

NBER WORKING PAPER SERIES

WHERE THE BOYS AREN'T: NON-COGNITIVE SKILLS, RETURNS TO SCHOOL AND  
THE GENDER GAP IN HIGHER EDUCATION

Brian A. Jacob

Working Paper 8964  
<http://www.nber.org/papers/w8964>

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
June 2002

The views expressed herein are those of the author and not necessarily those of the National Bureau of Economic Research.

© 2002 by Brian A. Jacob. 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.

Where the boys aren't: Non-cognitive skills, returns to school and the gender gap in higher education

Brian A. Jacob

NBER Working Paper No. 8964

June 2002

JEL No. I20, J16, J18

### **ABSTRACT**

Nearly 60 percent of college students today are women. Using longitudinal data on a nationally representative cohort of eighth grade students in 1988, I examine two potential explanations for the differential attendance rates of men and women -- returns to schooling and non-cognitive skills. The attendance gap is roughly five percentage points for all high school graduates. Conditional on attendance, however, there are few differences in type of college, enrollment status or selectivity of institution. The majority of the attendance gap can be explained by differences in the characteristics of men and women, despite some gender differences in the determinants of college attendance. I find that higher non-cognitive skills and college premiums among women account for nearly 90 percent of the gender gap in higher education. Interestingly, non-cognitive factors continue to influence college enrollment after controlling for high school achievement.

Brian A. Jacob

John F. Kennedy School of Government

Harvard University

79 JFK Street

Cambridge, MA 02138

and NBER

(617) 384-7968

Brian\_Jacob@harvard.edu

## 1. Introduction

Today nearly 60 percent of college students are women (Bae, Choy, Geddes, Sable & Synder 2000). Among low-income and minority students, young women are 25 percent more likely than young men to enroll in some form of postsecondary education.<sup>1</sup> While the enrollment gap differential was first evident in the late seventies, it has recently begun to attract attention from a broad spectrum of individuals including high school guidance counselors, college admissions officers, and policymakers (Koerner 1999; Mortenson 1999; Sommers 2000). Given that changes in the U.S. economy over the past two decades have significantly increased the return to college and cognitive ability (Katz & Murphy 1992; Levy & Murnane 1992; Murnane, Willet & Levy 1995), the relative decline in college enrollment among young men warrants investigation.<sup>2</sup>

While there has been an increasing recognition of this gap, there are no empirical studies that specifically seek to explain the phenomena. Some commentators have suggested that the predominance of young men in the military and prison explain the gap. However, in 1989 the attendance gap among 19-21 year olds was roughly eight percentage points. If *all* boys in the military or prison had attended college instead, the difference would still be roughly three percentage points (Census 1990; Justice 1997).

In this paper, I examine two potential explanations for the differential attendance rates of men and women—returns to schooling and non-cognitive skills. Several studies have found that the returns to college are higher for women, particularly among young workers (Murphy & Welch 1989; Murphy & Welch 1992). Kane (1994) reports that, among African-Americans,

---

<sup>1</sup> Based on calculations using the NELS data that are described in more detail below.

<sup>2</sup> While gender differences in college enrollment are commonly lamented in the media, it is not clear a priori that this is a problem. If this finding simply represents different employment preferences or tastes for schooling, it may

college earnings differentials rose faster for women than men in the eighties. While women have made considerable inroads into traditional white-collar male jobs, they have had significantly less success moving into skilled blue-collar occupations (England & Farkas 1986; Reskin & Roose 1990). For this reason, it is likely that a young man with high school diploma will have a better chance than a comparable young woman of finding a relatively high paying construction or manufacturing job.

While boys and girls today score roughly the same on most measures of cognitive ability,<sup>3</sup> boys generally receive lower grades, have more disciplinary problems, are more likely to be retained in grade and placed in special education, report lower school enjoyment and attachment and believe their teachers are less likely to encourage them (Kleinfeld 1998). These difficulties are often attributed to poor “non-cognitive” skills among boys, including the inability to pay attention in class, to work with others, to organize and keep track of homework or class materials and to seek help from others. In recent years, non-cognitive skills have received increased attention among economists in trying to explain labor market success (Murnane & Levy 1996; Duncan & Dunifon 1998). Low non-cognitive skills may reduce college attendance rates indirectly by decreasing the likelihood that a student graduates high school or leading to poor high school performance, which decreases the chance of being accepted to college or receiving financial aid. In addition, low non-cognitive skills may increase the non-pecuniary or “psychic” costs of college, thereby decreasing the likelihood of enrollment.

Using longitudinal data on a nationally representative cohort of eighth grade students in 1988, I find that the attendance gap is roughly five and ten percentage points for all high school

---

not be a matter for great societal concern. Moreover, women are still disproportionately represented in non-technical fields and continue to lag behind men in earnings (Blau 1998).

<sup>2</sup> In the 1996 National Assessment of Educational Progress (NAEP), boys scored higher in math and science while girls scored higher in reading and writing (Sommers 2000).

graduates and lower-income graduates respectively. Conditional on attendance, however, there are few differences in type of college, enrollment status or selectivity of institution. The majority of the attendance gap can be explained by differences in the characteristics of men and women, despite some gender differences in the determinants of college attendance. I find that higher non-cognitive skills and college premiums among women account for nearly 80 percent of the gender gap in higher education. Interestingly, non-cognitive factors continue to influence college enrollment after controlling for high school achievement.

The remainder of the paper is organized as follows. Section 2 lays out a simple model of college attendance. Sections 3 and 4 describe the data and empirical strategy used in this paper. Section 5 presents some descriptive statistics on the differences between men and women in terms of postsecondary experience including the type and selectivity of college attended. Sections 6 and 7 present the main findings and Section 8 concludes.

## **2. A model of college attendance**

The following model highlights several possible factors underlying the gender gap. To begin, assume that an individual has graduated high school and has only one period,  $k$ , during which to attend college. Further assume that students who decide to attend college will graduate.<sup>4</sup> The primary motivation of the model is the view that an individual will attend college if the benefits exceed the costs. Here we consider only the financial benefit of college measured in terms of lifetime earnings. Denote the earnings of a high school graduate in each period,  $t$ , as:

$$Y_{ht}, \quad t = k, k+1, \dots T$$

Denote the earnings of a college graduate in period  $t$  as:

---

<sup>4</sup> Altonji (1993) presents an interesting model that explicitly addresses the uncertainty inherent in the college attendance decision.

$$Y_{ct} \quad t = k, k+1, \dots T$$

The earnings benefit of college attendance, often referred to as the college premium, may now be written as:

$$\Pi_{ct} = Y_{ct} - Y_{ht} \quad t = k+1 \dots T$$

Assume that college takes one period to complete and that credit markets are perfect so that the individual can borrow or lend at interest rate  $r$ . The net expected present value of this premium is then:

$$E \left[ \sum_{j=1}^{T-k} \frac{Y_{c,k+j}}{(1+r)^j} - \sum_{j=1}^{T-k} \frac{Y_{h,k+j}}{(1+r)^j} \right] = E \left[ \sum_{j=1}^{T-k} \frac{\Pi_{c,k+j}}{(1+r)^j} \right]$$

The costs of college attendance consist of three components: (1) direct costs,  $d$ , such as tuition and fees net of financial aid; (2) forgone earnings during the period,  $Y_{hk}$ , referred to as opportunity costs and (3) “psychic” costs,  $h$ , that depend on a student’s non-cognitive skills (i.e., the higher a student’s non-cognitive ability, the more easily she is able to navigate college life and the less effort she must put forth to succeed). Given these costs and benefits, a utility maximizing decision-maker will attend college ( $C=1$ ) when the net expected present value of college attendance is positive.

$$C=1, \quad \text{if } E \left[ \sum_{j=1}^{T-k} \frac{\Pi_{c,j}}{(1+r)^j} - d - Y_{hk} - h \geq 0 \right]$$

$$C=0, \quad \text{otherwise.}$$

If we allow all costs and premiums to differ by gender and assume that  $\Pi_{c,t}$  is constant in all periods and  $T$  is large, then we can write the structural model of college attendance as

$$\Pr ( C ) = \Pr \left( \frac{\Pi_c^g}{r} - d^g - Y_{h,k}^g - h^g \geq 0 \right)$$

where the superscript  $g$  represents male or female. However, because I am primarily interested in the marginal impacts of various characteristics on the probability of college attendance, I will estimate the reduced form

$$\Pr(C_i) = F(X_i\beta^g) + \varepsilon_i$$

where  $C$  is an indicator of college attendance,  $X$  is a vector of observable characteristics that includes the college premium, opportunity costs, direct costs and psychic costs,  $\varepsilon$  is a stochastic error term and  $g$  subscripts male ( $m$ ) and female ( $f$ ). If one assumes that college attendance follows a linear probability model and restrict the coefficients  $\beta$  to be the same for men and women, but allow the intercepts to differ, then the reduced form can be rewritten as

$$\Pr(C_i) = (Female_i)\alpha + X_i\beta + \varepsilon_i$$

where  $Female$  is an indicator variable and  $\alpha$  measures the “gender gap.”

### 3. Data

Student enrollment information is drawn from the National Educational Longitudinal Study (NELS), a nationally representative sample of eighth graders in 1988 that were re-surveyed every two years until 1994. NELS contains detailed information on family background, cognitive ability, school achievement and attitudes toward school as early as junior high. The 1994 survey includes detailed information regarding postsecondary schooling and employment. For students who attended more than one college, I take the type, enrollment status and selectivity of the institution that the respondent attended for the longest period.

As a proxy for the college premium of individual  $i$  in state  $s$  with gender  $g$ , I use the earnings differential for 25-34 year old, full-time workers (at least 35 hours per week) of gender

g in state  $s$ . The differential is computed as the log difference between the median weekly earnings of college graduates (B.A. or higher) and high school graduates. These measures are constructed using census data drawn from the 1990 PUMS. This state level measure may be thought of as an average premium for individuals in the state.

I utilize four measures of non-cognitive ability. Two measures—middle school grades and the number of hours spent on homework per week in eighth grade—reflect effort and achievement in school. Grades are measured on a scale of zero to four, where a four indicates an “A” average, a three indicates a “B” average, etc. The idea here is that, conditional on cognitive ability, school achievement is determined by a variety of non-cognitive skills such as the ability to follow directions, work in groups, pay attention in class, and organize materials. Student behavior serves as another measure of non-cognitive skill because it reflects the ability to successfully negotiate the school environment. I create a composite measure of disciplinary incidents from a set of NELS questions that ask students to indicate how often during the last semester they were sent to the office for misbehaving, received a warning about their attendance, grades or behavior and got into a physical fight with another student.<sup>5</sup> Finally, an indicator of whether the child had ever been retained in grade during elementary school serves as a fourth measure of non-cognitive skills insofar as retention decisions are largely based on teacher perceptions of a student’s social maturity and behavior.

Among the 12,585 students in the NELS 1988-94 panel sample (i.e., students with survey data for all four waves), 67.3 percent of women had some postsecondary education by 1994 in comparison to only 62.3 percent of men. These attendance rates are somewhat higher than those documented in other national samples. For example, in the 1990 Census, nearly 51.4 percent of

---

<sup>5</sup> Principal components analysis was used to create a single composite from the NELS items BY55A-F described above.

19-21 year old women had attained some postsecondary experience compared with only 43.4 percent of males. The comparable figures for the 1993 October Current Popular Survey are 54.3 percent and 48.7 percent. One possible explanation for this difference involves sample attrition in NELS—i.e., those least likely to attend college have dropped out of the study and would not be included in the 1988-94 panel. In addition, it is possible that the detailed nature of the NELS survey questionnaire identified schooling experiences that individuals would not have self-reported in the simpler Census or CPS surveys. However, the magnitude of the gap closely resembles the numbers one obtains from other sources and a more detailed breakdown of differential attendance patterns by demographic group is consistent with data on the distribution of undergraduates nationally (Horn & Berktold 1998).

I delete 1,660 cases (13 percent of the total) because of missing data in one or more key variables such as student demographics, college attendance, high school graduation, non-cognitive skills and college premiums.<sup>6</sup> This leaves a total of 10,925 students, of which 9,734 graduated high school and 6,877 were high school graduates in the bottom three quartiles of the SES distribution (based on the entire NELS sample). While this reduced sample appears slightly more advantaged than the full panel sample, the differences do not seem unreasonably large. For example, 82 percent of the full panel graduated high school and 65 percent attended some college by 1994 in contrast to 86 percent and 68 percent in the reduced sample. Similarly, the students in the full sample came from families at the 51<sup>th</sup> percentile of the SES distribution while the average in the reduced sample was at the 53<sup>nd</sup> percentile. In order to estimate models that include measures of high school achievement, labor force participation in high school and family formation after high school, I imputed values for cases with missing data on these variables. In

the estimation, I include dummy variables to indicate whether the value of each variable was imputed.<sup>7</sup>

Summary statistics for men and women in the sample are shown in Table 1. There are few differences between boys and girls in terms of cognitive ability or family background. In contrast, boys and girls have considerably different non-cognitive skills. For example, 18 percent of boys were retained in elementary school compared with 12 percent of girls. Forty percent of boys have demonstrated some type of behavior problem compared with only 20 percent of girls. Boys are disproportionately represented in remedial classes, get lower grades, and spend less time on homework than girls. These differences carry over into high school where girls have higher grades and are more likely to be in an academic track. There are also differences in the college premiums earned by young men and women. Young female college graduates earn roughly 55 percent more than female high school graduates in contrast to male graduates who earn only 40 percent more than their less educated peers.

---

<sup>6</sup> Cases with missing data in the following variables were discarded: race, gender, college attendance, urban/rural location, high school graduation, middle school grades, homework, retention and behavior problems, state, tuition, unemployment, college density, college premiums and employment shares.

<sup>7</sup> Values are imputed using predictions from a linear regression that includes student demographics, standardized test scores, parental education and income, state and other measures of academic achievement. While subsequent estimates using imputed values may still be biased, it is often preferable to simply deleting cases with missing data (Little & Rubin 1987).

#### 4. Empirical strategy

The gender gap could be due to differences in either the characteristics of young men and women or the way in which men and women respond to these characteristics. For example, women may be more likely to attend college because they have greater cognitive ability than men, or because cognitive ability is a stronger determinant of college attendance for women than for men.

I first examine whether there are significant gender differences in the determinants of college attendance for men and women by estimating the following the following linear probability model<sup>8</sup> separately for men and women:

$$(1) \quad C_{ig} = X_{ig} \beta_g + \varepsilon_{ig}$$

where the  $i$  and  $g$  subscripts reflect individuals and gender groups,  $C$  is an indicator of college attendance,  $X$  is a vector of college determinants including cognitive ability, family background, non-cognitive skills, college costs and returns to schooling, and  $\varepsilon$  is a stochastic error term.

While this comparison can determine whether there are statistically significant differences between the magnitude of various determinants for men and women, it does not tell how large a role such coefficient differences play in explaining the gap. To do so, I decompose the gender gap into the proportion that can be attributed to differences in observable characteristics between boys and girls and the remaining proportion. Such decompositions have a long history in studies of wage discrimination, beginning with the work of Oaxaca (1973) who decomposed the wage differential between men and women into a component due to differences in observable characteristics and a component due to unexplained differences in the returns to these characteristics, which he interpreted in terms of discrimination (Blinder 1973; Jones 1983;

---

<sup>8</sup> Probit models yield virtually identical results so a LPM was used for ease of interpretation.

Reimers 1983; Cotton 1988; Neumark 1988; Oaxaca & Ransom 1994; Oaxaca & Ransom 1999).

In the case of college attendance, we can begin by writing the difference between males and females in the average probability of attending college

$$(4) \quad \bar{C}_f - \bar{C}_m = \bar{X}_f' \hat{\beta}_f - \bar{X}_m' \hat{\beta}_m$$

where the  $\hat{\beta}_f$  and  $\hat{\beta}_m$  are taken from (3). This attendance difference can be separated into two components—one due to differences in observed characteristics and one due to differences in the estimated coefficients of the two groups. To do this decomposition, however, one must choose which set of coefficients to use (male or female). Following Neumark (1988) and Oaxaca and Ransom (1994), I use common coefficients estimated from a pooled regression of males and females,  $\hat{\beta}_p$ ,

$$(5) \quad \bar{C}_f - \bar{C}_m = [\bar{X}_f - \bar{X}_m]' \hat{\beta}_p + \left[ \bar{X}_f' (\hat{\beta}_f - \hat{\beta}_p) - \bar{X}_m' (\hat{\beta}_m - \hat{\beta}_p) \right]$$

In this decomposition, the first term can be interpreted as the part of the attendance differential due to differences in characteristics. The second term can be interpreted as the part due to differences in coefficients.<sup>9</sup>

Finally, I examine the role of characteristics by estimating a variation of (1) that includes both men and women (assuming that the “prices” are equal across genders) and includes a dummy variable for female:

$$(6) \quad C_i = \alpha(\text{Female}_i) + X_i\beta + \varepsilon_i$$

---

<sup>9</sup> Persell et. al. (1992) describe student characteristics as assets and the coefficients as measures of “asset conversion,” emphasizing the role of societal forces such as discrimination in determining college attendance differences. Alternative explanations might focus on differences in preferences or opportunity costs.

In this model, the coefficient on *Female* reflects the magnitude of the gender gap. By entering covariates one-by-one, we can examine which observable characteristics (if any) mediate the relationship between gender and college attendance.

## **5. The postsecondary experiences of young men vs. women**

To begin, one might ask to what extent the college gender gap can be explained by high school graduation rates. For example, it may be the case that boys are simply less likely to graduate high school but, conditional on graduation, boys are equally (or more) likely to attend college. Table 2 shows that this is not the case.<sup>10</sup> While boys are slightly less likely to graduate high school than girls, the college attendance gap is roughly five percentage points among all high school graduates. Conversely, boys are more likely to be employed and in the military. Roughly the same number of boys and girls are neither in college, working nor in the military.

Conditional on attending college, women are equally likely to attend a four-year college. Women are roughly 33 percent more likely to attend a vocational or trade school, largely due to the predominance of women in programs for clerical workers. There are no significant differences between men and women in terms of full-time versus part-time enrollment status. Finally, conditional on attending a four-year college, men appear to enroll in slightly more selective institutions. Most notably, men attend colleges in which students score about ten points higher on the math portion of the SAT. Assuming the ability distribution of high school graduates is roughly equivalent for men and women, this slight advantage for women merely reflects the fact that males who attend college come from a somewhat higher point on the ability distribution on average. It is likely that the more noticeable advantage in math SAT scores also

reflects the preference for men to enroll in schools that are more heavily oriented towards math, science and engineering.

## **6. The Determinants of College Attendance for Men and Women**

As mentioned above, the gender gap could be due to differences in either the characteristics of young men and women or the way in which men and women respond to these characteristics. This section explores the latter. The reasons that boys and girls in the NELS cite for not attending college suggest that they approach the college attendance decision differently. For example, 37 percent of boys cite a “dislike of school” compared with 26 percent of girls. Similarly, 49 percent of boys say that they would rather work and earn money than go to college compared with 41 percent of girls and 23 percent of boys claim that they will not need more school for their jobs compared with 16 percent of girls.

In order to provide a more formal examination of the college enrollment decision, Table 3 presents the results from the linear probability model shown in equation (1), estimated separately for men and women. Note that the asterisks indicate whether the male and female coefficients are statistically different from each other. In general, the traditional determinants of college attendance appear significant and in the expected direction. Among women, for example, a 25 percentage point increase in the SES distribution (roughly one standard deviation in this sample) results in a 7.5 percentage point increase in the likelihood of attending college while a one standard deviation increase in math achievement coincides with more than a three percentage point increase in the probability of college attendance. Non-cognitive skills appear to have effects on college enrollment that are of the same magnitude as traditional determinants

---

<sup>10</sup> As described above, the NELS sample I use here is likely not representative of more marginal students due to sample attrition and missing data. It is possible that high school graduation may account for more of the gender gap

such as family background and cognitive ability. For example, a one standard deviation increase in middle school grades is associated with a 7.7 percentage point increase in the probability of enrollment for boys. Male high school graduates who repeated a grade during elementary school are roughly seven percentage points less likely than their peers to attend college. The measures of direct and opportunity costs and returns to schooling have less pronounced effects.

There are also some interesting differences between men and women. It appears that family background is a stronger determinant of attendance for men whereas cognitive ability (particularly reading ability) is a stronger determinant for women. While growing up in single parent household decreases the probability of college attendance by roughly three percentage points for boys, it does not have a significant impact on enrollment for girls. This is consistent with the literature that emphasizes the importance of male role models for boys' development (Hetherington, Bridges & Insabella 1998).

Finally, Table 3 provides some tentative evidence regarding the importance of opportunity costs in college enrollment. Insofar as manufacturing and construction are still largely male dominated industries, one would expect that the importance of these industries in a local economy would provide greater economic opportunities for low-skilled men than low-skilled women. If we believe that opportunity costs play an important role in the college attendance decision, we might expect the gender gap to be larger in states with a strong construction or manufacturing base. Similarly, the gender gap may be larger in rural areas where agricultural or other "male" jobs are more common. Consistent with this view, the point estimates for living in a rural area or in a state with a larger construction base are larger for boys than girls.

---

among this group. However, since this study focuses on high school graduates, this is not a significant concern here.

While one can reject the equivalence of the separate male and female regressions ( $F=5.84$ ,  $p<0.000$ ), the decomposition results presented in the bottom panel of Table 3 suggest that differences in observable characteristics, rather than differences in how men and women respond to these characteristics, explain the majority of the observed difference in college attendance rates. Of the total 4.9 percentage point gap, only 0.2 percentage points cannot be explained by differences in observable characteristics.

### **7. The role of non-cognitive skills and college premiums in the college gender gap**

Having demonstrated that the gender gap is due primarily to differences in observable characteristics, we now examine which of these characteristics explain the enrollment differential. Table 4 presents OLS estimates from equation (6). The Column 1 shows the unconditional gap, which is roughly five percentage points among all high school graduates and seven percentage points among lower SES graduates. In Columns 2 to 5, we add in potential explanatory variables. As expected, family background and cognitive ability do not substantially affect the gap. However, when we add measures of non-cognitive skills in Column 4, the gap drops almost three percentage points—roughly 40 percent—among all high school graduates. When we control for the college premium and other economic factors, the gap drops another two percentage points and is no longer statistically significant.<sup>11</sup> This suggests that among students with similar family background, cognitive ability and non-cognitive skills, in states with comparable college premiums for men and women, there is no appreciable difference between the college attendance rates of male and female high school graduates.

---

<sup>11</sup> Note that the college premium is driving the effect on the female coefficient because it is the only one of the economic variables that is correlated with gender and, in this model, the effect of explanatory variables is constrained to be identical for men and women.

The final two columns of Table 4 include controls for high school achievement, work experience, marriage and fertility. While these factors are clearly endogenous to the schooling decision, the results provide a useful description of the college attendance process. High school achievement is positively related to college attendance and working 20 hours a week or more during high school is negatively related to enrollment. Since girls have generally higher high school achievement and are less likely to work during this time, when we control for these factors, the gender gap shrinks even more. Marriage and fertility (before age 21) are negatively associated with enrollment and higher among women, so when we control for these factors, the gender gap increases.

Note that when we control for high school achievement and work experience as well as family formation, the influence of non-cognitive skill diminishes but remains significant. In fact, conditional on a wide range of high school achievement measures as well as 12<sup>th</sup> grade standardized test scores and work experience, the self-reported measure of middle school grades continues to exert a substantial influence on college enrollment.

## **8. Conclusions**

While the gender gap in higher education has received increasing attention in recent years, it has gone largely unexplained in the empirical research literature. The most common explanations of “where the boys are,” military and prisons, cannot explain the college enrollment differentials between men and women. This study suggests higher returns to college and greater non-cognitive skills among women account for nearly 90 percent of the gap.

Given the influence of non-cognitive ability on educational attainment, it would be interesting to know whether certain school characteristics foster the development of such skills

and if certain types of curriculum, pedagogy or learning environments are more effective for one gender than the other.<sup>12</sup> This would not only deepen our understanding of the differential college enrollment, but also may provide ways in which schooling could be restructured to foster greater success for boys. In addition, these findings raise questions about the differential trends in college attendance over the past three or four decades. Male and female enrollment among new high school graduates reached parity in the mid-to-late seventies, and female attendance has since outpaced that of men. While a variety of potential explanations come quickly to mind, including changes in marriage and fertility rates, labor force participation and returns to schooling, the simultaneity of these changes make it is extremely difficult to identify independent causal factors underlying the rise in female attendance.

Finally, the findings presented suggest that non-cognitive factors may play a larger role in other areas generally dominated by more concrete measures. Here we find that grades and disciplinary incidents have a substantial impact on the probability of enrolling in college, even after controlling for cognitive ability, family background, and high school achievement. The effect of non-cognitive skill is comparable in magnitude to socio-economic status and cognitive ability. Given the importance in college attendance decisions, it is likely that non-cognitive factors influence employment, occupation choice, and a variety of other labor market outcomes.

### **Acknowledgements**

I would like to thank Thomas DeLeire, Susan Dynarski, Robert LaLonde, Steven Levitt, Susan Mayer, and the participants of the Harris School workshop for many helpful comments and

---

<sup>12</sup> Prior studies have shown that communal school organization and small size are not only positively correlated with achievement and engagement of middle and high school students (Lee & Smith 1993; Lee & Smith 1995), but may influence gender differences in achievement (Lee, Chen & Smerdon 1996). Similarly, there is evidence that Catholic schools are particularly effective for low-SES and African-American males (Bryk, Lee & Holland 1993; Neal 1996).

suggestions. I am also grateful to Aurora D'Amico at NCES and Fay Booker at SRC for data assistance, and to Richard Ziebler at the College Board for providing some of the data used in this study. All remaining errors are my own. Please direct all correspondence to [Brian\\_Jacob@harvard.edu](mailto:Brian_Jacob@harvard.edu).

## References

- Altonji, J. G. (1993). The demand for and return to education when education outcomes are uncertain. *Journal of Labor Economics* 11(1), 48-83.
- Bae, Y., Choy, S., Geddes, C., Sable, J. & Synder, T. (2000). Trends in educational equity of girls and women. Washington, DC, National Center for Education Statistics, Department of Education.
- Blau, F. D. (1998). Trends in the well-being of american women, 1970-1995. *Journal of Economic Literature* 36(March), 112-165.
- Blinder, A. (1973). Wage discrimination: Reduced form and structural estimates. *Journal of Human Resources* 8(4), 436-455.
- Bryk, A. S., Lee, V. E. & Holland, P. B. (1993). *Catholic schools and the common good*. Cambridge, MA: Harvard University Press.
- Census (1990). Public use microdata sample. Washington, D.C., U.S. Bureau of the Census.
- Cotton, J. (1988). On the decomposition of wage differentials. *Review of Economics and Statistics* 70, 236-243.
- Duncan, G. J. & Dunifon, R. (1998). "soft-skills" and long-run labor market success. *Research in Labor Economics* 17, 123-149.
- England, P. & Farkas, G. (1986). *Households, employment and gender: A social, economic and demographic view*. New York: Aldine.
- Hetherington, E. M., Bridges, M. & Insabella, G. (1998). What matters? What does not? Five perspectives on the association between marital transitions and children's adjustment. *American Psychologist* 53(2), 167-184.

- Horn, L. J. & Berkthold, J. (1998). Profile of undergraduates in u.S. Postsecondary education institutions: 1995-96. Washington, D.C., National Center for Education Statistics.
- Jones, F. L. (1983). On decomposing the wage gap: A critical comment on blinder's method. *Journal of Human Resources* 18(1), 126-130.
- Justice (1997). Sourcebook of criminal justice statistics. Washington, D.C., U.S. Department of Justice.
- Kane, T. (1994). College entry by blacks since 1970: The role of college costs, family background and the returns to education. *Journal of Political Economy* 102(5), 878-911.
- Katz, L. F. & Murphy, K. M. (1992). Changes in relative wages, 1963-1987: Supply and demand factors. *The Quarterly Journal of Economics* 107(1), 35-78.
- Kleinfeld, J. (1998). The myth that schools shortchange girls: Social science in the service of deception.
- Koerner, B. I. (1999, February 8). Where the boys aren't. *U.S. News and World Report*, 47-55.
- Lee, V., Chen, X. & Smerdon, B. A. (1996). The influence of school climate on gender differences in the achievement and engagement of young adolescents. Washington, D.C., American Association of University Women Educational Foundation: 1-96.
- Lee, V. E. & Smith, J. B. (1993). Effects of school restructuring on the achievement and engagement of middle-grade students. *Sociology of Education* 66(July), 164-187.
- Lee, V. E. & Smith, J. B. (1995). Effects of high school restructuring and size on early gains in achievement and engagement. *Sociology of Education* 68(October), 241-270.
- Levy, F. & Murnane, R. J. (1992). U.S. Earnings levels and earnings inequality: A review of recent trends and proposed explanations. *Journal of Economic Literature* 30(September), 1333-1381.

- Little, R. J. A. & Rubin, D. B. (1987). *Statistical analysis with missing data*. New York: John Wiley & Sons.
- Mortenson, T. G. (1999). Where are the boys? The growing gender gap in higher education. *The College Board Review* August(188), 8-17.
- Murnane, R. & Levy, F. (1996). *Teaching the new basic skills*. New York: Martin Kessler Books.
- Murnane, R. J., Willet, J. B. & Levy, F. (1995). The growing importance of cognitive skills in wage determination. *The Review of Economics and Statistics*, 251-266.
- Murphy, K. M. & Welch, F. (1989). Wage premiums for college graduates: Recent growth and possible explanations. *Educational Researcher* 18(May), 17-26.
- Murphy, K. M. & Welch, F. (1992). The structure of wages. *Quarterly Journal of Economics* 107(February), 215-236.
- Neal, D. (1996). The effects of catholic schooling on educational achievement. Chicago, IL, University of Chicago.
- Neumark, D. (1988). Employer's discriminatory behavior and the estimation of wage discrimination. *Journal of Human Resources* 23(2), 279-295.
- Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International Economic Review* 14(3), 693-709.
- Oaxaca, R. & Ransom, M. (1994). On discrimination and the decomposition of wage differentials. *Journal of Econometrics* 61(1), 5-21.
- Oaxaca, R. L. & Ransom, M. R. (1999). Identification in detailed wage decompositions. *The Review of Economics and Statistics* 81(1), 154-157.

- Persell, C. H., Catsambis, S. & Cookson, P. W. J. (1992). Differential asset conversion: Class and gendered pathways to selective colleges. *Sociology of Education* 65(3), 208-225.
- Reimers, C. (1983). Labor market discrimination against hispanic and black men. *Journal of Economics and Statistics* 65(4), 570-579.
- Reskin, B. & Roose, P. (1990). *Job queues, gender queues: Explaining women's inroads into male occupations*. Philadelphia: Temple University Press.
- Sommers, C. H. (2000, May). The war against boys. *The Atlantic Monthly*, 59-74.

**Table 1**  
**Summary statistics**

<b>Variables</b>	<b>Men</b>	<b>Women</b>
College attendance rate	0.682 (0.466)	0.729 (0.445)
African-American	0.075 (0.263)	0.095 (0.293)
White	0.719 (0.449)	0.698 (0.459)
Hispanic	0.122 (0.328)	0.128 (0.334)
Urban	0.275 (0.447)	0.278 (0.448)
Rural	0.321 (0.467)	0.327 (0.469)
SES (percentile)	53.207 (28.649)	50.989 (28.878)
Family size	4.592 (1.367)	4.687 (1.416)
Single parent HH	0.182 (0.385)	0.182 (0.385)
8 <sup>th</sup> grade composite ability	51.850 (9.998)	52.147 (9.876)
8 <sup>th</sup> grade reading score	26.967 (8.481)	28.290 (8.470)
8 <sup>th</sup> grade math score	37.661 (12.052)	36.497 (11.645)
Middle school grades	2.932 (0.741)	3.065 (0.704)
Behavior problem	0.405 (0.491)	0.195 (0.396)
Behavior composite	0.296 (1.801)	-0.435 (1.250)
Hours homework/wk	5.877 (4.962)	6.208 (4.890)
Repeated a grade in elementary school	0.175 (0.380)	0.115 (0.319)
12 <sup>th</sup> grade composite ability	51.458 (9.602)	51.469 (9.266)
12 <sup>th</sup> grade math score	50.148 (14.176)	47.759 (13.786)
12 <sup>th</sup> grade reading score	32.584 (9.773)	34.283 (9.238)
Class rank (percentile)	0.479 (0.281)	0.564 (0.271)
High school grades	-7.359 (2.448)	-6.780 (2.378)
Academic track	0.703 (0.456)	0.717 (0.449)
Vocational track	0.068 (0.251)	0.050 (0.217)
Have children by 1994	0.069 (0.253)	0.159 (0.365)

Married by 1994	0.055 (0.227)	0.113 (0.316)
Worked 1-19 hrs in H.S.	0.484 (0.435)	0.564 (0.434)
Worked 20+ Hrs in H.S.	0.301 (0.394)	0.245 (0.370)
Hours of homework/wk in 12 <sup>th</sup> grade	0.974 (0.148)	0.985 (0.113)
Avg. tuition at 2 yr college in state (\$/100)	11.529 (5.495)	11.466 (5.385)
College density in state (# schools / 1000 residents)	0.006 (0.003)	0.007 (0.003)
1992 state unemployment rate (%)	7.389 (1.371)	7.394 (1.364)
Median weekly earnings (\$)	355.896 (43.351)	354.794 (42.039)
College premium	0.401 (0.067)	0.549 (0.061)
Construction share (%)	4.321 (0.707)	4.337 (0.704)
Manufacturing share (%)	16.178 (4.908)	16.163 (4.969)
Number of observations	5,158	5,767

Notes: Standard deviations shown in parentheses.

**Table 2**  
**The gender gap in higher education among high school graduates**

	Full sample		Bottom 75% of SES	
	Women	Men	Women	Men
<b>Probability of graduating HS</b>				
<b>Postsecondary experience</b>				
College	0.798**	0.749	0.741**	0.672
Employment, no college	0.172**	0.210	0.221**	0.276
Military, no college	0.003**	0.021	0.004**	0.027
<b>Type of college<sup>a</sup></b>				
4 year college	0.597	0.600	0.497	0.491
2 year college	0.294	0.304	0.365	0.381
Vocational school	0.050**	0.037	0.079**	0.053
Other / missing	0.050*	0.059	0.059**	0.075
<b>Enrollment status<sup>a</sup></b>				
Full-time	0.874	0.872	0.842	0.840
Half-time	0.079	0.082	0.103	0.102
Less than half-time	0.045	0.045	0.055	0.058
<b>Quality of 4-Year College<sup>b</sup></b>				
Composite quality measure	990**	1000	952**	962
Acceptance rate	0.718*	0.709	0.746	0.743
High school GPA > 3.0 (%)	62.9**	65.9	57.8**	61.0
SAT verbal – 25 <sup>th</sup> percentile	431	432	408	408
SAT verbal – 75 <sup>th</sup> percentile	543	545	520	523
SAT math – 25 <sup>th</sup> percentile	483**	489	456*	462
SAT math – 75 <sup>th</sup> percentile	604**	614	581**	591
ACT – 25 <sup>th</sup> percentile	20.2	20.4	19.3	19.5
ACT – 75 <sup>th</sup> percentile	35.3*	38.8	36.2*	42.1
<b>Number of observations<sup>c</sup></b>				
	5,161	4,573	3,703	3,174

Notes: The cells show the probabilities of participation in each of the postsecondary experiences. <sup>a</sup> Conditional on attending college. <sup>b</sup> Conditional on attending a four-year college. <sup>c</sup> The number of observations is the total sample and is lower for statistics presented on the type of college, enrollment status and college quality. \* indicates the coefficients are different at a 10% confidence level; \*\* indicates the coefficients are different at a 5% confidence level.

**Table 3**  
**The determinants of college attendance for men vs. women**

Independent variables	Men	Women
Black	0.067 (0.027)	0.070 (0.020)
Hispanic	0.082 (0.024)	0.059 (0.019)
Other ethnicity	0.030 (0.016)	0.030 (0.017)
Urban	0.010 (0.015)	0.006 (0.009)
Rural	-0.029 (0.014)	-0.019 (0.014)
SES (in percentiles)	0.004** (0.000)	0.003 (0.000)
Family size	-0.008 (0.006)	-0.014 (0.005)
Missing HH info	0.010 (0.057)	-0.049 (0.054)
Single parent	-0.028* (0.016)	0.004 (0.016)
8th grade math	0.003 (0.001)	0.003 (0.001)
Missing math test	0.194 (0.127)	-0.130 (0.252)
8th grade reading	0.001** (0.001)	0.004 (0.001)
Missing reading test	-0.216 (0.112)	0.153 (0.244)
Middle school grades	0.102 (0.010)	0.097 (0.10)
Behavior	-0.021 (0.006)	-0.015 (0.008)
Hours of homework	0.0004 (0.0010)	0.0012 (0.0009)
Repeated a grade in elementary school	-0.071 (0.018)	-0.089 (0.022)
College premium	0.213 (0.101)	0.232 (0.145)
Median weekly earnings (\$)	0.0003 (0.0002)	0.0006 (0.0002)
1992 state unemployment rate	0.004 (0.006)	0.006 (0.005)
Avg. tuition in 2-yr colleges in state (in \$100)	0.0007** (0.0007)	-0.0028 (0.0010)
College density in state (# schools per 1000 residents)	2.489 (2.594)	-0.686 (2.365)
Construction share of employment in state (%)	-0.026 (0.012)	-0.010 (0.012)
Manufacturing share of employment in state (%)	-0.002 (0.001)	-0.001 (0.001)
Constant	0.022 (0.157)	-0.099 (0.214)

Number of observations	4,573	5,161
R-squared	0.201	0.185
F-test of equivalence of regressions	F=4.99 (p<0.000)	

### Decomposition

Characteristics	$[\bar{X}_f - \bar{X}_m] \hat{\beta}_p$	0.047
Coefficients	$[\bar{X}_f'(\hat{\beta}_f - \hat{\beta}_p) - \bar{X}_m'(\hat{\beta}_m - \hat{\beta}_p)]$	0.002
Total	$\bar{C}_f - \bar{C}_m$	0.049

Notes: \* indicates the coefficients are different at a 10% confidence level; \*\* indicates the coefficients are different at a 5% confidence level.

**Table 4**  
**Regression of college attendance on gender, non-cognitive skills and the returns to schooling**

	1	2	3	4a	4b	5	6	7
<b>Female coefficient</b>								
HS graduates (n=9,734)	0.048 (0.007)	0.061 (0.007)	0.057 (0.007)	0.028 (0.007)	0.029 (0.016)	0.005 (0.016)	-0.005 (0.016)	0.011 (0.016)
HS graduates in the bottom 3 quartiles of the SES distribution (n=6,877)	0.069 (0.011)	0.079 (0.010)	0.075 (0.009)	0.044 (0.009)	0.041 (0.021)	0.019 (0.021)	0.003 (0.020)	0.023 (0.019)
<b>Demographics &amp; family background</b>	No	Yes						
<b>Cognitive ability</b>	No	No	Yes	Yes	Yes	Yes	Yes	Yes
<b>Non-cognitive skills</b>	No	No	No	Yes	No	Yes	Yes	Yes
<b>College premium &amp; other economic variables<sup>a</sup></b>	No	No	No	No	Yes	Yes	Yes	Yes
<b>High school achievement &amp; work experience</b>	No	No	No	No	No	No	Yes	Yes
<b>Family formation at age 20</b>	No	Yes						

Note: OLS regressions with robust s.e. in parentheses. <sup>a</sup> Economic variables include median weekly earnings, unemployment rate, tuition, college density and employment shares in manufacturing and construction.