb{E}(X_0|F = 1) - \mathbb{E}(X|F = 0)]\beta_0 + \mathbb{E}(X_0|F = 1)(\beta_o^1 - \beta_0)$$
$$= \Delta_{X,p}^{\mu} + \Delta_{X,e}^{\mu}.$$

Similarly, the educational response term, $\Delta_{E,R}^{\mu}$, can be written as the sum of a pure response effect $\Delta_{E,p}^{\mu}$ plus a reweighting error $\Delta_{E,e}^{\mu}$,

$$\Delta_{E,R}^{\mu} = \mathbb{E}(X|F = 1)\beta_1 - \mathbb{E}(X_0|F = 1)]\beta_o^1 - \mathbb{E}(X|F = 1)\beta_o^1 + \mathbb{E}(X|F = 1)\beta_o^1$$
$$= \mathbb{E}(X|F = 1)(\beta_1 - \beta_o^1) + [\mathbb{E}(X|F = 1) - \mathbb{E}(X_0|F = 1)]\beta_o^1$$
$$= \Delta_{E,p}^{\mu} + \Delta_{E,e}^{\mu} .$$

The specification error $\Delta_{X,e}^{\mu} = \mathbb{E}(X_0|F = 1)(\beta_o^1 - \beta_0)$ corresponds to the difference in the composition effects estimated by reweighting and by using simple regressions, where $\mathbb{E}(X_0|F = 1)$ is the mean of the reweighted sample. The reweighting error $\Delta_{E,e}^{\mu} = [\mathbb{E}(X|F = 1) - \mathbb{E}(X_0|F = 1)]\beta_o^1$ goes to zero in a large sample.

Because of the linearity of these expressions, the detailed decomposition or the apportionment of the composition and educational response effects to each explanatory variable is straightforward. In practice, this detailed reweighted decomposition can be obtained by running two decompositions: OB1) with the sample of girls $(F = 1)$ and the reweighted sample of boys looking like girls to get the pure wage structure effect, OB2) with the sample of boys $(F = 0)$ and the reweighted sample of boys looking like girls to get the pure composition effect.

4. Empirical Results

Before going on to the decomposition results, it is useful to show which of our explanatory variables are more significant and in explaining a cross-section in grade outcomes and how these relationships differ by gender. We estimate equation (1) separately by gender, and for the reweighted sample of boys (results not reported), for each of the nine GPA levels and for each of the three time periods to compute the decomposition results. We do this for each grade level (8^{th}, 10^{th}, and 12^{th}) and under two specifications for grade 12 students (with and without conditioning on our measure of subjective ability) for a total of 324 GPA regressions. To conserve space we report selected estimated coefficients only for seniors in the 2000s and only for the two GPA levels where the gender achievement gaps were largest, that is for the A and C+ grades, and only for boys (not reweighted) and girls.[40]

4.1 Determinants of Top and Below Average GPA

Tables 4a and 4b report the estimated coefficients of the explanatory variables listed in Table 3, along with t-statistics. In Table 4a, the dependent variable is equal to 100 if the student gets an A, and 0 otherwise, so that the coefficients indicate the added probability of getting an A associated with the explanatory variables. In Table 4b, we estimate the covariates of getting exactly a C+. Specification 1 includes educational expectations, under the assumption that students take their abilities and other limitations into account when formulating their expectations. Specification 2 explores the consequences of this assumption by explicitly controlling for subjective school ability and for educational aspirations formed without possible limitations resulting from ability or other constraints.[41] We do find that the explanatory power of expectations is reduced when we control for student ability and student aspirations, yet educational expectations remain among the most significant explanatory variables. In Table 4a, wanting and expecting to attend graduate school, especially among boys, is associated with a higher probability, in the 4 to 14 percentage point range of getting an A. Conversely, expecting to go to a two-year college shows a negative association, in the −7 to −11 percentage point range with getting A. Note that expecting to go to a four-year college is so widespread that it has little explanatory power. In Table 4b, consistent with our threshold model, expecting to go to a two-

[40] The detailed estimates are presented in Appendix Tables A3a and A3b. The coefficients estimated on the reweighted sample are available upon request; they are generally close to the ones estimated on the sample of boys.
[41] Appendix Table A3 also reports the estimates from specification 3 that exclude expectations altogether, leaving the type of high school program to capture plans for the future.

year college is associated with a higher probability, in the 2 to 6 percentage point range range of getting a C+ , again especially for boys.[42] The types of high school program, thought to be part of a student's plans for the future, show similarly strong associations with these GPA levels and significant differences across genders. Girls in academic high school programs are more likely to get an A, and less likely to get a C+ than boys.

As in Balsa et al. (2011), alcohol binging is associated with a significantly lower probability of getting an A, about −4 percentage point, and a higher probability of getting a C+, about 1-4 percentage point. Similar effects are found for smoking variables, in the −3 to −6 percentage point range for getting an A and the +2 percentage point for a C+. We view these correlations as symptomatic of time impatience or caring less about the future. Focusing on family background variables, we find that controlling for school ability (going from Specification 1 to 2) substantially reduces the impact of parental education on students' probabilities of getting an A or a C+, although that association remains significant for girls.[43] To the extent that parental education is capturing the family socio-economic status, these results are consistent with past research (e.g., Cameron and Heckman (2001) and Reynolds and Pemberton (2001)), showing that the biggest influence of parental resources on the children's education operates through academic performance. Other important family influences, more impervious to the addition of subjective school ability, are the actual presence of parents in the household. The father not living in the same household and the mother working have significant effects (about −1 to −4 percentage point) on the probability of getting an A, and positive effects on the probability of getting a C+ (about 1-2 percentage point). Interestingly the effect of the absent father is somewhat greater for girls, and that of the mother working is somewhat greater for boys. Consistent with Buchmann and DiPrete (2006), we find that these effects have increased from the 1980s to the 2000s. In comparison to the above regressors, the effects of the variables related to working during school are generally less significant and show some of the non-linear patterns found in the literature.

[42] This interesting new finding would be masked if the dependent variable was getting at least C+. In this case, expecting to go to a four-year college dominates.

[43] Specification 3 in Table A3, which excludes expectations, shows larger effects of parental education.

4.2 Decomposition results

4.2.1 Overview

To succinctly summarize the aggregate decomposition results, we first present them in the form of graphs that display the gender differentials for the entire GPA distribution. Figure 5 first shows the female/male differences in percentage for each GPA levels, $\Delta_{O,R}^{c}, c = 1,\ldots 9$, traced by lines; positive numbers indicate a larger value for females, negative numbers indicate a larger value for males. The numbers behind the line graphs show that the female advantage in the percentage of seniors getting A's increases from 3.7 percentage points in the 1980s, to 4.7 points in the 1990s to 6.1 points in the 2000s.[44] At the same time, the male advantage in the C+ grade decreases from 4.4 percentage points in the 1980s, to 3.9 points in the 1990s to 3.2 points in the 2000s. Among 10th and 8th graders, the female advantage in the percentage of 10th and 8th students getting A's increases from 3.6 and 5.2 points, respectively in the 1990s, to 4.7 and 6.0 points, respectively in the 2000s. The similar decreases in male advantage in mediocre grades for 10th and 8th graders are from 3.1 and 2.4 percentage points in the 1990s to 2.6 and 1.5 percentage points in the 2000s. Thus for seniors, the changes in gender differentials to be accounted for correspond to 38 percent (2.3 percentage points) of the differential in top grades in the 2000s, and to 40 percent (1.3 percentage points) for mediocre grades.

Figure 5 also show the decomposition of the gender differentials into the composition effects, $\Delta_{X,p}^{c}$, the portion "explained" by gender differences in characteristics, and the educational response effects, $\Delta_{E,p}^{c}$, the portion attributed to the fact that the relationship between characteristics and GPA levels differs by gender.[45] For each time period, positive bars indicating the excess percentage of girls in a GPA level (negative bars indicating the excess percentage of boys) are divided into two; the bottom darker (blue) portion corresponds to the composition effects and the upper lighter (beige) portion corresponds to the educational response effects.

[44] See Table 5, Tables A4 and A5. These numbers are a bit different from the ones reported in Table 1, Panel B, row 3 (3.2, 4.4, and 5.4) because for the analysis, we restrict the sample to those observations for which we have complete data.

[45] The remainder terms, the specification errors $\Delta_{X,e}^{c}$ and the reweighting errors $\Delta_{E,e}^{c}$ are reported in Tables 5 and 6, and Appendix Tables A4 and A5. They account for the discrepancies between the bars and the lines (e.g. GPA=4 for 8th graders). The specification and reweighting errors are generally found to be at least an order of magnitude smaller than the main effects $\Delta_{X,p}^{c}$ and $\Delta_{E,p}^{c}$.

For 12th graders in Panel A, the portions attributable to composition effects generally increase over the three time periods, especially at the top of the grade distribution.[46] Averaging over all GPA levels, the "explained" part grew from a mere 10% of the total gender differential in the 1980s to 32% in the 1990s and to 37% in the 2000s. For the 10th and 8th graders, in Panel B and C, respectively, the specification of the educational response functions includes two proxies for school aptitude and a school misbehavior index not available for seniors.[47] With better measures of cognitive and non-cognitive skills, an even larger share of the gender differentials is accounted for by gender differences in the explanatory variables, especially at the bottom of the GPA distribution.[48] Among 10th (and 8th) graders, averaging over all GPA levels, the "explained" part accounts for almost half of the total gender differential: more precisely 54.1% (41.7%) in the 1990s and 31.1% (62.1%) in the 2000s.

4.2.2 Tabular Decomposition Results for Selected GPA Levels

In Table 5, we present the detailed decompositions for the 2000-2009 time period in a more classic tabular form, which includes standard errors, for the GPA levels where the gender differentials are the largest (A and C+) for seniors. Table 6 presents similar decomposition results for the 10th and 8th grade students. The results for the earlier periods can be found in Appendix Tables A4 and A5. We begin by discussing the composition effects in the upper panel of each table.

Because the female/male difference in school aptitude is negative, the effects of subjective school ability for seniors, and of the two retrospective measures of aptitude for 10th and 8th graders, go in the wrong direction: Their coefficients are positive for top grades and negative for mediocre grades. This reduces the part of the gender differentials, negative for top grades and positive for mediocre grades, accounted for by the explanatory variables. For

[46] To conserve space, we display only Specification 1. The decomposition results from Specification 2 are presented in Table 5.

[47] The list of variables available for 10th graders and 8th graders is the following: dummies for race (white/non-white), SMSA, ever held back, smoked cigarettes per day (4), alcohol binging last two weeks (4), sibling not same household, father not same household, mother not same household, mother working (3), father's education (7), mother's education (7), worked during school, average hours of work (6), average earnings (7), type of high school program (4), indexes for school misbehavior last year, school too hard last year, educational expectations (army, vocational, go to college, complete 4 year college). So the main differences with Specification 2 for seniors are the absence of the number of siblings, of the aspirations for post-secondary choices, and the expectations of going to graduate school.

[48] There are a few exceptions in the middle of the grade distribution where the gender differentials are quite small: GPA levels C+ (2.7) and B (3) in 2000s for 10th graders; GPA level B (3) for 8th graders. For 10th graders, we thus present the detailed decomposition results for the A and C grades, where the gender gaps are the largest.

example in Table 5, going from Specification 1 to Specification 2 see the reduction in "Total Explained".[49] We also note that race, living in a standard metropolitan statistical area (SMSA), and family background variables are other sets of "contrarian" or "swimming upstream" variables: these variables work to the advantage of boys (because there are more black girls, more girls with absent father, etc.,) and reduce the percentage of girls with top grades and of boys with mediocre grades.[50] That is, if girls were as confident as boys about their school ability, if they lived in similar families, if there were as few Black girls living in SMSA as boys, the girls' grades would be even higher. In the 2000s, there would be from 0.6 to 0.9 percent more girls than boys earning A's. Although minor, it is still interesting to find that some high achieving girls are "swimming upstream."

Our major result is that educational expectations are the most important factor accounting for gender differentials in academic achievement. Table 5 shows that, for A grades, gender differences in expectations account for 2.03 out of 2.44 percentage points of the "explained" by gender differences in characteristics (in Specification 1) in the 2000s, this is up from 1.13 in the 1990s and from 0.23 in the 1980s.[51] For the C+ grades, the numbers are -1.19 out of -1.22 percentage points of the "explained" in the 2000s, up from -0.66 in the 1990s, and from -0.08 in the 1980s. As noted above, controlling for subjective school ability (Specification 2) reduces the absolute magnitude of the gender differentials accounted for by expectations, but not the portion explained. So, not only has the explanatory power of educational expectations gone from virtually negligible to almost all from the 1980s to the 2000s, but they also account for almost all the changes to be accounted for (2.3 and 1.3 percentage pointes) noted above.

In Table 6 (for 10^{th} and 8^{th} graders), we find similarly impressive results in the accounting power of expectations: for the A grades, the part explained by expectations is 1.26 out of 1.35 "Total Explained" for 10^{th} graders in the 2000s, up from 1.04 in the 1990s, and it is 1.53 out of 2.21 "Total Explained" for 8^{th} graders in the 2000s, up from 1.07 in the 1990s . Overall these results convey the same message as the one suggested by Table 2: here, even after controlling for a host of other factors, gender differences in educational expectations (and

[49] This effect is similar to the gender differences in educational attainment on the gender pay gap. In recent years, gender differences in education reduce the explained part of the gender pay gap. In the 2000s, the Total Explained corresponds to more 40% of the gender achievement gap in Specification 1, but only 17% in Specification 2.
[50] Comparing Table 5 and Tables A4 and A5 show that these effects increased over time.
[51] See Tables A4 and A5.

changes therein) account for the largest share (and the most salient changes over time) in gender differentials in academic achievement.

There is another set of variables that has consistent explanatory power going in the right direction across all specifications. Gender differences in non-cognitive skills measured by smoking and alcohol binging among seniors, and with the addition of school misbehavior among 10th and 8th graders account for 0.56, 0.58, and 1.29 of the gender gap in A grades, respectively.[52] For mediocre grades, the numbers are of similar order of magnitude, from -0.54 to -0.36 among 10th graders, and from -0.50 to -0.54 among 8th graders. We thus find that gender differences in non-cognitive skills are the second most important factor to account for "explained" gender differences in academic achievement. However as shown in Appendix Tables A4 and 5, we find some evidence that these effects are decreasing over time, at least for 12th and 10th graders.

Finally, it is interesting to consider the contribution of changes in gender differences in educational responses presented in the bottom panels of Tables 5 and 6, noting that the interpretation of these differences crucially depends on the omitted category in each case. The most important difference is linked to the type of high school program attended, where the omitted category is "other (not specified) high school." As we saw in previous tables, not only are girls increasingly attending college preparatory high school, but here we show that they are benefiting more (in terms of grades) from it than boys. This differential educational response adds to the total effect of "plans for the future" factors in accounting for gender differences in academic achievement.

The effects of gender differences in educational responses associated with family background is more difficult to interpret because departures from the omitted category (families with father present, mother present, one sibling, mother not working, both parents with high school education) are a more complex affair and the results are sensitive to which number of siblings is the omitted category (especially in the 1990s).[53] Nevertheless, they indicate that family background generally bolsters the response of high achieving girls by comparison with boys. A similar effect seems to apply to "work during school," where the omitted categories are

[52] We note however that these effects are constant over time.

[53] Such sensitivity is not surprising given that even using an instrumental variable strategy that exploit exogenous variation in family size, Conley and Glauber (2006) find a strong effect of sibship size on second-born boys' grade retention, but no effect on first-born boys.

not working, zero hours of work and zero wages.[54] Working during school seems to act as a complement rather than a distraction for high-achieving girls.

4.2.3 Graphical Detailed Decomposition Results for all GPA Levels

Figure 6a and 6b display the results of the detailed decomposition for each grade and GPA level for each category of factors. Figure 6a shows the two most important categories of factors: plans for the future (includes type of high school program and educational expectations) and student attributes (race, SMSA, smoking and binging, school ability and misbehavior where available), while Figure 6b focuses on the family environment and working during school. The conventions are as for Figure 5: the lines trace the magnitude of the gender gap, and the bars for each GPA levels are divided into two, the darker (blue) one capturing the composition effects and the lighter (beige) the educational response effects. In some instances, either effect can be negative, as explained above.[55]

The overall message emerging from Figure 6a is the same as the one we took away from Tables 5a and 5b. The effects of "Plans for the future" displayed in Panels A, are by far the most important explanatory factors contributing to both the composition and educational response effects, generally with the right signs, except for the very low GPA levels. More girls than boys are aiming for professions that require a graduate degree, more girls are getting A's. More boys than girls are aiming for skilled worker jobs and protective service occupations, more boys are getting C+'s.[56] This message is even stronger among 10th graders and 8th graders. The composition effects associated with plans for the future are generally accounting for more than 50% of the gender differentials, both at the low and high end of the GPA distribution. For these younger students, plans after high school are arguably further in the future and thus less likely endogenous (in the sense of resulting from cognitive dissonance issues). Youth with lower GPA are less likely to say that they will not go to college because of their lower GPA, given that many believe that there is still time for improvement. For 10th and 8th graders, Figure 6 shows in Panel B that students' attributes are a non-negligible set of factors. Here, they include not only

[54] Note that working during school is always the factor that has the least economic significance in the composition effects.
[55] The distance between the height of the bars and the symbol on the line corresponds to the portion of the gender differential accounted for by other factors.
[56] See Table A1.

smoking and alcohol binging, but also school misbehavior which accounts for a sizeable share of the gender differentials at the lowest GPA levels.

Figure 6b illustrates that relatively little is left to explain by family environment once plans for the future and aptitude (for the 10th and 8th graders) are included. Worth a mention is the "swimming upstream" more difficult family environment among high achieving girls present for all grade levels and some educational responses effects to that family enviroment among boys with B grades (and with C grades among seniors). Working during school is less frequent among 10th and 8th graders and has a negligible effect there. Among seniors, it bolsters high achieving girls while keeping more boys in the B grades, but these effects are very small.

In summary, the decomposition results show a marked improvement, over the three time periods, in the model's ability to account for gender differences in academic achievement. This finding is essentially due to the increasing explanatory power of gender differences in "plans for the future" and applies equally well to grade 12, grade 10, and grade 8 students. Indeed, when the expectation variables are omitted, we can account for relatively little of the gender differences.[57] A startling aspect here comes from the fact that the explanatory power of the educational expectations is as great for 8th graders as it is for 12th graders. To the extent that educational expectations of 8th graders are less likely endogenous with respect to GPA levels than those of seniors, this is welcome news for the validation of the model. Certainly, among 12th graders, Specification 2, which controls for subjective school ability and educational aspirations, grants less explanatory power to the model, but this does not diminish the relative importance of educational expectations, to the contrary. By comparison with the cohorts of students studied in Goldin, Katz, and Kuziemko (2006), the educational expectations of young women has risen even more in the 2000s, and are the most likely explanation for why girls are leaving boys behind in terms of earning top grades in high school. By comparison, boys' disruptive behavior has a sizeable and persistent effect, but it does not grow over time.

6. Conclusion

Using a long-lived series of detailed cross-sectional surveys of high school students, this paper set out to identify which factors among a set of plausible culprits,— plans for the future,

[57] Detailed results are available upon request. To give an example, while specification 1 allows us to account for more than a third of the gender differences in A's (2.4 out of 6.07 points) in the 2000s among 12th graders (Table 4a). Without the expectations variables, this is reduced to 0.81 points.

non-cognitive traits, family environment, and labor market work during school—, are relatively more important in accounting for changes over the past three decades in the gender achievement gap, especially at the top of the GPA distribution. By comparison with other studies that simply focus on college-going, we are able to better distinguish the drivers of the academic achievement of boys and girls because of the full range of post-secondary options available in our data, from serving in the army, to attending a vocational, a two-year college, a four-year college and even graduate or professional schools. Indeed, most of the identification of the effect of educational expectations comes from either the two-year college or the graduate school options. Wanting to "go to college" is simply a too common aspiration to be informative.

Our findings show that the predominance of girls at the top of the GPA distribution is rooted in their higher educational expectations, themselves linked to career plans that include a graduate degree (such a law or medical degree). More precisely, in the 2000s, "Plans for the Future" is the most important set of explanatory factors accounting for the girls' higher share of A's at the three grade levels (12^{th}, 10^{th}, and 8^{th} graders). This set of factors is important enough to account for all of the increase of 2.3%, from the 1980s to 2000s, in the gender difference in the percentage of seniors earnings A's.

By comparison with girls, more boys think that they are likely to enter military service or to attend a vocational school. Because the career plans of boys include more predominantly male occupations (craftsmen, protective service and military service occupations, engineers and architects) that do not require advanced degrees, their lower share of high grades is consistent with the "threshold" model that we propose. In an era where much emphasis for improving students' achievement is placed on schools and teachers, this paper offers a long term view, which highlights the role of students' motivation and gender differences therein. Clearly, among 8^{th} and 10^{th} graders, the second dominant factor accounting for the lower grades of boys is a measure of the frequency of having been set to the office or to detention over the previous year. This suggests that motivation and misbehavior may go hand-in-hand. We note that there are many on-going field experiments, such as SDRC's "Future to Discover", whose preliminary results indeed seem to suggest that boys' plans for the future are more moveable than girls.

References:

Bachman, Jerald G., Patrick M. O'Malley, John E. Schulenberg, Lloyd D. Johnston, Alison L. Bryant, and Alicia C. Merline (2002) *The Decline of Substance Use in Young Adulthood- Changes in Social Activities, Roles and Beliefs,* Mahwah, N.J.: Lawrence Erlbaum Assoc.

Balsaa, Ana I., Laura M. Giuliano, and Michael T. French (2011) "The Effects of Alcohol Use on Academic Achievement in High School," *Economics of Education Review,* 30: 1 –15.

Becker, Gary S., William H. J. Hubbard, and Kevin M. Murphy (2010) "The Market for College Graduates and the Worldwide Boom in Higher Education of Women," *American Economic Review,* 100 (2): 229–233

Bedard, Kelly and Insook Cho (2010) "Early Gender Test Score Gaps across OECD countries," *Economics of Education Review*, 29(3): 348-363.

Bell, Lisa, Gary Burtless, Janet Gornick and Timothy M. Smeeding (2007) "Failure to Launch: Cross-National Trends in the Transition to Economic Independence," chap. 2 in *The Price of Independence: the Economics of Early Adulthood*, edited by Sheldon Danziger and Cecilia Elena Rouse. New York: Russell Sage Foundation.

Bertrand, Marianne and Jessica Pan (2011), "The Trouble with Boys: Social Influences and the Gender Gap in Disruptive Behavior", Working paper Booth School of Business.

Bishop, John H. (2006) "Drinking from the Fountain of Knowledge: Student Incentive to Study and Learn – Externalities, Information Problems and Peer Pressure," in *Handbook of the Economics of Education,* Elsevier - North Holland, Vol. 2: 909-944.

Buchmann, Claudia and Thomas A. DiPrete (2006). "The Growing Female Advantage in College Completion: The Role of Family Background and Academic Achievement," *American Sociological Review*, 71: 515-541.

Cameron, Stephen V. and James J. Heckman (2001) "The Dynamics of Educational Attainment For Black Hispanic, and White Males," *Journal of Political Economy*, 109(3): 455-499.

Checchi, Daniele and Luca Flabbi (2007) "Intergenerational Mobility and Schooling Decisions in Germany and Italy: The Impact of Secondary School Tracks," IZA DP No. 2876.

Charles, K., and Luoh, M.-C. (2003), "Gender Differences in Completed Schooling", *Review of Economics and Statistics*, 85(3), 559–77.

Cho, Donghun, (2007) "The Role of High School Performance in Explaining Women's Rising College Enrollment," *Economics of Education Review*, 26: 450–462

Conger, D. and M.C. Long (2010) "Why are Men Falling Behind? Explanations for the Gender Gap in College," *Annals of the American Academy of Political and Social Science*, 627(1): 184-214.

Conley, Dalton and Rebecca Glauber (2006) "Parental Educational Investment and Children's Academic Risk: Estimates of the Impact of Sibship Size and Birth Order from Exogenous Variation in Fertility," *Journal of Human Resources*, 41(4): 722-737.

Cornwell, Christopher, David, Mustard, and Jessica Van Parys (2013) "Noncognitive Skills and the Gender Disparities in Test Scores and Teacher Assessments: Evidence from Primary School," *Journal of Human Resources*, 48(1): 236-264.

Dahl, Gordon B. and Enrico Moretti (2008) "The Demand for Sons." *Review of Economic Studies* 75 (4): 1085-1120.

Dee, Thomas (2005) "A Teacher Like Me: Does Race, Ethnicity, or Gender Matter? " *American Economic Review*, 95:158-166.

Duckworth, A. L., and M. E. P. Seligman (2006) "Self-Discipline Gives Girls the Edge: Gender in Self-Discipline, Grades, and Achievement Test Scores" *Journal of Educational Psychology* 98(1), 198-208.

Dustmann, Christian (2004) "Parental Background, Secondary School, Track Choice, and Wages," *Oxford Economic Papers*, 56: 209–230.

Fortin, Nicole M. (2006) "Higher Education Policies and the College Premium: Cross-State Evidence from the 1990s." *American Economic Review*, 96: 959-987.

Fortin, Nicole M., Thomas Lemieux and Sergio Firpo (2011) "Decomposition Methods," in O. Ashenfelter and D. Card (eds.) *Handbook of Labor Economics*, Vol. 4A, Amsterdam: North-Holland , 2011, 1-102.

Goldin, C. and L.F. Katz (2002) "The Power of the Pill: Oral Contraceptives and Women's Career and Marriage Decisions," *Journal of Political Economy*, 110: 730-770.

Goldin, Claudia, Lawrence F. Katz, and Ilyana Kuziemko (2006) "The Homecoming of American College Women: The Reversal of the College Gender Gap," *Journal of Economic Perspectives*, 20: 133–156.

Greene, J. P. and M. A. Winters (2006) "Leaving Boys Behind: Public High School Graduation Rates", Manhattan Institute for Policy Research, No. 48.

Guiso, Luigi, Ferdinando Monte, Paola Sapienza, and Luigi Zingales, (2008) "Culture, Gender, and Math," *Science*, 320(5880): 1164-1165.

Haveman, R. and B. Wolfe (1995) "The Determinants of Children's Attainments: A Review of Methods and Findings," *Journal of Economic Literature*, 33(4):1829-1878.

Herbert, Jennifer and Deborah Stipek (2005) "The Emergence of Gender Differences in Children's Perceptions of their Academic Competence," *Applied Developmental Psychology*, 26: 276–295.

Ingels, S.J., and B.W. Dalton (2008). *Trends Among High School Seniors, 1972–2004* (NCES 2008-320). National Center for Education Statistics, Institute for Education Sciences, U.S. Department of Education. Washington, DC.

Jacob, Brian A. (2002) "Where the Boys Aren't: Non-Cognitive Skills, Returns to School and the Gender Gap in Higher Education," *Economics of Education Review*, 21: 589–598

Jacob, Brian A. and Tamara Wilder (2012) "Educational Expectations and Attainment," In *Whither Opportunity? Rising Inequality and the Uncertain Life Chances of Low-Income Children*, ed. by Greg J. Duncan and Richard J. Murnane. New York, NY: Russell Sage Press.

Kramarz, Francis, Machin, Stephen, and Amine Ouazad (2008) "What Makes a Test Score? The Respective Contributions of Pupils, Schools, and Peers in Achievement in English Primary Education," IZA Discussion Paper No. 3866.

Lai, Fang (2010) "Are Boys Left Behind? The Evolution of the Gender Achievement Gap in Beijing's Middle Schools," *Economics of Education Review*, 29(3): 383-399.

LoGerfo, L, A Nichols, D Chaplin (2006) *Gender Gaps in Math and Reading Gains during Elementary and High School by Race and Ethnicity*, The Urban Institute.

Machin, Steve, and Sarah McNally (2005). Gender and Student Achievement in English Schools. *Oxford Review of Economic Policy*, 21(3), 357–372.

National Center for Education Statistics (NCES). (2004). *The High School Transcript Study: A Decade of Change in Curricula and Achievement, 1990–2000.* NCES 2004–455, by Robert Perkins, Brian Kleiner, Stephen Roey, and Janis Brown. Project Officer: Janis Brown. Washington, DC: 2004.

Organization for Economic Cooperation and Development (OECD) (2008). "The Reversal of Gender Inequalities in Higher Education: An On-going Trend" by Stéphan Vincent-Lancrin, Chap. 10 in *Higher Education to 2030, Vol.1: Demography.*

Reynolds, J.R., and Pemberton, J. (2001). "Rising College Expectations Among Youth in the United States: A Comparison of the 1979 and 1997 NLSY". *Journal of Human Resources*, 36(4):703–726.

Stange, Kevin M. (2008) "An Empirical Investigation of the Option Value of College Enrollment," forthcoming in the *American Economic Journal.*

Stinebrickner, Todd and Ralph Stinebrickner (2009) "Learning about Academic Ability and The College Drop-Out Decision", NBER Working Paper 14810, National Bureau of Economic Research: Cambridge, MA .

Zafar, Basit (2011) "How Do College Students Form Expectations?" *Journal of Labor Economics* , 29(2): 301-348.

Table 1. Vocational Expectations of 12th Graders by Gender

Years	1976-1988		1989-1999		2000-2009	
Kind of work respondent thinks will be doing at age 30	Boys	Girls	Boys	Girls	Boys	Girls
Percentage in the Labor Force[a]	99.9	93.2	99.7	97.7	99.6	98.3
Laborer, Service worker, Sales clerk	2.1	7.5	1.6	4.1	2.4	4.2
Skilled or semi-skilled worker, Protective services (including Military), Farm owner	37.6	4.4	30.5	4.8	27.4	5.0
Clerical or office worker	1.7	21.0	1.2	9.0	0.9	2.7
Owner of small business, Sales Representative, Manager or Administrator	17.3	14.4	17.9	13.5	16.9	12.8
Professional without doctoral degree[b]	27.9	37.4	32.1	42.2	36.1	48.3
Professional with doctoral degree (or equiv)[c]	13.5	15.3	16.7	26.4	16.4	27.1
Total in the Labor Force	100.0	100.0	100.0	100.0	100.0	100.0
Number of observations	18369	19343	11667	12560	9242	10396

Note: Module 4 respondents were asked "What kind of work do you think that you will be doing when you are 30 years old? Mark the one that comes closest to what you expect to be doing." Sixteeen choices were possible including "Full-time Homemaker" and "Don't know"(6.7 percent of respondents). We regroup the 14 other choices into 6 categories for conciseness. More examples of specific jobs were given than reported.

[a] Computed as 100 minus the percentage of respondents choosing "Full-time homemaker or housewife".

[b] Registered Nurse, Librarian, Engineer, Architect, Social Worker, Technician, Accountant, Actor, Artist, Musician given as examples.

[c] Lawyer, Physician, Dentist, Scientist, College Professor given as examples.

Table 2. Difference-and-Differences Estimates in Academic Performance and Plans for the Future - 12th graders

Time period	1976-1988	1989-1999	Change over time (2)-(1)	2000-2009	Change over time (4)-(2)	Change over time (4)-(1)
	(1)	(2)	(3)	(4)	(5)	(6)
A: Average grades						
Girls	**3.004**	**3.106**	0.102	**3.218**	0.112	0.214
	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
Boys	**2.804**	**2.907**	0.103	**3.030**	0.123	0.225
	(0.002)	(0.002)	(0.003)	(0.003)	(0.004)	(0.003)
Difference	0.200	0.199	-0.001	0.189	-0.010	-0.011
	(0.003)	(0.003)	(0.004)	(0.004)	(0.005)	(0.005)
B: Proportion with A grade						
Girls	**0.100**	**0.143**	0.043	**0.192**	0.048	0.091
	(0.001)	(0.004)	(0.001)	(0.004)	(0.002)	(0.002)
Boys	**0.069**	**0.099**	0.030	**0.137**	0.038	0.068
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
Difference	0.032	0.044	0.012	0.054	0.011	0.023
	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.002)
C: Proportion definitely will attend graduate or professional school[a]						
Girls	**0.101**	**0.205**	0.104	**0.249**	0.044	0.147
	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Boys	**0.099**	**0.150**	0.051	**0.168**	0.018	0.069
	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
Difference	0.003	0.055	0.053	0.081	0.026	0.078
	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.002)
Number of observations	207,152	160,403		118,173		

Note: Self-reported grades in 9 categories (D, C-,C,C+,B-,B,B+,A-,A) are translated into the numbers 1, 1.7,2,2.3,2.7,3,3.3,3.7 and 4 following standard institutional practice.

[a] The numbers for other post-secondary choices are reported in Table 2.

Table 3. Means of Selected Core Variables by Gender - 12th graders

Core Variables	1976-1988			1989-1999			2000-2009		
	Boys	Girls		Boys	Girls		Boys	Girls	
Subjective School Ability Index	0.652	0.651		0.658	0.654		0.664	0.658	*
Education Aspirations: want to attend (binary dummy)									
Army	0.203	0.092	*	0.177	0.079	*	0.179	0.078	*
Vocational	0.284	0.219	*	0.207	0.141	*	0.203	0.124	*
Two-year college	0.206	0.293	*	0.214	0.256	*	0.240	0.266	*
Four-year college	0.635	0.650	*	0.744	0.810	*	0.773	0.850	*
Graduate or professional	0.416	0.432	*	0.529	0.613	*	0.519	0.625	*
Educational Expectations: index of likeness to attend[a]									
Army	0.281	0.102	*	0.215	0.078	*	0.202	0.079	*
Vocational	0.319	0.264	*	0.268	0.210	*	0.274	0.208	*
Two-year college	0.338	0.364	*	0.362	0.370	*	0.383	0.386	
Four-year college	0.584	0.585		0.702	0.758	*	0.737	0.816	*
Graduate or professional	0.389	0.385	*	0.471	0.530	*	0.490	0.571	*
High school program: Academic	0.487	0.514	*	0.550	0.611	*	0.518	0.589	*
General	0.300	0.307	*	0.283	0.272	*	0.328	0.298	*
Vocational	0.155	0.120	*	0.107	0.068	*	0.081	0.049	*
Other	0.059	0.060		0.059	0.049	*	0.073	0.065	*
Cigarettes smoking: Less than one-half pack	0.212	0.260	*	0.258	0.260		0.217	0.201	*
Alcohol binging last two weeks: Two to nine times	0.307	0.167	*	0.231	0.127	*	0.197	0.121	*
Father not same household	0.169	0.185	*	0.201	0.228	*	0.207	0.244	*
Mother working all the time	0.201	0.234	*	0.353	0.398	*	0.462	0.495	*
Father: Some high school	0.145	0.154	*	0.101	0.110	*	0.098	0.108	*
Father: Completed College	0.190	0.176	*	0.230	0.214	*	0.253	0.225	*
Mother: Some high school	0.126	0.149	*	0.082	0.101	*	0.071	0.082	*
Mother: Completed College	0.164	0.146	*	0.234	0.211	*	0.290	0.257	*
Works over school year	0.848	0.798	*	0.801	0.792	*	0.755	0.756	
Number of observations	74230	79942		60469	66875		50549	57202	

Note: Asterisk indicates statistically significant gender difference at the 5% level. Means of other variables and other categories are reported in Table A2.

[a] The four categories of likeliness are: definitively won't, probably won't, probably will, definitively will.

Table 4a. Selected Coefficients of LPM on Specific Grades - 12th graders - 2000-2009

Dependent variable: A (93-100)	Specification 1				Specification 2			
Explanatory Variables	Boys		Girls		Boys		Girls	
Subjective School Ability					44.457	(53.75)	70.881	(72.68)
Education Aspirations: want to attend (binary dummy)								
Army					-2.522	(-5.06)	-2.214	(-3.11)
Vocational					-0.138	(-0.29)	1.152	(1.89)
Two-year college					0.235	(0.54)	-0.198	(-0.42)
Four-year college					-1.739	(-3.98)	-1.321	(-2.50)
Graduate or professional					1.644	(4.34)	2.071	(5.04)
Educational Expectations: index of likeness to attend[a]								
Army	-2.328	(-4.47)	-0.132	(-0.16)	0.310	(0.47)	2.342	(2.40)
Vocational	-3.945	(-7.18)	-3.929	(-6.52)	-3.522	(-5.62)	-3.938	(-5.67)
Two-year college	-9.946	(-21.99)	-11.409	(-24.47)	-7.536	(-14.45)	-7.042	(-12.68)
Four-year college	3.672	(6.07)	4.384	(6.40)	0.150	(0.23)	0.572	(0.77)
Graduate or professional	13.711	(24.63)	10.543	(18.49)	8.660	(14.07)	4.102	(6.43)
High school program:								
Academic	5.905	(9.39)	9.090	(12.99)	1.287	(2.08)	2.010	(2.97)
General	-1.433	(-2.30)	-0.288	(-0.41)	-2.827	(-4.65)	-1.885	(-2.78)
Vocational	2.468	(3.12)	4.776	(4.87)	-0.251	(-0.33)	1.334	(1.42)
Cigarettes smoking: Less than one-half pack	-4.928	(-12.39)	-7.789	(-17.68)	-3.635	(-9.39)	-6.107	(-14.48)
Alcohol binging last 2 weeks:								
Two to nine times	-4.984	(-12.01)	-5.247	(-9.81)	-4.664	(-11.57)	-4.007	(-7.83)
Father not same household	-1.228	(-3.07)	-2.319	(-5.69)	-0.870	(-2.24)	-1.895	(-4.87)
Mom working: All the time	-3.855	(-8.26)	-4.764	(-9.46)	-3.513	(-7.75)	-3.844	(-7.98)
Father: Completed College	0.910	(2.04)	2.883	(6.00)	0.133	(0.31)	1.708	(3.72)
Mother: Completed college	1.420	(3.30)	2.945	(6.27)	0.662	(1.58)	1.445	(3.22)
Works over school year	-3.092	(-2.55)	0.325	(0.22)	-2.206	(-1.87)	2.141	(1.50)
Constant	25.242	(28.44)	28.758	(29.58)	-1.310	(-1.25)	-13.363	(-11.36)
R-squared	0.116		0.126		0.166		0.202	
Number of observations	49328		56156		49328		56156	

Note: Dependent variables is set to 100 if the student has a GPA of 4, and to 0 otherwise. The base group for Alcohol binging and cigarettes smoking is none, mom working is not working, father's and mother's education is high school, for high school program is other. The coefficients of the other variables and categories included in the regression are reported in Appendix Table A3.

[a] The four categories of likeliness are: definitively won't, probably won't, probably will, definitively will.

Table 4b. Selected Coefficients of LPM on Specific Grades - 12th graders - 2000-2009

Dependent variable: C+(77-79)	Specification 1				Specification 2			
Explanatory Variables	Boys		Girls		Boys		Girls	
Subjective School Ability					-20.698	(-28.22)	-20.586	(-31.84)
Education Aspirations: want to attend (binary dummy)								
Army					-0.290	(-0.66)	0.287	(0.61)
Vocational					0.241	(0.56)	0.733	(1.81)
Two-year college					-0.285	(-0.75)	0.488	(1.56)
Four-year college					1.298	(3.35)	-0.761	(-2.17)
Graduate or professional					-0.773	(-2.30)	-0.147	(-0.54)
Educational Expectations: index of likeness to attend[a]								
Army	2.328	(5.15)	0.469	(0.89)	2.296	(3.96)	0.134	(0.21)
Vocational	0.473	(0.99)	0.876	(2.27)	0.125	(0.23)	0.175	(0.38)
Two-year college	5.951	(15.15)	3.446	(11.57)	4.884	(10.56)	1.995	(5.42)
Four-year college	-5.343	(-10.17)	-3.373	(-7.70)	-4.019	(-6.92)	-1.387	(-2.82)
Graduate or professional	-5.101	(-10.55)	-3.123	(-8.57)	-2.762	(-5.06)	-1.606	(-3.80)
High school program: Academic	-4.472	(-8.19)	-5.978	(-13.36)	-2.344	(-4.28)	-3.900	(-8.69)
General	-0.181	(-0.33)	-1.912	(-4.21)	0.450	(0.83)	-1.421	(-3.16)
Vocational	-1.720	(-2.50)	-1.507	(-2.41)	-0.456	(-0.67)	-0.635	(-1.02)
Cigarettes smoking: Less than one-half pack	1.865	(5.40)	2.941	(10.44)	1.275	(3.71)	2.431	(8.69)
Alcohol binging last 2 weeks: Two to nine times	1.533	(4.26)	1.709	(5.00)	1.373	(3.84)	1.362	(4.02)
Father not same household	1.044	(3.00)	1.728	(6.64)	0.870	(2.52)	1.587	(6.15)
Mother working: All the time	1.542	(3.81)	1.509	(4.69)	1.380	(3.43)	1.269	(3.97)
Father: Some high school	0.826	(1.64)	2.117	(5.59)	0.624	(1.25)	1.876	(5.00)
Mother: Some high school	1.500	(2.59)	0.712	(1.67)	1.290	(2.25)	0.547	(1.29)
Works over school year	0.996	(0.95)	2.283	(2.40)	0.601	(0.58)	1.713	(1.82)
Constant	7.884	(10.23)	6.896	(11.10)	20.066	(21.54)	19.544	(25.05)
R-squared	0.051		0.051		0.066		0.069	
Number of observations	49328		56156		49328		56156	

Note: Dependent variables is set to 100 if the student has a GPA of 2.3, and to 0 otherwise. T-statistics are in parentheses. The base group for alcohol binging and cigarettes smoking is none, mom working is not working, father's and mother's education is high school, for high school program is other. The coefficients of the other variables and categories included in the regression are reported in Appendix Table A3.

[a] The four categories of likeliness are: definitively won't, probably won't, probably will, definitively will.

Table 5. Detailed Decomposition Results - 12th Graders
Percentage Female/Male Difference for Selected GPA Levels

	Specification 1				Specification 2			
2000-2009	A (93-100): 4		C+ (77-79): 2.3		A (93-100): 4		C+ (77-79): 2.3	
Total Differential	6.063	(0.007)	-3.152	(0.005)	6.063	(0.007)	-3.152	(0.005)
Composition Effects:								
Total Explained	2.395	(0.004)	-1.224	(0.003)	1.034	(0.004)	-0.590	(0.003)
Race, SMSA	-0.293	(0.001)	0.254	(0.001)	-0.232	(0.001)	0.224	(0.001)
Own School Ability					-0.463	(0.002)	0.215	(0.001)
Smoking, Binging	0.562	(0.001)	-0.198	(0.001)	0.489	(0.001)	-0.163	(0.001)
Family Background	-0.348	(0.001)	0.194	(0.001)	-0.233	(0.001)	0.143	(0.001)
Work	0.060	(0.001)	-0.026	(0.001)	0.079	(0.001)	-0.038	(0.001)
High school program	0.385	(0.001)	-0.257	(0.001)	0.185	(0.001)	-0.163	(0.001)
Educ. Expectations	2.029	(0.003)	-1.192	(0.003)	1.207	(0.003)	-0.809	(0.003)
Specification Error	-0.005	(0.008)	0.102	(0.006)	-0.048	(0.007)	0.112	(0.006)
Educational Responses:								
Total Unexplained	3.593	(0.007)	-1.933	(0.005)	4.802	(0.007)	-2.489	(0.005)
Race, SMSA	-1.414	(0.013)	-0.493	(0.009)	-1.509	(0.013)	-0.507	(0.009)
Own School Ability					-0.089	(0.001)	0.005	(0.000)
Smoking, Binging	-0.467	(0.005)	0.335	(0.003)	-0.333	(0.005)	0.320	(0.003)
Family Background	1.591	(0.027)	0.313	(0.019)	1.086	(0.025)	0.706	(0.019)
Work	1.590	(0.013)	0.303	(0.009)	1.616	(0.012)	0.087	(0.009)
High school program	2.646	(0.028)	-0.905	(0.020)	1.406	(0.027)	-1.503	(0.020)
Educ. Expectations	-0.117	(0.002)	0.153	(0.002)	1.081	(0.021)	-1.331	(0.016)
Constant	-0.236	(0.043)	-1.639	(0.030)	1.542	(0.045)	-0.267	(0.034)
Reweighting Error	0.080	(0.003)	-0.097	(0.001)	0.276	(0.003)	-0.185	(0.001)

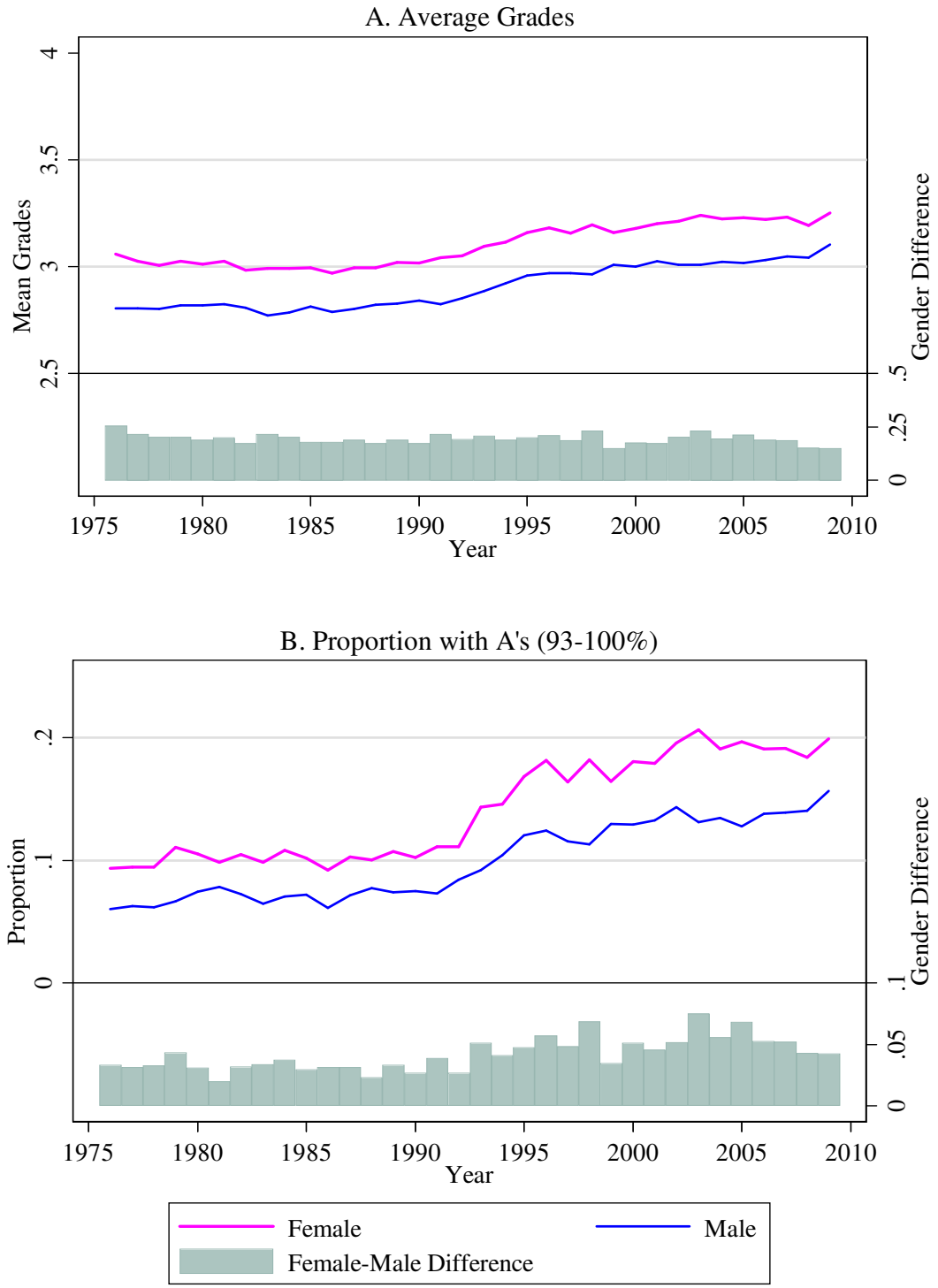
Note: Standard errors are in parentheses. Reweighted decomposition follows methodology of section 4. In specification 2, educational aspirations are included among the variables in the educational expectations category.

Table 6. Detailed Decomposition Results - 10th and 8th Graders
Female/Male Percentage Difference for Selected GPA Levels

	10th graders				8th graders			
2000-2009	A (93-100): 4		C (73-76): 2		A (93-100): 4		C+ (77-79): 2.3	
Total Differential	4.740	(0.006)	-2.586	(0.005)	5.957	(0.007)	-1.481	(0.005)
Composition Effects:								
Total Explained	1.348	(0.004)	-1.424	(0.003)	2.207	(0.004)	-0.864	(0.003)
Race, SMSA	-0.138	(0.001)	0.042	(0.000)	-0.158	(0.001)	0.133	(0.000)
School Hard, Held back	-0.654	(0.001)	0.146	(0.001)	-0.632	(0.001)	0.096	(0.001)
Misbehavior, Smoking	0.577	(0.001)	-0.358	(0.001)	1.286	(0.002)	-0.542	(0.002)
Family Background	-0.230	(0.001)	0.148	(0.001)	-0.308	(0.001)	0.180	(0.001)
Work	0.006	(0.001)	0.038	(0.001)	-0.024	(0.001)	-0.059	(0.001)
High school program	0.527	(0.001)	-0.224	(0.001)	0.518	(0.001)	-0.138	(0.001)
Educ. Expectations	1.260	(0.002)	-1.215	(0.002)	1.525	(0.003)	-0.534	(0.002)
Specification Error	0.153	(0.006)	-0.343	(0.005)	0.187	(0.007)	-0.320	(0.006)
Educational Responses:								
Total Unexplained	3.231	(0.006)	-0.721	(0.004)	3.068	(0.007)	-0.090	(0.005)
Race, SMSA	-0.746	(0.012)	-0.548	(0.009)	-0.690	(0.013)	-0.382	(0.009)
School Hard, Held back	-0.194	(0.002)	0.143	(0.001)	-0.097	(0.002)	0.137	(0.002)
Misbehavior, Smoking	-0.456	(0.004)	0.142	(0.003)	-0.039	(0.004)	0.347	(0.003)
Family Background	0.754	(0.019)	-0.360	(0.014)	1.543	(0.021)	0.731	(0.015)
Work	0.171	(0.005)	-0.190	(0.004)	0.462	(0.006)	-0.043	(0.004)
High school program	0.016	(0.012)	0.180	(0.009)	1.534	(0.008)	-0.071	(0.006)
Educ. Expectations	0.027	(0.002)	0.218	(0.002)	-0.215	(0.002)	0.102	(0.002)
Constant	3.658	(0.026)	-0.307	(0.019)	0.569	(0.027)	-0.910	(0.020)
Reweighting Error	0.315	(0.003)	-0.099	(0.001)	0.495	(0.003)	-0.208	(0.001)

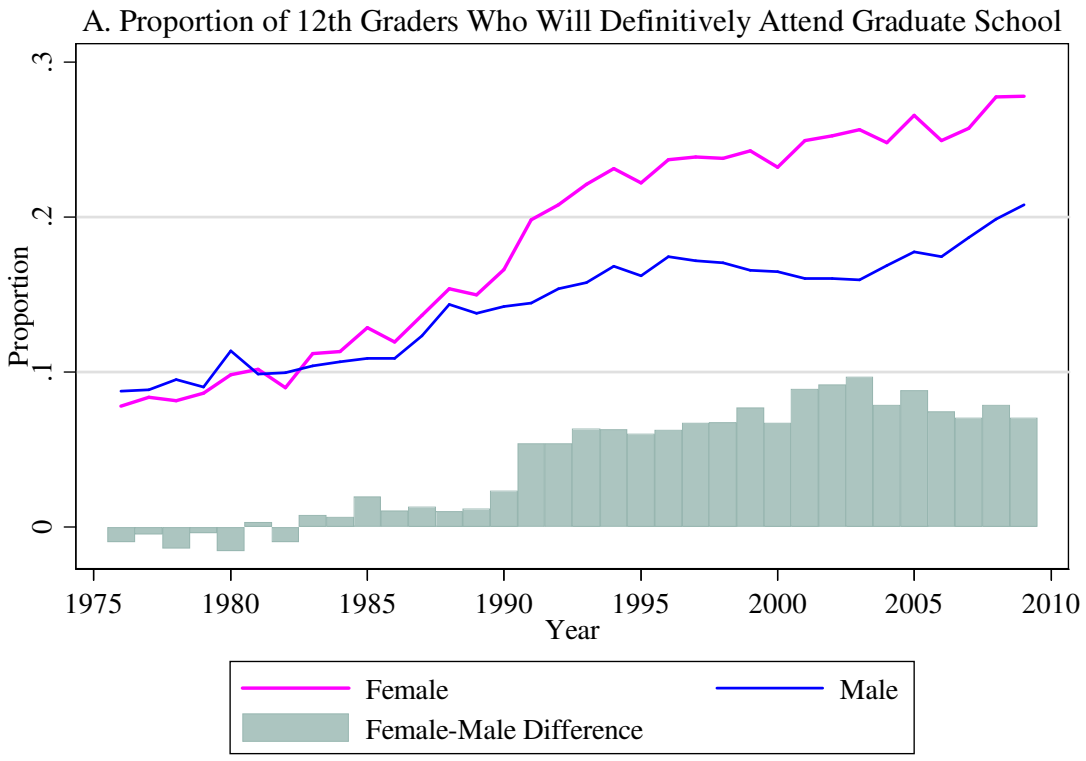
Note: Standard errors are in parentheses. Reweighted decomposition follows methodology of section 4.

Figure 1. Self-Reported Grades of High School Seniors by Gender



Note: In Figure 1A, self-reported grades in 9 categories (D, C-,C,C+,B-,B,B+,A-,A) are translated into the numbers 1, 1.7,2,2.3,2.7,3,3.3,3.7 and 4 following standard institutional practice.

Figure 2. Educational Expectations of High School Students



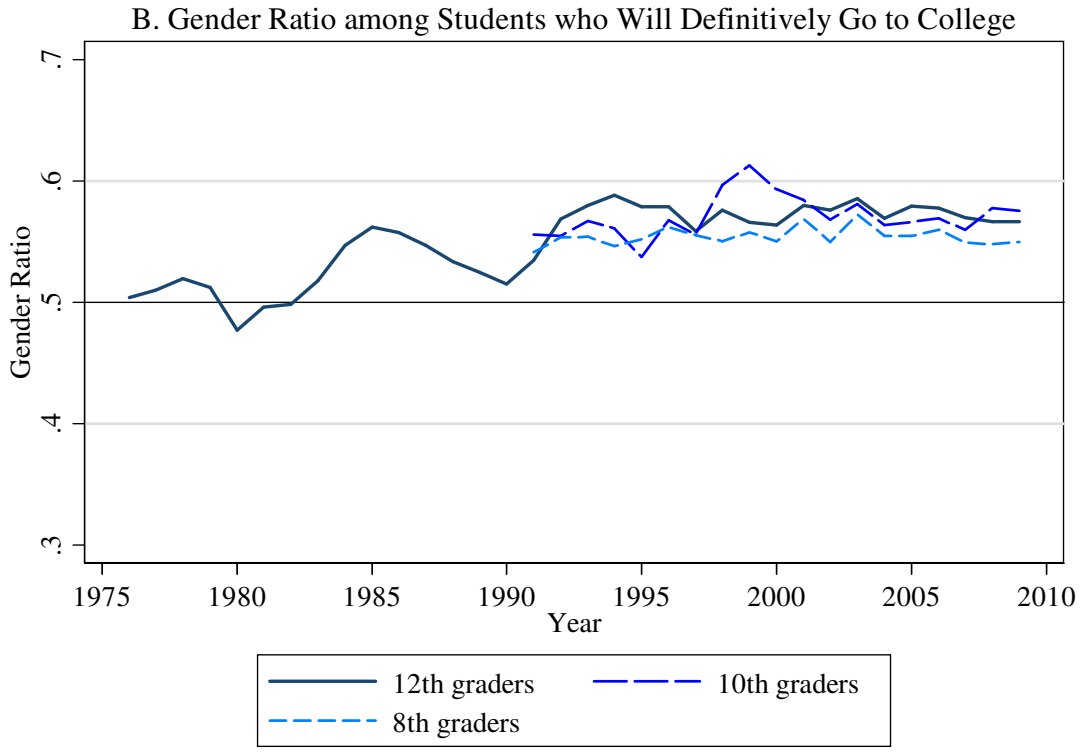
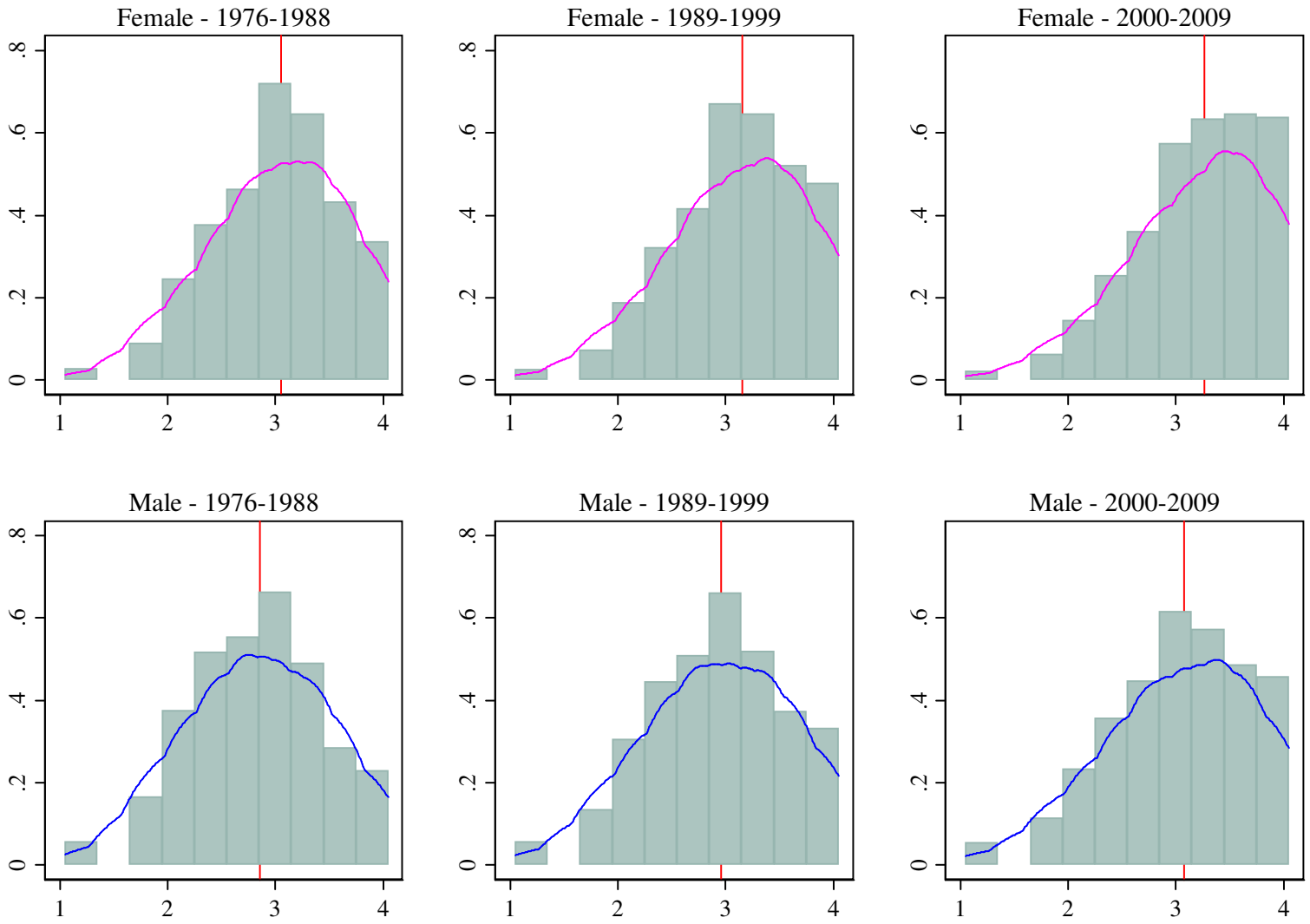


Figure 3. Male and Female Densities of Self-Reported Grades among 12th Graders

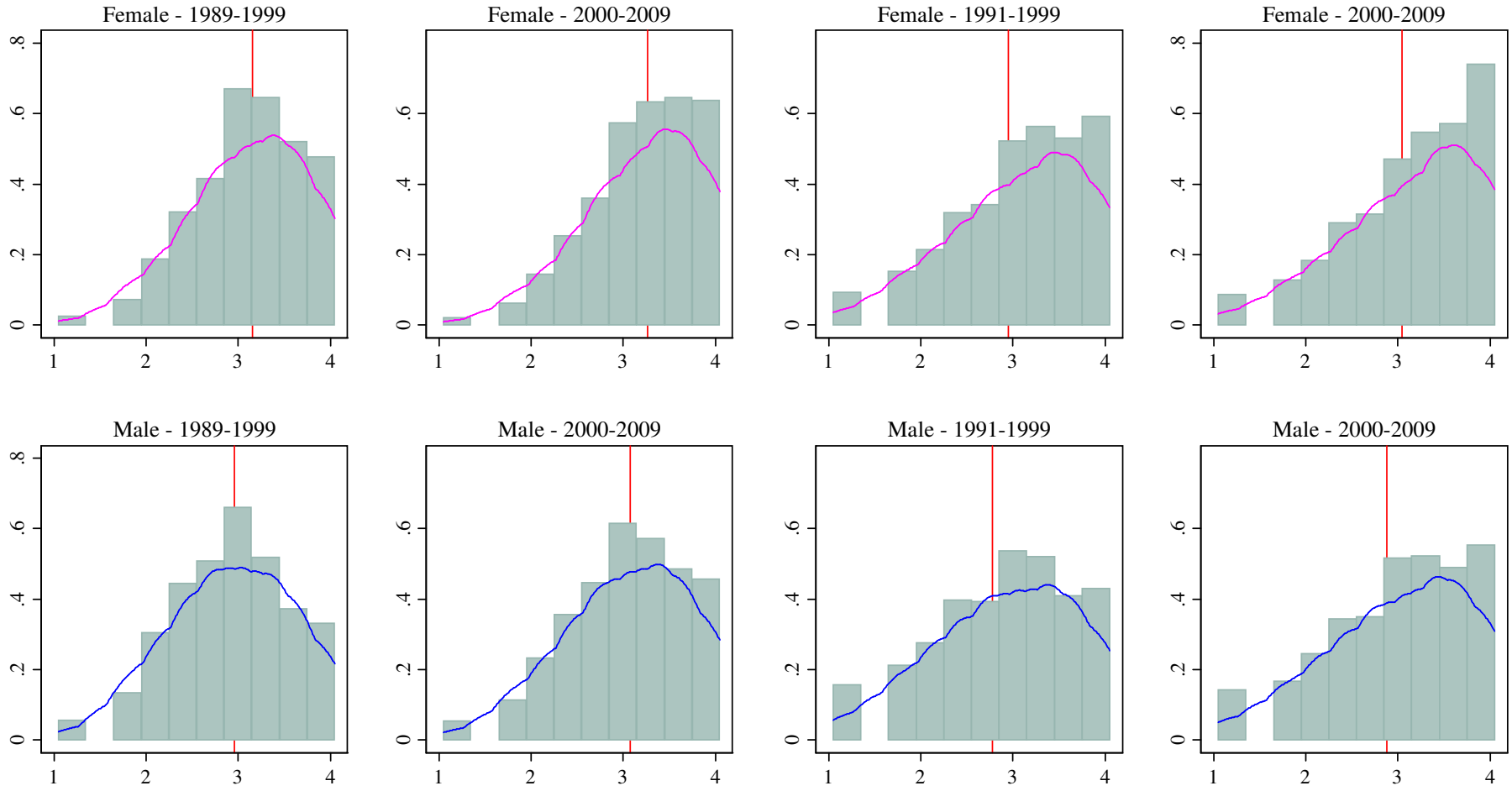


Note: Average grades is indicated by vertical line. Histogram which corresponds to actual data is overlaid with a kernel density.

Figure 4. Male and Female Densities of Self-Reported Grades among 10th and 8th Graders

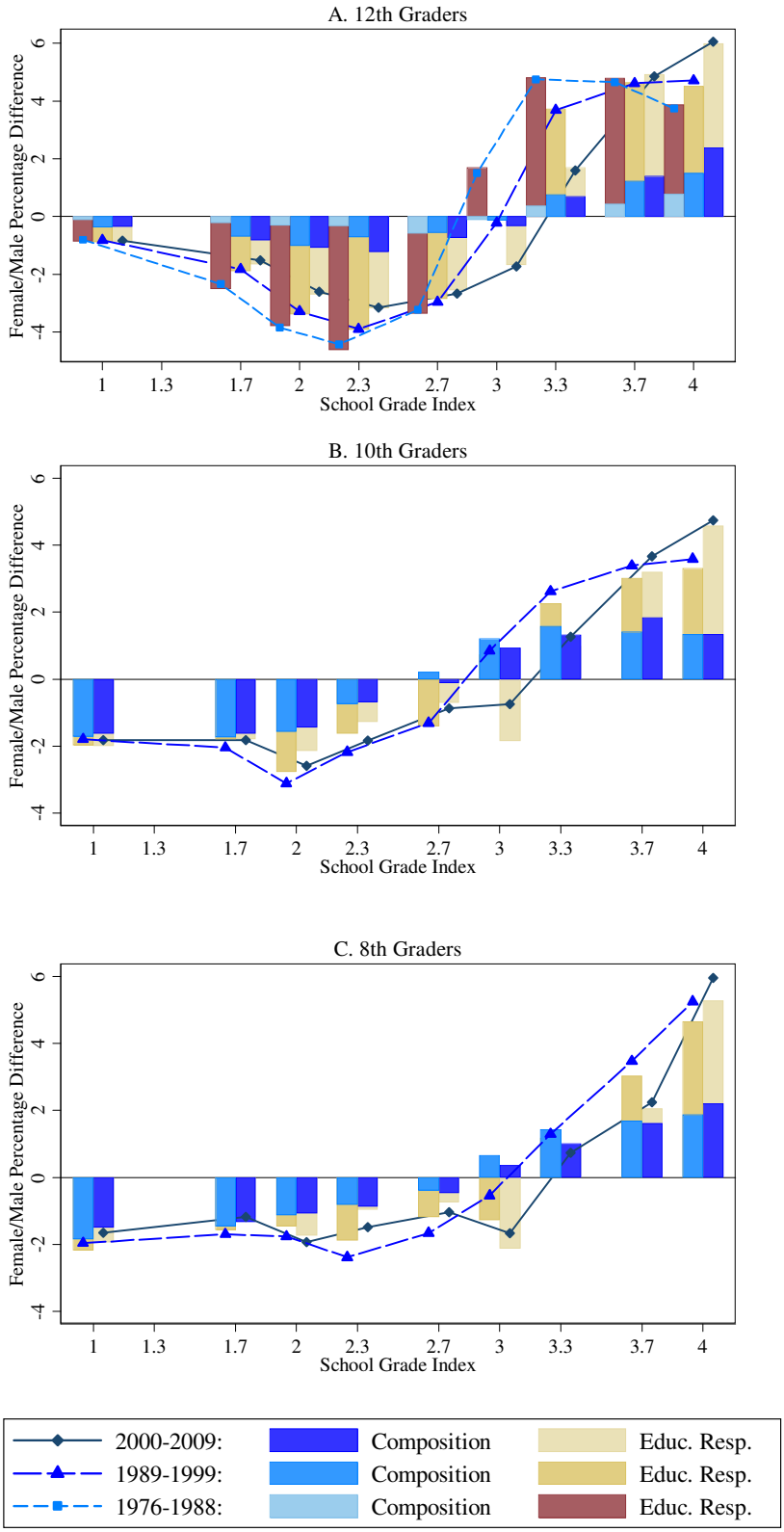
A. 10th Graders

B. 8th Graders



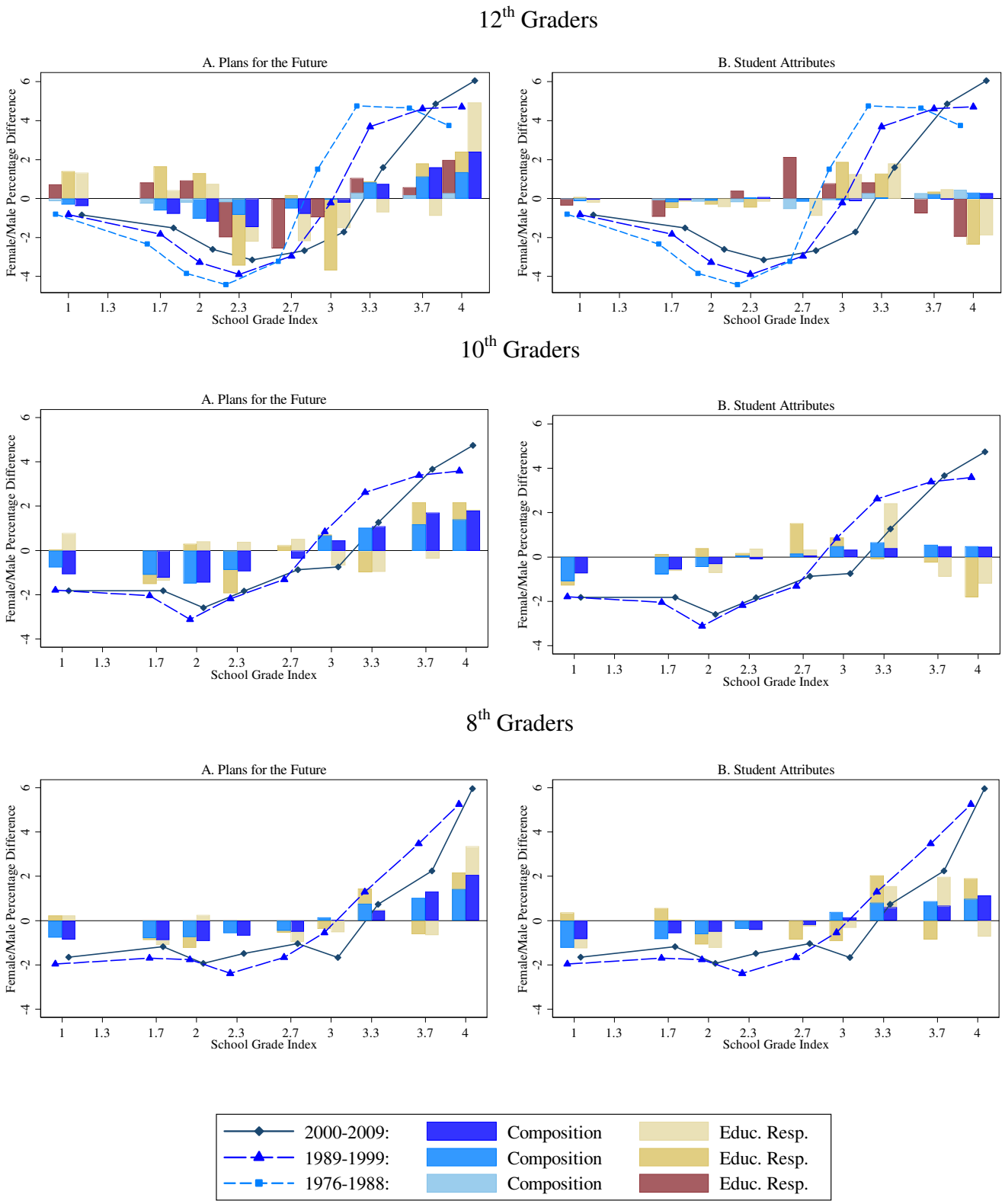
Note: Average grades is indicated by vertical line. Histogram which corresponds to actual data is overlaid with a kernel density.

Figure 5. Aggregate Decomposition of Gender Differences in GPA Levels



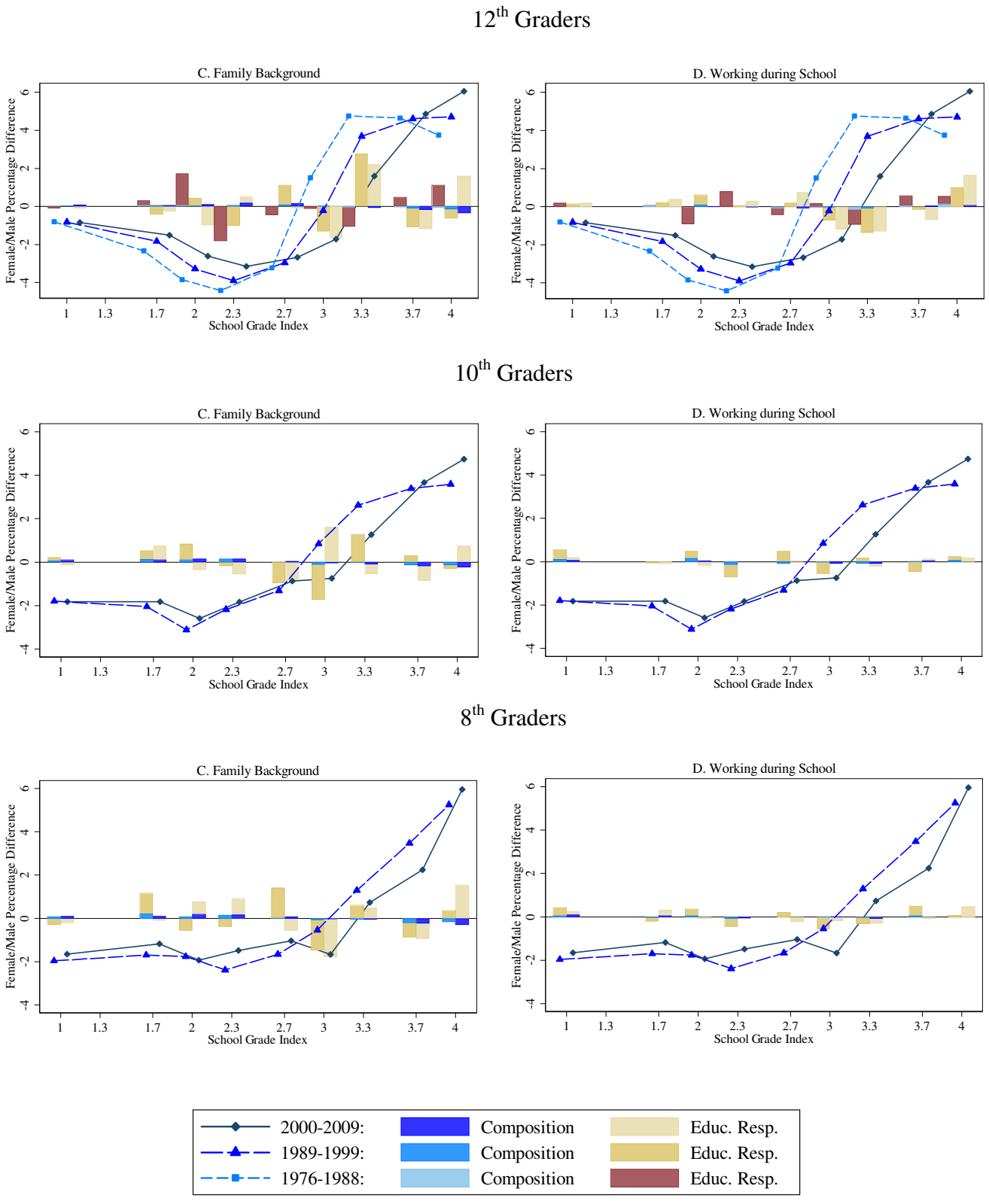
Note: Self-reported grades in 9 categories (D, C-,C,C+,B-,B,B+,A-,A) are translated into the numbers 1, 1.7,2,2.3,2.7,3,3.3,3.7 and 4 following standard institutional practice. The lines in the figure show the raw gender differences; the different shaded bars indicate attributable to composition or educational responses effects.

Figure 6a. Detailed Decomposition of Gender Differences in GPA Levels



Note: Self-reported grades in 9 categories (D, C-,C,C+,B-,B,B+,A-,A) are translated into the numbers 1, 1.7,2,2.3,2.7,3,3.3,3.7 and 4 following standard institutional practice. . The lines in the figure show the raw gender differences; the different shaded bars indicate attributable to composition or educational responses effects.

Figure 6b. Detailed Decomposition of Gender Differences in GPA Levels



Note: Self-reported grades in 9 categories (D, C-,C,C+,B-,B,B+,A-,A) are translated into the numbers 1, 1.7,2,2.3,2.7,3,3.3,3.7 and 4 following standard institutional practice. . The lines in the figure show the raw gender differences; the different shaded bars indicate attributable to composition or educational responses effects.

Table A1. Proportion and Average Wages of 25 to 39 years old from MORG-CPS and IPUMS-USA in the same occupational categories as MTF

	Years	1976-1988				1989-1999				2000-2009			
		Men		Women		Men		Women		Men		Women	
	Proportion	%	Wage	%	Wage	%	Wage	%	Wage	%	Wage	%	Wage
	in the Labor Force	95.0		68.83		93.46		75.01		92.1		74.7	
	of Employed Workers in the Occupation Category												
1	Laborer, Service worker, Sales clerk	12.1	11.60	20.7	8.52	15.0	10.03	20.6	8.67	17.5	11.19	23.2	9.11
2	Skilled or semi-skilled worker, Protective services (including Military), Farm owner	39.5	15.76	11.1	10.46	37.9	13.86	9.2	7.87	34.3	14.16	6.6	10.79
3	Clerical or office worker	5.6	15.65	30.6	11.71	5.8	13.80	25.9	9.19	5.9	13.32	21.0	12.07
4	Owner of small business[a], Sales Representative, Manager or Administrator	26.1	19.91	16.2	15.21	23.0	19.08	17.7	8.39	21.6	22.06	17.0	18.16
5	Professional without doctoral degree	14.4	20.31	20.1	16.01	16.1	20.01	25.0	11.54	18.3	21.88	29.6	17.89
6	Professional with doctoral degree (or equiv)	2.3	21.79	1.4	19.47	2.2	23.59	1.7	13.33	2.5	25.10	2.6	23.14
	Total Employed	100.0		100.0		100.0		100.0		100.0		100.0	

Note: Percentage of the workforce in the military is from the IPUMS-USA (1970, 1980, 1990, 2000) and the American Community Surveys (2000-2009). Percentages in the other occupations and wages are average occupational real hourly wages in 2000 dollars from the MORG-CPS for the corresponding years.

[a] Self-employed in class of worker.

Table A2. Means of Select Core Variables by Gender - 12th graders

Core Variables	1976-1988			1989-1999			2000-2009		
	Boys	Girls		Boys	Girls		Boys	Girls	
Self-reported Grades:									
D (69 or below): 1	0.014	0.006	*	0.015	0.006	*	0.014	0.005	*
C- (70-72): 1.7	0.045	0.022	*	0.036	0.018	*	0.031	0.016	*
C (73-76): 2	0.104	0.065	*	0.086	0.053	*	0.065	0.039	*
C+ (77-79): 2.3	0.146	0.101	*	0.126	0.087	*	0.099	0.068	*
B- (80-82): 2.7	0.165	0.133	*	0.149	0.120	*	0.129	0.103	*
B (83-86): 3	0.203	0.218	*	0.202	0.200		0.187	0.169	
B+ (87-89): 3.3	0.154	0.201	*	0.160	0.197	*	0.175	0.190	*
A- (90-92): 3.7	0.093	0.140	*	0.119	0.165	*	0.154	0.203	*
A (93-100): 4	0.076	0.113	*	0.108	0.155	*	0.147	0.207	*
Race: Black	0.083	0.097	*	0.085	0.105	*	0.084	0.107	*
Live in MSA	0.683	0.683		0.731	0.738	*	0.755	0.759	
Cigarettes smoking per day: None	0.715	0.673	*	0.678	0.694	*	0.749	0.774	*
Less than one-half pack	0.212	0.260	*	0.258	0.260		0.217	0.201	*
One to 1½ pack	0.070	0.064	*	0.060	0.044	*	0.030	0.023	*
Two packs or more	0.003	0.002	*	0.005	0.002	*	0.004	0.002	*
Alcohol binging last 2 weeks: None	0.534	0.713	*	0.635	0.775	*	0.686	0.780	*
Once	0.129	0.111	*	0.109	0.092	*	0.100	0.094	*
Two to nine times	0.307	0.167	*	0.231	0.127	*	0.197	0.121	*
Ten or more times	0.030	0.008	*	0.025	0.006	*	0.017	0.006	*
Siblings not same household	0.243	0.235	*	0.326	0.312	*	0.329	0.311	*
Siblings: None	0.046	0.042		0.056	0.051	*	0.060	0.052	
One	0.268	0.256	*	0.326	0.311	*	0.323	0.300	*
Two	0.262	0.254		0.271	0.270		0.282	0.272	*
Three or more	0.424	0.449	*	0.344	0.365	*	0.332	0.374	*
Mother not same household	0.075	0.066	*	0.098	0.084	*	0.096	0.089	*
Father not same household	0.169	0.185	*	0.201	0.228	*	0.207	0.244	*
Mom working: No	0.312	0.299	*	0.198	0.184	*	0.146	0.140	*
Some of the time	0.312	0.302	*	0.254	0.242	*	0.206	0.196	*
Most of the time	0.175	0.164	*	0.195	0.176	*	0.185	0.170	*
All the time	0.201	0.234	*	0.353	0.398	*	0.462	0.495	*
Father education: less than primary	0.062	0.076	*	0.034	0.046	*	0.031	0.041	*
Some high school	0.145	0.154	*	0.101	0.110	*	0.098	0.108	*
Completed high school	0.320	0.320		0.285	0.298	*	0.288	0.304	*
Some college	0.156	0.153		0.195	0.191		0.182	0.180	
Completed College	0.190	0.176	*	0.230	0.214	*	0.253	0.225	*
Graduate or prof.	0.127	0.121	*	0.155	0.141	*	0.147	0.142	

Note: Asterisk indicates statistically significant gender difference at the 5% level. This table completes Table 3 by presenting the means of the variables not reported there but included in the regression.

(continued next page)

Table A2. Means of Select Core Variables by Gender - 12th graders (continued)

Core Variables	1976-1988			1989-1999			2000-2009		
	Boys	Girls		Boys	Girls		Boys	Girls	
Mother education: less than primary	0.032	0.042	*	0.027	0.035	*	0.027	0.034	*
Some high school	0.126	0.149	*	0.082	0.101	*	0.071	0.082	*
Completed high school	0.441	0.416	*	0.339	0.333		0.277	0.280	
Some college	0.166	0.175	*	0.210	0.215		0.210	0.222	
Completed College	0.164	0.146	*	0.234	0.211	*	0.290	0.257	*
Graduate or prof.	0.071	0.072		0.108	0.104		0.125	0.124	
Works over school year	0.848	0.798	*	0.801	0.792	*	0.755	0.756	
Average hours of work: None	0.177	0.222	*	0.223	0.223		0.271	0.260	*
5 or less hours	0.101	0.101		0.097	0.095		0.097	0.096	
6 to 10 hours	0.099	0.103		0.095	0.107		0.097	0.107	*
11 to 20 hours	0.262	0.300	*	0.260	0.303	*	0.252	0.281	*
21 to 30 hours	0.234	0.203	*	0.220	0.205	*	0.194	0.191	
More than 30 hours	0.128	0.071	*	0.104	0.067	*	0.090	0.065	*
job: None	0.227	0.281	*	0.269	0.283	*	0.311	0.314	
$1-5	0.037	0.046		0.018	0.022		0.010	0.010	
$6-10	0.040	0.045		0.025	0.028		0.036	0.044	*
$11-50	0.289	0.326	*	0.176	0.215	*	0.114	0.140	*
$51-75	0.253	0.200	*	0.128	0.151	*	0.086	0.106	*
$76-125	0.106	0.080	*	0.222	0.209	*	0.197	0.213	*
$126+	0.047	0.022	*	0.162	0.092	*	0.246	0.174	*

presenting the means of the variables not reported there but included in the regression.

Table A3a. Selected Coefficients of LPM on Specific Grades - 12th graders - 2000-2009

Dependent variable: A (93-100)	Specification 1				Specification 2				Specification 3			
Explanatory Variables	Boys		Girls		Boys		Girls		Boys		Girls	
Subjective School Ability					44.457	(53.75)	70.881	(72.68)				
Education Aspirations: want to attend (binary dummy)												
Army					-2.522	(-5.06)	-2.214	(-3.11)				
Vocational					-0.138	(-0.29)	1.152	(1.89)				
Two-year college					0.235	(0.54)	-0.198	(-0.42)				
Four-year college					-1.739	(-3.98)	-1.321	(-2.50)				
Graduate or professional					1.644	(4.34)	2.071	(5.04)				
Educational Expectations: index of likeness to attend[a]												
Army	-2.328	(-4.47)	-0.132	(-0.16)	0.310	(0.47)	2.342	(2.40)				
Vocational	-3.945	(-7.18)	-3.929	(-6.52)	-3.522	(-5.62)	-3.938	(-5.67)				
Two-year college	-9.946	(-21.99)	-11.409	(-24.47)	-7.536	(-14.45)	-7.042	(-12.68)				
Four-year college	3.672	(6.07)	4.384	(6.40)	0.150	(0.23)	0.572	(0.77)				
Graduate or professional	13.711	(24.63)	10.543	(18.49)	8.660	(14.07)	4.102	(6.43)				
High school program:												
Academic	5.905	(9.39)	9.090	(12.99)	1.287	(2.08)	2.010	(2.97)	11.151	(17.88)	14.138	(20.31)
General	-1.433	(-2.30)	-0.288	(-0.41)	-2.827	(-4.65)	-1.885	(-2.78)	-0.695	(-1.10)	0.4851	(0.67)
Vocational	2.468	(3.12)	4.776	(4.87)	-0.251	(-0.33)	1.334	(1.42)	-0.363	(-0.46)	2.7861	(2.82)
Other (base)												
Race: Black	-7.534	(-13.06)	-10.330	(-18.51)	-5.828	(-10.37)	-8.729	(-16.35)	-6.924	(-11.80)	-9.287	(-16.49)
SMSA	-5.404	(-14.96)	-8.525	(-22.02)	-5.270	(-15.02)	-8.227	(-22.23)	-4.072	(-11.13)	-6.812	(-17.45)
Smoked cigarettes per day: None (base)												
Less than one-half pack	-4.928	(-12.39)	-7.789	(-17.68)	-3.635	(-9.39)	-6.107	(-14.48)	-6.040	(-14.96)	-8.990	(-20.17)
One to 1½ pack	-3.911	(-4.23)	-7.383	(-6.64)	-2.513	(-2.79)	-5.382	(-5.06)	-5.990	(-6.38)	-9.730	(-8.65)
Two packs or more	13.153	(5.27)	-1.031	(-0.25)	14.792	(6.09)	2.882	(0.74)	12.306	(4.84)	-2.945	(-0.71)
Alcohol binging last 2 weeks: None (base)												
Once	-4.087	(-7.92)	-4.502	(-7.94)	-3.991	(-7.96)	-3.864	(-7.13)	-3.993	(-7.60)	-4.199	(-7.29)

Note: Dependent variables is set to 100 if the student has a GPA of 4, and to 0 otherwise. T-statistics are in parentheses. This table completes Table 4 by presenting all the coefficients not included in the regression.

Table A3a. Selected Coefficients of LPM on Specific Grades - 12th graders - 2000-2009 (continued)

Dependent variable: A (93-100)	Specification 1				Specification 2				Specification 3			
Explanatory Variables	Boys		Girls		Boys		Girls		Boys		Girls	
Alcohol binging last 2 weeks:												
Two to nine times	-4.984	(-12.01)	-5.247	(-9.81)	-4.664	(-11.57)	-4.007	(-7.83)	-5.040	(-11.94)	-4.902	(-9.03)
Ten or more times	-4.409	(-3.68)	-6.093	(-2.77)	-3.398	(-2.92)	-3.150	(-1.50)	-4.749	(-3.89)	-5.698	(-2.55)
Siblings not same household	0.078	(0.21)	0.243	(0.62)	0.382	(1.08)	0.295	(0.78)	0.093	(0.25)	0.008	(0.02)
Siblings: One (base)												
None	1.491	(2.10)	-1.119	(-1.38)	1.166	(1.69)	-1.847	(-2.39)	1.672	(2.31)	-0.718	(-0.87)
Two	-0.796	(-2.04)	-1.107	(-2.59)	-0.624	(-1.65)	-0.914	(-2.24)	-1.086	(-2.73)	-1.335	(-3.08)
Three or more	-1.615	(-4.17)	-1.797	(-4.36)	-1.493	(-3.96)	-1.476	(-3.75)	-2.214	(-5.63)	-2.594	(-6.22)
Don't know	-0.909	(-0.32)	-8.712	(-2.53)	1.816	(0.65)	-5.119	(-1.55)	-1.714	(-0.59)	-8.836	(-2.53)
Father not same household	-1.228	(-3.07)	-2.319	(-5.69)	-0.870	(-2.24)	-1.895	(-4.87)	-1.145	(-2.81)	-2.495	(-6.04)
Mother not same household	0.169	(0.31)	-1.853	(-3.05)	0.620	(1.16)	-1.460	(-2.51)	-0.790	(-1.41)	-2.726	(-4.42)
Mom working: No (base)												
Some of the time	-3.779	(-7.32)	-2.718	(-4.81)	-3.433	(-6.84)	-2.360	(-4.37)	-4.296	(-8.17)	-3.080	(-5.38)
Most of the time	-4.192	(-7.87)	-4.730	(-8.07)	-3.593	(-6.94)	-3.742	(-6.68)	-5.092	(-9.39)	-5.397	(-9.08)
All the time	-3.855	(-8.26)	-4.764	(-9.46)	-3.513	(-7.75)	-3.844	(-7.98)	-4.616	(-9.72)	-4.994	(-9.78)
Father education: less than prima	0.170	(0.16)	-2.468	(-2.52)	1.407	(1.36)	-0.826	(-0.88)	-0.154	(-0.14)	-3.039	(-3.06)
Some high school	-1.731	(-2.98)	-1.728	(-2.92)	-1.265	(-2.24)	-0.959	(-1.69)	-2.209	(-3.74)	-2.281	(-3.80)
Completed high school (base)												
Some college	0.358	(0.78)	0.581	(1.19)	0.003	(0.01)	-0.201	(-0.43)	1.022	(2.18)	1.298	(2.63)
Completed College	0.910	(2.04)	2.883	(6.00)	0.133	(0.31)	1.708	(3.72)	2.226	(4.93)	4.185	(8.61)
Graduate or professional	2.635	(4.66)	2.917	(4.88)	1.074	(1.95)	0.408	(0.71)	6.312	(11.09)	5.714	(9.48)
Mother education: less than prim	-1.717	(-1.50)	-3.926	(-3.63)	-0.144	(-0.13)	-2.827	(-2.73)	-1.434	(-1.23)	-4.050	(-3.69)
Some high school	-2.298	(-3.45)	-2.685	(-4.02)	-1.876	(-2.90)	-2.068	(-3.24)	-2.355	(-3.47)	-3.464	(-5.11)
Completed high school (base)												
Some college	-1.215	(-2.73)	0.310	(0.67)	-1.480	(-3.42)	-0.169	(-0.38)	-0.749	(-1.65)	0.936	(1.98)
Completed College	1.420	(3.30)	2.945	(6.27)	0.662	(1.58)	1.445	(3.22)	2.348	(5.36)	4.174	(8.78)
Graduate or professional	0.789	(1.34)	1.973	(3.16)	-0.178	(-0.31)	0.408	(0.68)	3.330	(5.60)	4.649	(7.39)
Works over school year	-3.092	(-2.55)	0.325	(0.22)	-2.206	(-1.87)	2.141	(1.50)	-4.307	(-3.49)	-0.629	(-0.42)

Note: Dependent variables is set to 100 if the student has a GPA of 4, and to 0 otherwise. T-statistics are in parentheses. This table completes Table 4 by presenting all the coefficients not included in the regression.

Table A3a. Selected Coefficients of LPM on Specific Grades - 12th graders - 2000-2009 (continued)

Dependent variable: A (93-100)	Specification 1				Specification 2				Specification 3			
Explanatory Variables	Boys		Girls		Boys		Girls		Boys		Girls	
Average hours of work: None												
5 or less hours	4.386	(3.96)	5.993	(4.26)	2.817	(2.62)	3.205	(2.39)	5.714	(5.07)	7.773	(5.45)
6 to 10 hours	1.529	(1.39)	2.543	(1.84)	0.687	(0.64)	0.744	(0.56)	2.781	(2.49)	3.983	(2.83)
11 to 20 hours	-0.607	(-0.58)	0.227	(0.17)	-1.310	(-1.29)	-0.955	(-0.74)	0.068	(0.06)	1.268	(0.93)
21 to 30 hours	-0.279	(-0.26)	-0.261	(-0.19)	-0.758	(-0.73)	-1.564	(-1.19)	-0.360	(-0.33)	-0.153	(-0.11)
More than 30 hours	2.062	(1.82)	-0.511	(-0.35)	1.061	(0.97)	-1.793	(-1.28)	1.535	(1.33)	-0.772	(-0.52)
Average earnings per week from job: None												
$1-5	1.578	(1.10)	-0.816	(-0.73)	1.111	(0.78)	-0.637	(-0.57)	4.162	(2.47)	2.795	(1.57)
$6-10	-0.270	(-0.31)	-0.211	(-0.33)	-0.507	(-0.58)	-0.404	(-0.64)	2.181	(2.12)	-2.540	(-2.50)
$11-50	-0.367	(-0.56)	-0.483	(-1.00)	-0.244	(-0.38)	-0.343	(-0.71)	1.913	(2.51)	0.107	(0.14)
$51-75	0.031	(0.04)	-0.546	(-1.02)	0.288	(0.41)	-0.392	(-0.74)	2.013	(2.40)	-1.360	(-1.60)
$76-125	-0.350	(-0.54)	-0.605	(-1.21)	-0.225	(-0.35)	-0.407	(-0.82)	0.114	(0.15)	-2.039	(-2.57)
$126+	-0.406	(-0.62)	-0.205	(-0.39)	-0.159	(-0.25)	-0.002	(0.00)	-0.470	(-0.61)	-2.247	(-2.70)
Constant	25.242	(28.44)	28.758	(29.58)	-1.310	(-1.25)	-13.363	(-11.36)	20.693	(23.19)	24.617	(25.17)
R-squared	0.116		0.126		0.166		0.202		0.082		0.099	
Number of observations	49328		56156		49328		56156		49328		56156	

Note: Dependent variables is set to 100 if the student has a GPA of 4, and to 0 otherwise. T-statistics are in parentheses. This table completes Table 4 by presenting all the coefficients not included in the regression.

Table A3b. Coefficients of LPM (100%) on Specific Grades - 12th graders - 2000-2009

Dependent variable: C+(77-79)	Specification 1				Specification 2				Specification 3			
Explanatory Variables	Boys		Girls		Boys		Girls		Boys		Girls	
Subjective School Ability					-20.698	(-28.22)	-20.586	(-31.84)				
Education Aspirations: want to attend (binary dummy)												
Army					-0.290	(-0.66)	0.287	(0.61)				
Vocational					0.241	(0.56)	0.733	(1.81)				
Two-year college					-0.285	(-0.75)	0.488	(1.56)				
Four-year college					1.298	(3.35)	-0.761	(-2.17)				
Graduate or professional					-0.773	(-2.30)	-0.147	(-0.54)				
Educational Expectations: index of likeness to attend[a]												
Army	2.328	(5.15)	0.469	(0.89)	2.296	(3.96)	0.134	(0.21)				
Vocational	0.473	(0.99)	0.876	(2.27)	0.125	(0.23)	0.175	(0.38)				
Two-year college	5.951	(15.15)	3.446	(11.57)	4.884	(10.56)	1.995	(5.42)				
Four-year college	-5.343	(-10.17)	-3.373	(-7.70)	-4.019	(-6.92)	-1.387	(-2.82)				
Graduate or professional	-5.101	(-10.55)	-3.123	(-8.57)	-2.762	(-5.06)	-1.606	(-3.80)				
High school program: Academic	-4.472	(-8.19)	-5.978	(-13.36)	-2.344	(-4.28)	-3.900	(-8.69)	-7.663	(-14.31)	-7.884	(-17.91)
General	-0.181	(-0.33)	-1.912	(-4.21)	0.450	(0.83)	-1.421	(-3.16)	-0.755	(-1.38)	-2.293	(-5.04)
Vocational	-1.720	(-2.50)	-1.507	(-2.41)	-0.456	(-0.67)	-0.635	(-1.02)	-0.289	(-0.42)	-0.787	(-1.26)
Other (base)												
Race: Black	6.870	(13.71)	5.158	(14.46)	6.020	(12.08)	4.699	(13.28)	6.472	(12.84)	4.792	(13.45)
SMSA	2.864	(9.13)	2.512	(10.15)	2.792	(8.97)	2.454	(10.00)	2.089	(6.64)	1.909	(7.73)
Smoked cigarettes per day: None (base)												
Less than one-half pack	1.865	(5.40)	2.941	(10.44)	1.275	(3.71)	2.431	(8.69)	2.567	(7.40)	3.412	(12.10)
One to 1½ pack	2.579	(3.21)	2.612	(3.67)	1.949	(2.44)	1.992	(2.82)	3.904	(4.84)	3.619	(5.09)
Two packs or more	-0.091	(-0.04)	-0.064	(-0.02)	-0.840	(-0.39)	-1.334	(-0.52)	0.621	(0.28)	0.948	(0.36)
Alcohol binging last 2 weeks: None (base)												
Once	0.752	(1.68)	0.703	(1.94)	0.698	(1.57)	0.525	(1.46)	0.677	(1.50)	0.598	(1.64)

Note: Dependent variables is set to 100 if the student has a GPA of 4, and to 0 otherwise. T-statistics are in parentheses. This table completes Table 4 by presenting all the coefficients not included in the regression.

Table A3a. Selected Coefficients of LPM on Specific Grades - 12th graders - 2000-2009 (continued)

Dependent variable: C+(77-79)	Specification 1				Specification 2				Specification 3			
Explanatory Variables	Boys		Girls		Boys		Girls		Boys		Girls	
Alcohol binging last 2 weeks:												
Two to nine times	1.533	(4.26)	1.709	(5.00)	1.373	(3.84)	1.362	(4.02)	1.532	(4.22)	1.594	(4.64)
Ten or more times	1.828	(1.75)	4.423	(3.14)	1.381	(1.34)	3.540	(2.54)	2.053	(1.96)	4.362	(3.09)
Siblings not same household	0.019	(0.06)	-0.294	(-1.17)	-0.118	(-0.38)	-0.318	(-1.27)	-0.004	(-0.01)	-0.197	(-0.78)
Siblings: One (base)												
None	-0.701	(-1.14)	0.175	(0.34)	-0.556	(-0.91)	0.372	(0.73)	-0.780	(-1.26)	0.035	(0.07)
Two	0.308	(0.91)	0.349	(1.28)	0.222	(0.66)	0.306	(1.13)	0.497	(1.46)	0.421	(1.54)
Three or more	0.497	(1.48)	0.440	(1.67)	0.443	(1.33)	0.350	(1.34)	0.901	(2.67)	0.709	(2.69)
Don't know	1.196	(0.48)	-0.202	(-0.09)	-0.051	(-0.02)	-1.239	(-0.57)	1.795	(0.72)	-0.160	(-0.07)
Mother not same household	0.579	(1.21)	-0.051	(-0.13)	0.392	(0.83)	-0.175	(-0.45)	1.054	(3.01)	1.799	(6.88)
Father not same household	1.044	(3.00)	1.728	(6.64)	0.870	(2.52)	1.587	(6.15)	1.223	(2.55)	0.285	(0.73)
Mom working: No (base)												
Some of the time	0.068	(0.15)	0.841	(2.33)	-0.095	(-0.21)	0.750	(2.10)	0.343	(0.76)	0.963	(2.66)
Most of the time	1.148	(2.48)	1.641	(4.38)	0.861	(1.87)	1.373	(3.70)	1.650	(3.54)	1.858	(4.94)
All the time	1.542	(3.81)	1.509	(4.69)	1.380	(3.43)	1.269	(3.97)	1.928	(4.73)	1.568	(4.85)
Father education: less than prima	1.164	(1.26)	1.444	(2.31)	0.616	(0.67)	0.945	(1.52)	1.490	(1.60)	1.650	(2.62)
Some high school	0.826	(1.64)	2.117	(5.59)	0.624	(1.25)	1.876	(5.00)	1.201	(2.37)	2.335	(6.15)
Completed high school (base)												
Some college	0.153	(0.38)	-0.223	(-0.71)	0.311	(0.78)	0.015	(0.05)	-0.227	(-0.56)	-0.503	(-1.61)
Completed College	-0.302	(-0.78)	-0.222	(-0.72)	0.062	(0.16)	0.123	(0.40)	-1.132	(-2.92)	-0.706	(-2.30)
Graduate or professional	-0.444	(-0.90)	-0.382	(-1.00)	0.296	(0.61)	0.331	(0.87)	-2.342	(-4.79)	-1.339	(-3.51)
Mother education: less than prim	-0.387	(-0.39)	0.172	(0.25)	-1.139	(-1.16)	-0.191	(-0.28)	-0.459	(-0.46)	0.221	(0.32)
Some high school	1.500	(2.59)	0.712	(1.67)	1.290	(2.25)	0.547	(1.29)	1.667	(2.86)	1.006	(2.35)
Completed high school (base)												
Some college	-0.717	(-1.86)	-0.675	(-2.27)	-0.593	(-1.55)	-0.529	(-1.79)	-1.003	(-2.58)	-0.929	(-3.11)
Completed College	-0.977	(-2.61)	-1.504	(-5.01)	-0.631	(-1.70)	-1.073	(-3.60)	-1.564	(-4.16)	-1.975	(-6.56)
Graduate or professional	-1.299	(-2.55)	-1.062	(-2.66)	-0.852	(-1.69)	-0.620	(-1.57)	-2.608	(-5.11)	-1.988	(-4.99)
Works over school year	0.996	(0.95)	2.283	(2.40)	0.601	(0.58)	1.713	(1.82)	1.798	(1.70)	2.598	(2.72)

Note: Dependent variables is set to 100 if the student has a GPA of 4, and to 0 otherwise. T-statistics are in parentheses. This table completes Table 4 by presenting all the coefficients not included in the regression.

Table A3a. Selected Coefficients of LPM on Specific Grades - 12th graders - 2000-2009 (continued)

Dependent variable: C+(77-79)	Specification 1				Specification 2				Specification 3			
Explanatory Variables	Boys		Girls		Boys		Girls		Boys		Girls	
Average hours of work: None												
5 or less hours	-2.086	(-2.17)	-3.144	(-3.50)	-1.362	(-1.43)	-2.317	(-2.60)	-2.927	(-3.02)	-3.749	(-4.16)
6 to 10 hours	-2.198	(-2.31)	-3.021	(-3.41)	-1.796	(-1.90)	-2.481	(-2.82)	-2.994	(-3.12)	-3.527	(-3.97)
11 to 20 hours	-1.714	(-1.89)	-2.463	(-2.87)	-1.383	(-1.53)	-2.084	(-2.45)	-2.195	(-2.40)	-2.831	(-3.28)
21 to 30 hours	-1.678	(-1.81)	-2.115	(-2.42)	-1.437	(-1.56)	-1.700	(-1.96)	-1.723	(-1.85)	-2.153	(-2.45)
More than 30 hours	-0.746	(-0.76)	-2.652	(-2.83)	-0.263	(-0.27)	-2.229	(-2.40)	-0.438	(-0.44)	-2.541	(-2.70)
Average earnings per week from job: None												
$1-5	1.578	(1.10)	-0.816	(-0.73)	1.111	(0.78)	-0.637	(-0.57)	1.885	(1.30)	-0.413	(-0.37)
$6-10	-0.270	(-0.31)	-0.211	(-0.33)	-0.507	(-0.58)	-0.404	(-0.64)	-0.034	(-0.04)	-0.090	(-0.14)
$11-50	-0.367	(-0.56)	-0.483	(-1.00)	-0.244	(-0.38)	-0.343	(-0.71)	-0.374	(-0.57)	-0.454	(-0.93)
$51-75	0.031	(0.04)	-0.546	(-1.02)	0.288	(0.41)	-0.392	(-0.74)	-0.201	(-0.28)	-0.609	(-1.13)
$76-125	-0.350	(-0.54)	-0.605	(-1.21)	-0.225	(-0.35)	-0.407	(-0.82)	-0.317	(-0.48)	-0.624	(-1.24)
$126+	-0.406	(-0.62)	-0.205	(-0.39)	-0.159	(-0.25)	-0.002	(0.00)	-0.291	(-0.44)	-0.176	(-0.33)
Constant	7.884	(10.23)	6.896	(11.10)	20.066	(21.54)	19.544	(25.05)	10.620	(13.86)	8.434	(13.63)
R-squared	0.051		0.051		0.066		0.069		0.035		0.042	
Number of observations	49328		56156		49328		56156		49328		56156	

Note: Dependent variables is set to 100 if the student has a GPA of 2.3, and to 0 otherwise. T-statistics are in parentheses. This table completes Table 4 by presenting the coefficients not reported there but included in the regression.

Table A4. Detailed Decomposition Results - 12th Graders
Percentage Female/Male Difference for Selected GPA Levels

A: 1976-1988	A (93-100): 4		C+ (77-79): 2.3		A (93-100): 4		C+ (77-79): 2.3	
Total Differential	3.747	(0.005)	-4.429	(0.007)	3.747	(0.005)	-4.429	(0.007)
Composition Effects:								
Total Explained	0.787	(0.003)	-0.321	(0.004)	0.549	(0.003)	-0.097	(0.004)
Race, SMSA	-0.151	(0.000)	0.234	(0.001)	-0.095	(0.000)	0.190	(0.001)
Own School Ability					-0.081	(0.001)	0.068	(0.001)
Smoking, Binging	0.602	(0.001)	-0.402	(0.002)	0.479	(0.001)	-0.300	(0.002)
Family Background	-0.083	(0.001)	0.004	(0.001)	-0.032	(0.001)	-0.033	(0.001)
Work	0.143	(0.001)	0.024	(0.001)	0.149	(0.001)	0.025	(0.001)
High school program	0.065	(0.001)	-0.098	(0.001)	0.006	(0.000)	-0.044	(0.001)
Educ. Expectations	0.211	(0.002)	-0.083	(0.003)	0.123	(0.002)	-0.004	(0.003)
Specification Error	0.001	(0.005)	0.152	(0.007)	-0.011	(0.005)	0.172	(0.007)
Educational Responses:								
Total Unexplained	3.099	(0.005)	-4.294	(0.005)	3.200	(0.004)	-4.435	(0.005)
Race, SMSA	-1.243	(0.007)	0.222	(0.008)	-1.290	(0.013)	0.064	(0.008)
Own School Ability					0.493	(0.001)	0.002	(0.000)
Smoking, Binging	-0.701	(0.004)	0.166	(0.004)	-0.240	(0.005)	0.106	(0.004)
Family Background	1.119	(0.018)	-1.830	(0.021)	1.196	(0.026)	-1.549	(0.021)
Work	0.412	(0.009)	0.766	(0.011)	1.569	(0.012)	0.899	(0.010)
High school program	1.700	(0.019)	-1.618	(0.021)	1.381	(0.027)	-1.758	(0.021)
Educ. Expectations	0.006	(0.002)	-0.179	(0.002)	0.796	(0.022)	0.660	(0.012)
Constant	1.807	(0.029)	-1.821	(0.033)	1.501	(0.046)	-2.859	(0.035)
Reweighting Error	-0.139	(0.001)	0.034	(0.001)	-0.009	(0.002)	-0.068	(0.001)
B: 1989-1999	A (93-100): 4		C+ (77-79): 2.3		A (93-100): 4		C+ (77-79): 2.3	
Total Differential	4.711	(0.006)	-3.898	(0.005)	4.711	(0.006)	-3.898	(0.005)
Composition Effects:								
Total Explained	1.499	(0.003)	-0.713	(0.003)	0.517	(0.003)	-0.103	(0.003)
Race, SMSA	-0.258	(0.001)	0.284	(0.001)	-0.212	(0.001)	0.257	(0.001)
Own School Ability					-0.355	(0.002)	0.209	(0.001)
Smoking, Binging	0.560	(0.001)	-0.246	(0.001)	0.427	(0.001)	-0.167	(0.001)
Family Background	-0.145	(0.001)	0.062	(0.001)	-0.058	(0.001)	0.013	(0.001)
Work	-0.010	(0.001)	0.002	(0.001)	0.019	(0.001)	-0.016	(0.001)
High school program	0.220	(0.001)	-0.237	(0.001)	0.031	(0.001)	-0.121	(0.001)
Educ. Expectations	1.132	(0.002)	-0.579	(0.002)	0.665	(0.002)	-0.279	(0.002)
Specification Error	0.114	(0.006)	-0.085	(0.007)	0.030	(0.006)	0.021	(0.006)
Educational Responses:								
Total Unexplained	3.035	(0.006)	-3.204	(0.005)	3.885	(0.005)	-3.627	(0.005)
Race, SMSA	-1.772	(0.010)	-0.419	(0.009)	-2.250	(0.010)	-0.342	(0.009)
Own School Ability					-0.046	(0.000)	0.008	(0.000)
Smoking, Binging	-0.582	(0.004)	-0.020	(0.004)	-0.378	(0.004)	-0.123	(0.004)
Family Background	-0.493	(0.019)	-0.996	(0.018)	-0.470	(0.018)	-0.937	(0.018)
Work	0.998	(0.011)	0.081	(0.010)	0.583	(0.011)	0.195	(0.010)
High school program	1.001	(0.026)	-2.365	(0.024)	0.862	(0.025)	-2.819	(0.024)
Educ. Expectations	0.036	(0.002)	-0.247	(0.002)	0.887	(0.016)	-0.080	(0.015)
Constant	3.845	(0.036)	0.762	(0.033)	4.695	(0.037)	0.471	(0.036)
Reweighting Error	0.062	(0.002)	-0.057	(0.001)	0.279	(0.002)	-0.189	(0.001)

Note: Standard errors are in parentheses. Reweighted decomposition follows methodology of section 4. In specification 2, educational aspirations are included among the variables in the educational expectations category.

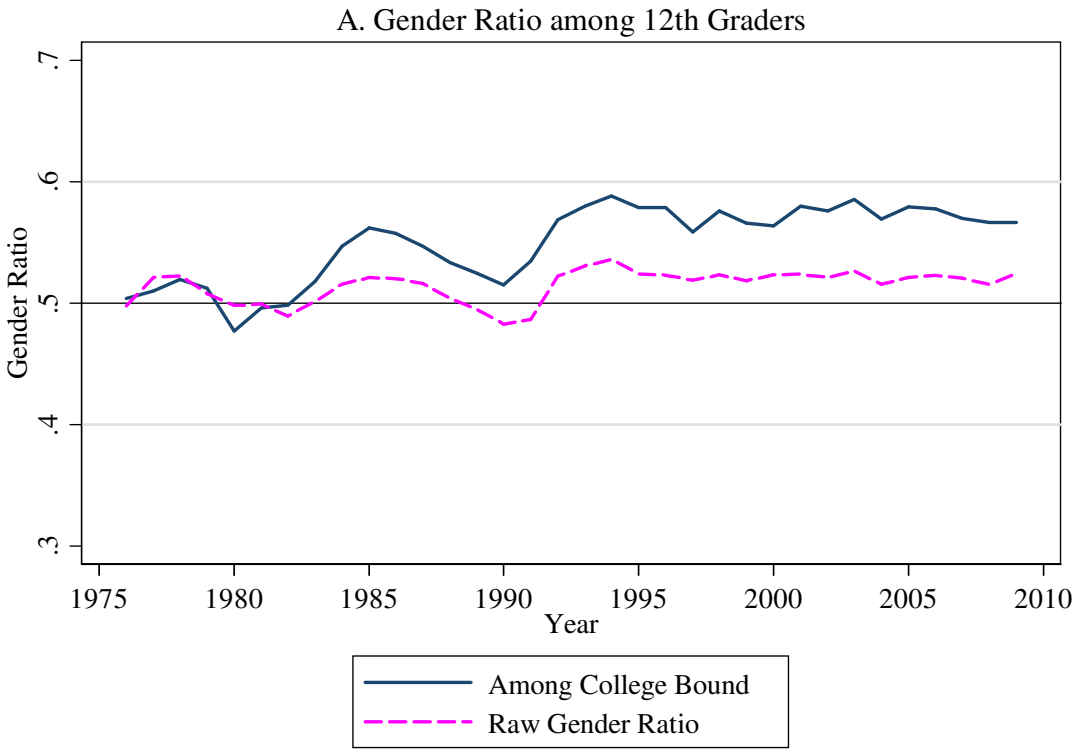
Table A5. Detailed Decomposition Results - Composition Effects
Percentage Female/Male Difference for Selected GPA Levels

	10th graders				8th graders			
1991-1999	A (93-100): 4		C (73-76): 2		A (93-100): 4		C+ (77-79): 2.3	
Total Differential	3.585	(0.006)	-3.110	(0.005)	5.246	(0.006)	-2.383	(0.005)
Composition Effects:								
Total Explained	1.328	(0.003)	1.574	(0.003)	1.872	(0.004)	-0.800	(0.003)
Race, SMSA	-0.183	(0.001)	0.087	(0.001)	-0.196	(0.001)	0.132	(0.001)
School Hard, Held back	-0.454	(0.001)	0.108	(0.001)	-0.322	(0.001)	0.059	(0.001)
Misbehavior, Smoking	0.640	(0.002)	-0.537	(0.002)	1.164	(0.002)	-0.502	(0.002)
Family Background	-0.140	(0.001)	0.106	(0.001)	-0.177	(0.001)	0.150	(0.001)
Work	0.083	(0.001)	0.128	(0.001)	-0.016	(0.001)	-0.070	(0.001)
High school program	0.343	(0.001)	-0.254	(0.001)	0.354	(0.001)	-0.124	(0.001)
Educ. Expectations	1.040	(0.002)	-1.212	(0.002)	1.065	(0.002)	-0.445	(0.002)
Specification Error	0.004	(0.006)	-0.242	(0.006)	0.165	(0.007)	-0.258	(0.006)
Educational Responses:								
Total Unexplained	1.981	(0.006)	-1.184	(0.005)	2.783	(0.006)	-1.075	(0.005)
Race, SMSA	-1.476	(0.010)	0.300	(0.009)	1.058	(0.011)	-0.387	(0.010)
School Hard, Held back	-0.020	(0.002)	-0.023	(0.002)	-0.129	(0.003)	0.050	(0.002)
Misbehavior, Smoking, B	-0.332	(0.005)	0.075	(0.004)	-0.115	(0.004)	0.379	(0.004)
Family Background	-0.156	(0.017)	0.724	(0.015)	0.369	(0.019)	-0.393	(0.016)
Work	0.165	(0.006)	0.373	(0.005)	0.078	(0.007)	-0.386	(0.006)
High school program	0.785	(0.013)	0.086	(0.011)	0.917	(0.008)	-0.084	(0.007)
Educ. Expectations	0.007	(0.002)	0.204	(0.002)	-0.173	(0.002)	0.096	(0.002)
Constant	3.009	(0.025)	-2.923	(0.022)	0.779	(0.025)	-0.350	(0.021)
Reweighting Error	0.272	(0.002)	-0.110	(0.001)	0.426	(0.003)	-0.068	(0.001)

Note: Standard errors are in parentheses. Reweighted decomposition follows methodology of section 4.

Figure A1. Trends by Gender

Figure A2. Gender Ratios



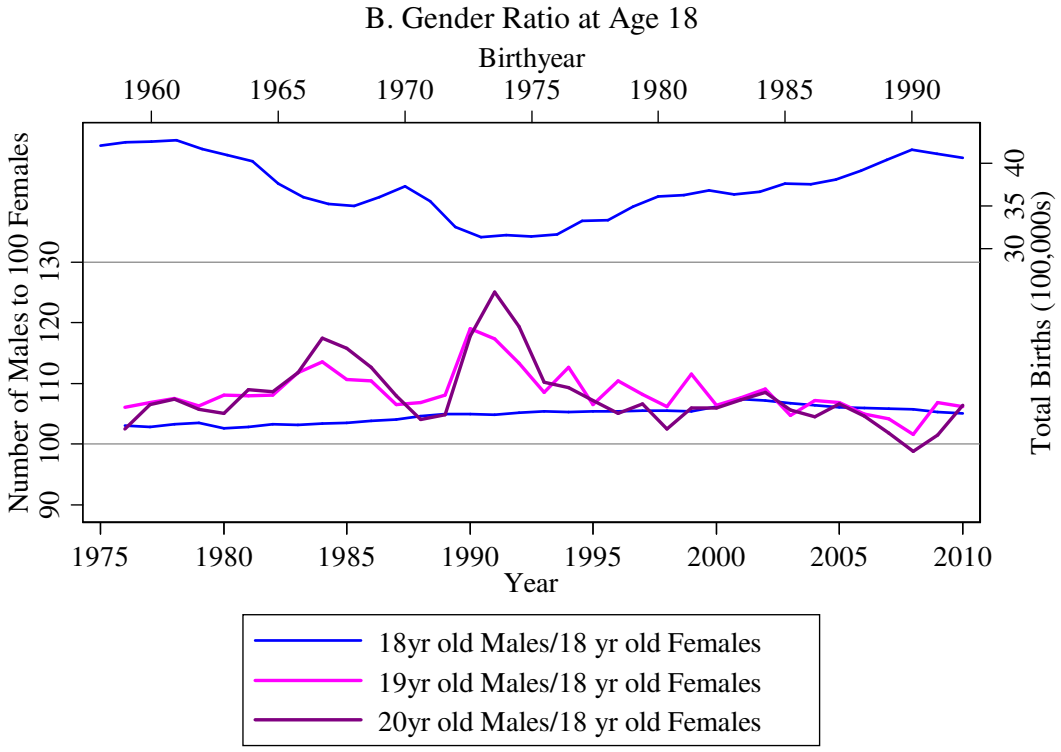
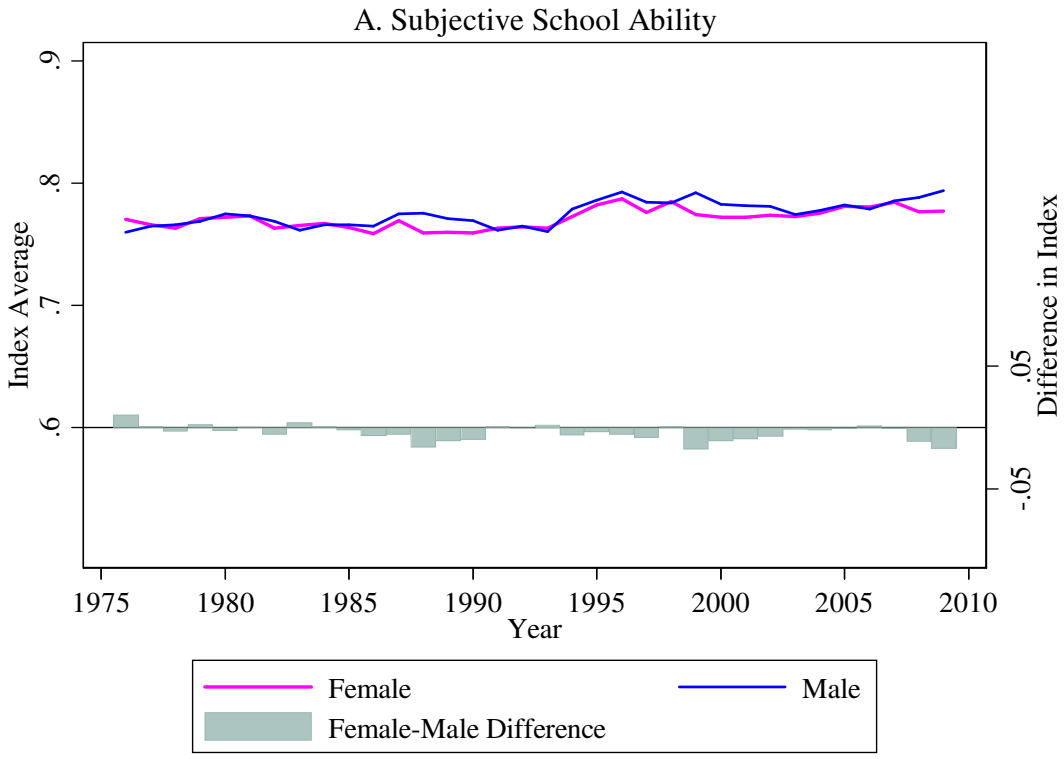
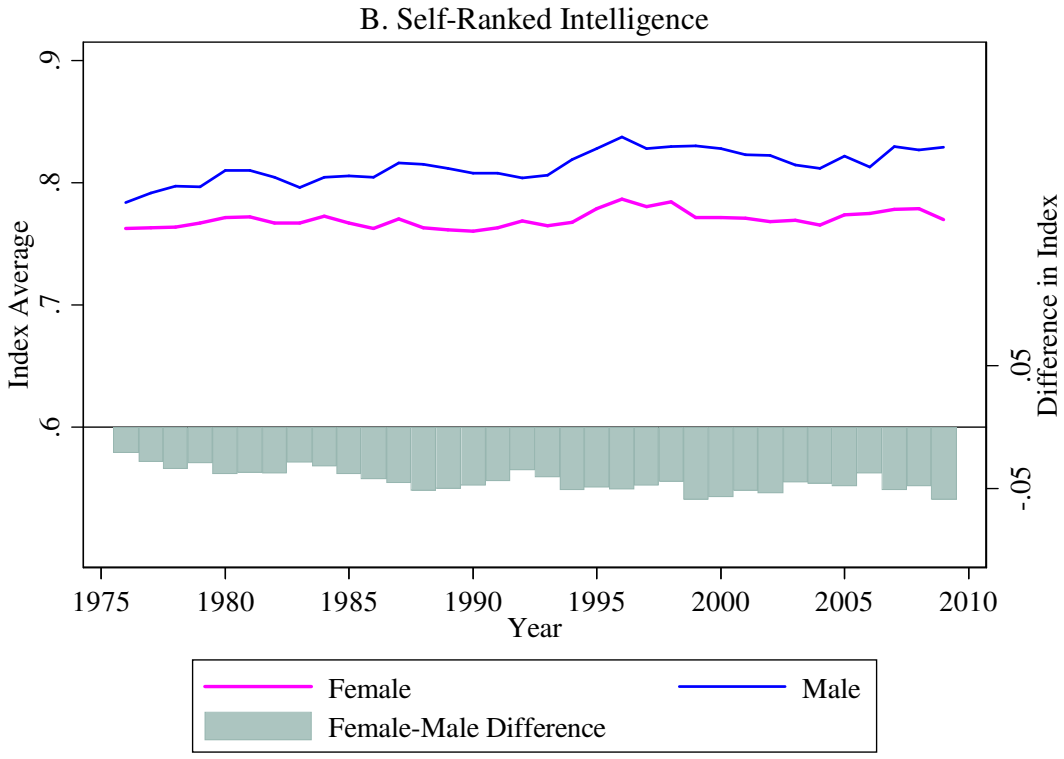


Figure A3. Average Subjective School Ability and Intelligence by Gender and Gender Gap





Appendix B. Simple "Threshold" Model of High School Performance

We present below a simple behavioral threshold model of academic performance where the changing aspirations and expectations for the future of girls would lead them to capture a larger proportion of high grades. Over the last three decades there has been sustained effort (Manski and Wise, 1983; Manski, 1993; Dominitz, Manski, and Fischhoff, 2001; Manski, 2004) to understand the formation of students' expectations and to ascertain the importance of these expectations in their decision to enroll in college. Recent studies (Stange, 2008; Zafar, 2011; Jacob and Wilder, 2012) set in longitudinal settings have focused on learning and beliefs updating, trying to address on the first part of the puzzle. In our cross-sectional setting, we rely instead on changes over time across gender in the educational expectations of students in order to evaluate the importance of these expectations for the high academic achievement that opens the door to college-going and to graduate school attendance. We argue that these changes are different by gender for reasons exogenous to the early education process; rather they are rooted in changes in the labor market opportunities for women.[59]

Models of high school performance in economics are usually set as derivatives of the Mincerian human capital investment model where individuals choose their level of schooling to maximize their life-cycle earnings. Here we want think of decisions taken earlier in life when labor market outcomes are not as concrete or narrow as returns to college, but would come out as the answer to the typical question: "What do you want to do when you grow up?"At this stage, parents are likely still involved in the education of their children, perhaps actively assisting them with homework, helping them set goals, and manage their time. The model thus borrows from the model of intergenerational transmission of income status (Becker and Tomes, 1979), the idea that the other generation's utility, the parents in this case, enters the decision maker's objective function. The parents' utility from their offspring's educational expectations depends on their own characteristics, such as their own level of education, as well as on the school ability or aptitude of each child. Over the last three decades, exogenous changes in the opportunities for women in the labor market have led many parents to have higher educational expectations for their daughters. As shown in Chen, Fortin, Phipps, and Oreopoulos (2011), in the 2000s, parents of primary school students had higher educational expectations for their daughters than for their

[59] A caveat here is the disproportionate increase in ADHD among boys (Elder and Lubotsky, 2009; Chen, Fortin, Oreopoulos and Phipps, 2011). But in our data, we actually see a decrease over time in school misbehavior

sons. Frenette and Zeman (2007) find that parental expectations about the highest education level that their child will attain accounts for a notable share of gender differences in university attendance. Further the role of parental expectations could explain why first generation immigrant boys suffer less from the boys' underachievement problem.[60]

As with signaling models of educational choices (Spence, 1973), we want to allow for the fact that, given a level of aptitude or ability, an individual may find it optimal to aim for the minimum GPA needed to reach a career or educational goal. Implicit in our framework is the fact that basic school ability is revealed quite early in the pupil's schooling experience . This contrasts with learning models (Stange, 2008) where academic ability for college is revealed slowly over time and where individuals revised their schooling decisions. We think that both models provide an adequate representation of the behavior of some subsets of individuals and for different levels of skills. The updating of educational expectations is perhaps more salient among college/university students who face more fateful choices about which major to pursue or whether or not to continue their studies (Zafar, 2011) than among high school students for whom the salient choice is whether or not to drop out (once they reach minimum school leaving age). Despite updating by some individuals, results in Stange (2008) and Jacob and Wilder (2012) support the idea that the majority of individuals are actually successful at enacting their early educational plans. Jacob and Wilder (2012) report that only 35% of high school students update their educational expectations from grade 8 to grade 10; from grade 10 to 12, that percentage is only 25%. Importantly, because we use cross-sectional data, the model is set in a static framework.

We assume that a pupil comes to secondary school with a basic aptitude for school (A_i) that was largely revealed during elementary school and with plans for future education likely influenced by parental desires, and aims for a GPA (G_i) in the range $R_j(G_i) = I[l_j \leq G_i \leq u_j]$, where I is the indicator function, that will allow him/her to pursue further schooling/career plans ($S_j j = 1, \ldots, J$). As illustrated in Figure B.1, the simple functional form for $R_j(G_i)$ with multiple thresholds assumes with probability one that a student with a GPA in the indicated range will be able to pursue her educational plans.[61] The student chooses a level of effort and target GPA to

[60] Wilson, Burgess and Briggs (2011) have also suggested aspiration-based explanations to account for ethnic differences in academic performance. See also Zafar (2011).
[61] In reality, the discontinuities do not need to be as sharp as illustrated in Figure B.1.

maximize her utility from schooling plans minus costs, $C(G_i)$, plus an intergenerational utility component,

$$Max_{\{G,E\}}\ (1-\alpha)U_i\{S_j - C(G_i)\} + \alpha U_i^p(S_j, A_i),$$
$$\text{subject to } S_j(G_i) = I\left[l_j \le G_i \le u_j\right], \qquad j = 1,\ldots,J$$
$$C(G_i) = f(E_i, A_i).$$

The component $U_i^p(S_j, A_i)$ represents the utility to the parents of having an offspring of ability A_i in reach of educational level S_j, and $\alpha \in [0,1]$ is the weight placed by the student on parental utility. This last parameter is potentially important is assessing gender differences, as psychologists argue that girls place more importance on pleasing adults than boys. Assuming separability of schooling plans and costs, however, parental utility merely acts to scale the rewards of a schooling plan. The effect of parents' other characteristics might have similar rescaling effects on the cost function.[62]

Importantly, the cost of getting a particular grade, $C(G_i) = f(E_i, A_i)$, is decreasing non-linearly with ability, $\frac{\partial f(E_i,A_i)}{\partial A_i} < 0$, and $\frac{\partial^2 f(E_i,A_i)}{\partial A_i^2} > 0$. The cost of academic achievement is increasing with effort E_i, $\frac{\partial f(E_i,A_i)}{\partial E_i} > 0$, a but there may be some complex non-linear interactions between effort and ability, $\frac{\partial^2 f(E_i,A_i)}{\partial E_i \partial A_i} \gtreqless 0$, possibly different by gender, that we do not attempt to model directly here. The mapping $R_j(G_i)$ of GPA into educational plans may include a more complex step function than the one above, where there are different probabilities of attaining educational choice S_j by GPA level. What is important in leading some students to optimally choose lower levels of GPA are the thresholds in access to educational choice by GPA, as shown in the example of Figure B.1.

In Figure B.1, we present an example of three potential educational choices for student *i*. We illustrate the special case where utility, $U_i(S_j) = w_i * I\left[l_j \le G_i\right]$, is a simple step function corresponding to different wage levels chosen as arbitrary examples: for *j*=two-year college ($w_i = 4, l_{2-yr} = 2$), four-year college ($w_i = 7.5, l_{4-yr} = 3$), and graduate school ($w_i = 10, l_{grad} = 4$). The cost functions illustrated in Figure B1, subsume in their functional forms the level of effort needed for high, medium and low ability students to achieve that higher GPAs,

[62] The role of teachers in this model would be similar to that of parents in lowering the cost of academic achievement and enhancing its benefits by motivating students to succeed.

showing that it is more costly for lower ability students to obtain high GPAs. Thus the choices of GPA, $G_i^*(A_i, S_j)$, which maximizes the utility net of achievement cost for each ability level, are the lower bound of each educational choice. That is, the low ability student will target a GPA of 2 to access two-year colleges, the medium ability a GPA of 3 to access four-year colleges, and the high ability student a GPA of 4 aiming to attend graduate school. Letting $G_i^{max}(A_i) = G_i^*(A_i, S_J)$ be the highest grade that a student, with a given level of ability A_i, can attain when the student has the highest educational aspirations $j = J$ (e.g. graduate school), then a student's optimal choice of GPA may reflect potential educational under-achievement $G_i^*(A_i, S_k) < G_i^{max}(A_i)$.

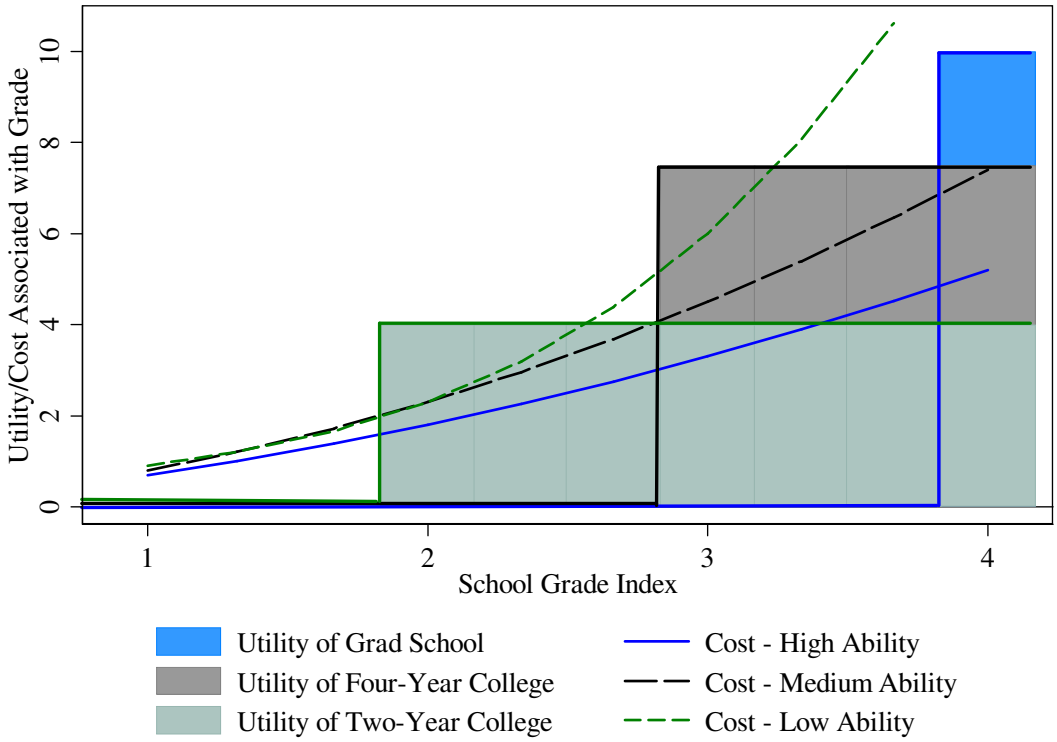


Figure B1. Utility and Cost of Academic Achievement

The model helps rationalize the relative underperformance of boys as the consequence of career choices that require lower levels of educational attainment. This potential educational underachievement may be of concern when students' educational aspirations are limited by lack of information, borrowing constraints, time impatience, or other intertemporal optimization errors. This model contrasts with both the early childhood development (ECD) branch and

education production function (EPF) branch of the literature on cognitive achievement in children (Todd and Wolpin, 2003), where the goal is to maximize achievement under some cost constraints.

Appendix B - References:

Becker, Gary and N. Tomes (1979) "An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility," *Journal of Political Economy*, 87:1153-89.

Chen, Kelly, Nicole Fortin, Philip Oreopoulos and Shelley Phipps (2011) "School Start Age and Hyperactivity in Canadian Children," mimeo, Dalhousie University.

Dominitz, J., C. Manski, and B. Fischhoff (2001): "Who Are Youth At-Risk?: Expectations Evidence in the NLSY-97," in *Social Awakenings:A dolescents' Behaviors as Adulthood Approaches,* ed. by R. Michael. New York: Russell Sage Foundation, 230-257.

Elder, Todd and Darren H. Lubotsky (2009) "Kindergarten Entrance Age and Children's Achievement: Impacts of State Policies, Family Background, and Peers," *Journal of Human Resources*, 44(3): 641-683.

Frenette, Marc and Klarka Zeman (2007) "Why are Most University Students Women? Evidence based on Academic Performance, Study Habits and Parental Influences," Analytical Studies Branch Research Paper Series No 303. Ottawa: Statistics Canada.

Manski, Charles F., and David Wise (1983) *College Choice in America*. Cambridge: Harvard University Press.

Manski, Charles F. (1993): "Adolescent Econometricians: How Do Youth Infer the Returns to Schooling?" in *Studies of Supply and Demand in Higher Education*, ed. by C. Clotfelter and M. Rothschild. Chicago: University of Chicago Press, 43-57.

Manski, Charles F. (2004) "Measuring Expectations," *Econometrica*, 72(5): 1329-1376.

Spence, A. Michael (1973) "Job Market Signaling," *Quarterly Journal of Economics*, 87 (3): 355-374.

Todd, Petra E. and Kenneth I. Wolpin (2003) "On the Specification and Estimation of the Production Function for Cognitive Achievement", *Economic Journal*, 113(485): F3–F33.

Wilson, Deborah, Simon Burgess and Adam (2011) "The Dynamics of School Attainment of England's Ethnic Minorities", *Journal of Population Economics.* 24(2): 681-700.