CENTRE for ECONOMIC PERFORMANCE

CEP Discussion Paper No 1220

May 2013

Explaining Cross-Racial Differences in the Educational Gender Gap

Esteban Aucejo

THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

Abstract
The sizable gender gap in college enrollment, especially among African Americans, constitutes a puzzling empirical regularity that may have serious consequences on marriage markets, male labor force participation and the diversity of college campuses. For instance, only 35.7 percent of all African American undergraduate students were men in 2004. Reduced form results show that, while family background covariates cannot account for the observed gap, proxy measures for non-cognitive skills are crucial to explain it. Moreover, a sequential model of educational attainment indicates that males have actually higher preferences for education than females after controlling for latent factors (i.e. cognitive and non-cognitive skills). The model also shows that cognitive skills strongly affect the decision to move from one school level to the next, especially after finishing high school, but cannot account for disparities between genders. On the contrary, the substantial differences in the distribution of non-cognitive skills between males and females make these abilities critical to explain the gender gap in educational attainment across and within races.

Keywords: Gender gap, college enrollment, non-cognitive skills, cognitive skills, race
JEL Classifications: I2, J15, J16

This paper was produced as part of the Centre's Education and Skills Programme. The Centre for Economic Performance is financed by the Economic and Social Research Council.

Acknowledgements
I would like to thank Peter Arcidiacono, Pat Bayer, Federico Bugni, Joe Hotz, Marjorie McElroy, Steve Pischke and Seth Sanders. Any errors are my own.

Esteban Aucejo is an Associate of the Centre for Economic Performance and a Lecturer (Assistant Professor) in Economics, London School of Economics and Political Science.

Published by
Centre for Economic Performance
London School of Economics and Political Science
Houghton Street
London WC2A 2AE

.

1 Introduction

The gender composition of U.S. college campuses has changed dramatically since the 1950's. In 1950, males represented 68% of college enrollees but by 1970 this number had fallen to 52% [1]. Goldin et al. (2006) indicate that the elimination of institutional and social barriers that prevented women from pursuing higher education coupled with changes in expectations, labor force participation, age at first marriage and improvements in high school performance contributed to a convergence between genders.

But not only did women catch up to men in terms of college enrollment, the last three decades have seen women pass men, and by a substantial margin. According to the National Center of Educational Statistics (NCES), females represented 57% of the total fall enrollment in degree-granting institutions in 2004. Moreover, gender disparities are robust to four and two years colleges, and are larger for those who graduate. However, the gap size varies substantially across races; for example, 56% and 64% of all white and African American undergraduates enrolled in 2004 were women[2]. Indeed, the fact that the difference in the total population proportion[3] of white and black females attending college (13%) is smaller than the proportion between black females and black males (17%) denotes the importance of the gender imbalances. In this regard, a detailed analysis of this empirical regularity, especially among African Americans is key for two main reasons. First, substantial gender differences within race could suggest the importance of developing public policies that target specific subgroups of the population (e.g. black males) that may be at higher risk of poor educational outcomes. Second, the sizable gender disparities may have considerable consequences for many future outcomes. A growing decline in the number of college educated males would imply that women who would like to marry a man with a similar education background will face a marriage squeeze[4], particularly in the case of black women[5]. In addition, the lack of college educated men has implications in labor markets. Among black males between 24 and 50 years old,

[1] U.S. Department of Education, NCES, Digest of Education Statistics, 2008. Table 189

[2] U.S. Department of Education, NCES, Digest of Education Statistics, 2005. Table 23.1. See also McDaniel et al (2011) for an analysis of historical trends in the educational gender gap among African Americans.

[3] The previous percentages correspond to the sample of just college enrollees.

[4] Evidence indicates that college females are having trouble to find a date on campus. According to the National Longitudinal Survey of Freshman, while only 35.8% of black males reported not to have had a date (during the junior year) with people met on campus, 50.3% of black females did so. Substantial anecdotical evidence can be found in newspaper articles; see for example: "The New Math on Campus" (The New York Times, February 2010); "A Tough Time to Be a Girl: Gender Imbalance on Campuses" (The Chronicle of Higher Education, July 2008).

[5] For instance, 42.1% of 24 to 45 years old college educated black women have never got married, as compared to 27.9% of white women. Source: IPUMS-CPS (2009)

only 65% of high school dropouts or graduates were able to secure a full-time job, while 87% of Bachelor's degree holders did so in 2008 (IPUMS-CPS). Finally, the shortage of black males may weaken efforts to increase campuses diversity.

The National Longitudinal Survey of Youth 1997 (NLSY97) shows similar gender imbalances. Reduced form results indicate that family background covariates can fully account for differences in college enrollment across races; however, these variables have no impact on the gender gap. On the contrary, the inclusion of proxy measures for non-cognitive/socio-emotional[6] skills at quite early stages of schooling career (i.e. grade retention, suspensions at school, GPA in grade eight, involvement in fights and precocious sex) can fully explain the gender disparities in college enrollment for all racial groups.

In order to further investigate the importance of non-cognitive and cognitive skills, a sequential model of educational attainment, together with a measurement system for the identification of latent factors (i.e. cognitive and non-cognitive abilities) are estimated. This approach provides four main advantages. First, it incorporates into the analysis the key fact that postsecondary attainment is the result of previous educational decisions; enrolling in college depends on graduating from high school which also depends on finishing grade 10 and so on. Second, a complete profile of the schooling career path of males and females can be recovered, thus helping to identify the educational levels where boys are more likely to leave education. Third, the inclusion of two latent factors associated with cognitive and non-cognitive skills makes it possible to control for dynamic selection and to deal with measurement error, given that noisy proxies likely provide biased estimates. Fourth, the effects of non-cognitive abilities can be distinguished from the cognitive ones[7]; therefore, changes in the relative importance of these skills can be determined at the different educational transitions[8]. Cameron and Heckman (1998) and (2001), Heckman, Stixrud and Urzua (2006), and Cunha and Heckman (2008), among others, have shown the relevance of factor models in order to account for the effects of skills and family background characteristics in educational attainment. For instance, Heckman, Stixrud and Urzua (2006) show that both cognitive and non-cognitive skills have considerable effects on graduating from four-year institutions. Heckman et al's approach is methodologically extended in this work to examine differences across genders.

[6] Cognitive skills reflect an individual's ability to think. The terms socio-emotional and non-cognitive skills, used as synonyms in this manuscript, reflect the ability to: understand and manage feelings, follow appropriate social behaviors and develop manners. According to psychologists, socio-emotional abilities are critical because they facilitate engagement in learning, promote positive peer relationships, buffer children against risk and benefit mental health.

[7] Covariates such as school GPA, which are generally included in OLS regressions strategies, are most likely a function of cognitive and non-cognitve abilities.

[8] Isolating the effect of each skill may also help to explain differences in socioeconomic outcomes. Moreover, policy recommendations may be different if non-cognitive abilities turn out to be important for educational attainment; according to Cunha et al. (2005) these skills are more malleable at later stages of life than the cognitive ones.

The estimation results reveal that disparities in skills, especially in non-cognitive ones, can explain the gender gap across all races. Indeed, males are shown to have “higher preferences” for educational attainment than females after controlling for the latent factors. For example, gender differences in college enrollment among African Americans would be 29% higher if black males had preferences for schooling similar to those of black females. This result is consistent with part of the economics literature that finds disparate incentives for educational attainment between males and females. For instance, Becker et al. (2010) argue that the expected benefits of schooling are still higher for males than for females; and Hubbard (2011) shows that college premium for women is not larger than for men once topcoding biases (in CPS survey) are corrected.

In addition, estimation results show that the puzzling gap size between African American males and females is mainly explained by the substantial gender differences in non-cognitive skills distribution[9]. For instance, simulation exercises indicate that if black gender disparities in skills mirrored the white ones, then the size of the gap would be the same for both races.

Policy recommendations that intend to improve educational attainment or close the gender gap may depend on the relative importance of one ability over the other. In this regard, cognitive skills show a greater impact (conditional on reaching certain grade) on the probability of transitioning from one schooling level to the following than non-cognitive skills do, especially after finishing high school. However, the substantial disparities in the distribution of non-cognitive skills between males and females make these abilities more relevant in terms of the gender gap size. Results indicate that if young men had the non-cognitive distribution of women, they would close the gender differences in educational attainment. On the contrary, this outcome could not be obtained if, instead, cognitive skills distributions were equalized. Finally, an analysis of the changes in the mean of the factors distributions at each transition of schooling career[10] suggests that selection into college is driven by both skills but with a higher emphasis on cognitive ones. However, the considerable gender disparities in non-cognitive abilities prevent many males (relative to females) from finishing high school; which is a necessary step to enroll in postsecondary education. To sum up, differences in skills between males and females at early stages in life can fully explain the disparities in educational attainment.

This work builds on Jacob (2002) which is the first paper that attributes to non-cognitive skills a key role in explaining the gender gap in college enrollment. However, his findings are somewhat weaker than the ones presented in this manuscript. For instance, Jacob shows that these skills only account for 42% of the male-female disparities, while in this paper non-cognitive skills fully account for the gender gap. Many reasons could explain the differences. First, the NLSY97 includes a more complete set of measures of non-cognitive skills than Jacob’s dataset (NELS 88:94). Moreover, his

[9] This result also applies for the other racial groups.

[10] Remember that in each schooling transition a given proportion of students is leaving the education system, which leads to changes in the distribution of skills (i.e. selection process).

sample is only based on high school graduates in which college attendance rates are higher than those documented in other national surveys. While the 1990 Census shows that 51.4% of 19-21 year old women had some postsecondary education, the corresponding number in Jacob's dataset is 67.3%.

The rest of the document is organized as follows: section 2 describes the data and the gender gap in detail. Section 3 shows reduced form results. Section 4 presents a factor model of educational attainment. Section 5 describes the estimation outcomes. Section 6 discusses implications of the model. Section 7 concludes.

2 Data: NLSY97

2.1 The Gender Gap in Depth

The empirical strategy of this paper is based on the National Longitudinal Survey of Youth of 1997 (NLSY97); therefore, before characterizing in more detail the gender gap in postsecondary attainment, a brief description of this database is suitable. The NLSY97 is a nationwide representative sample of youths who were 12 to 17 years old when they were first surveyed in 1997. It collects vast information on family background characteristics, educational experiences and labor market behavior through time, with the aim to document the transition from school to work and into adulthood of the survey participants.

The NLSY97 shows that gender disparity in educational attainment is an empirical regularity that can be found at all ethnic groups; however, its magnitude varies substantially across races. Table 1 indicates that the proportion of white males (before age 25) enrolled in college (52.8%) was considerably smaller than white females (63.1%)[11]. However, a wider gap can be shown among African Americans, being the percentage of black men and women enrolled in college 32.6% and 49.7%, respectively. In addition, notice that girls are significantly less likely to be high school dropouts than boys; for instance, 28.5% of black males could not obtain a high school diploma, while only 18.9% of black females were in that same situation. This last outcome is surprising given the high pregnancy rates of African American adolescents (17%)[12].

Table 1 also shows that disparities in college enrollment between black and white men (20.2%) are only 3.1% points higher than the gender gap among African Americans (-17.1%). Moreover, differences in postsecondary attainment between white and black females (13%) are substantially

[11] Percentages are expressed as a proportion of the total subsample populations (e.g. total white males). Part of the statistics presented in the introduction of the paper were expressed, instead, as proportion of undergraduate subsample populations (e.g. whites enrolled in college).

[12] National Vital Statistics System. U.S. Department of Health and Human Resources, CDC (2006). African American females of age 19 or less.

smaller than those between black females and males[13]. The fact that the size of the gender gap among African Americans is almost as big as the racial gap suggests that gender imbalances are relevant.

Educational Attainment as % of Total Demographic Group (Before age 25)						
	White		Black		Hispanic	
	Males	Females	Males	Females	Males	Females
High School Dropout	14.7%	12.2%	28.5%	18.9%	22.9%	19.2%
Only High School Degree	32.5%	24.7%	38.9%	31.4%	37.3%	35.5%
College Enrollment	52.8%	63.1%	32.6%	49.7%	39.8%	45.3%
Observations	2072	1908	1025	1078	843	829

Table 1: Educational attainment before age 25 as % of total demographic subsample. Percentages are now expressed as a proportion of the total subsample populations (e.g. white male). Previous tables were expressed as a proportion of undergraduate subsample populations (e.g. whites enrolled in college). Data: NLSY97

In order to provide robustness checks, summary statistics of gender disparities were also analyzed using the National Center of Educational Statistics (NCES) database. The NCES collects information on enrollment, major and graduation rates open by gender and race from each postsecondary institution in the US. NCES data shows patterns quite similar to the ones described above. For example, the difference in the percentage of undergraduate fall enrollment in degree-granting institutions between white males and females was -11.8% [14] in 2004; while among African Americans was -28.6% (i.e. approximately 2 out of 3 black students in college were females). Moreover, disparities are even bigger if degrees granted by two or four years institutions are considered (see Appendix A). This fact may suggest that apart from enrolling in higher proportion, females are more likely to persist and/or finish their studies in a shorter period of time than males. An analysis of career paths indicates that majors related to health professions and liberal arts and sciences are highly dominated by females; however, males are still a majority in engineering and computer science[15].

[13] This result is consistent with the findings on the CPS survey for the years 2000 and 2001, population of 18 and 19 years old by school enrollment status.

[14] Notice that this proportion was calculated based on the college enrolled sample. More specifically, the percentages of white males and white females enrolled in degree granting institutions conditional on total white enrollment in 2004 were 44.1% and 55.9% respectively; then, the difference was -11.8%. On the contrary, the proportions presented based on the NLSY97 were obtained considering the unconditional sample.

[15] See table A1 to A3 of Appendix A for more detailed information on gender differences across races, degrees granted and majors.

To sum up, the NLSY97 and the NCES statistics provide consistent evidence related to differences in educational attainment between males and females. Including into the analysis the racial dimension is critical due to, for example, the higher gender disparities among African Americans than the rest of the ethnic groups.

The empirical strategy of this paper makes intensive use of family background characteristics and proxies for cognitive and non-cognitive skills; therefore, the following two subsections provide a detailed description of these variables based on NLSY97 data.

2.2 Family Background Covariates

Family background characteristics have substantial impact on enrollment differentials based on race. However, these factors are less likely to account for the gender gap given that males and females belong on average to the same type of families. Table 2 shows means and standard deviations of selected family characteristics[16]: mother education[17], number of household members under the age of 18, family structure (i.e. dummy variable for broken family at age 12) and parenting style (i.e. dummy variables for authoritarian, authoritative, uninvolved or permissive). Parenting style is included in the analysis with the aim to control for any differential effects that parents-sons/parents-daughters relationships may have in educational attainment. Psychologist Diana Baumrind (1991) has identified four patterns of parenting styles based upon two aspects of parenting behavior: control and warmth. 1) Authoritarian Parenting: little warmth and highly controlling, 2) Authoritative Parenting: warmth but firm, 3) Permissive Parenting: warmth but undemanding, 4) Uninvolved Parenting: not warmth and undemanding.

A brief analysis of table 2 shows no surprising results. For instance, white mothers are more educated than their black and Hispanic counterparts. In addition, a test of differences in means cannot reject the null hypothesis of equal means across genders conditional on race. Similarly, white families present a smaller number of young household members, but again there are no statistical differences across genders. Finally, black kids belong in much higher proportion to broken families (at age of 12) than whites and Hispanics.

[16] Family income was not included due to the extensive number of missing values and inconsistencies. For example, it was found that siblings that lived in the same house and with the same parents reported quite different amounts. Empirical results do not change if family income is included in the empirical strategies.

[17] Mother education takes value 0 if a mother does not finish high school, 1 if she only finishes high school, 2 if she has some college, and 3 if she finishes four years college.

	Family Background: Means and Standard Deviations (NLSY97)					
	White		Black		Hispanic	
	Males	Females	Males	Females	Males	Females
Mother Education	1.62	1.62	1.19	1.20	0.90	0.90
	(0.987)	(0.979)	(0.927)	(0.923)	(0.965)	(0.984)
Family Members	2.25	2.25	2.60	2.61	2.61	2.63
Under Age 18	(1.08)	(1.14)	(1.45)	(1.45)	(1.30)	(1.35)
Broken Family	0.412	0.444	0.797	0.786	0.494	0.502
	(0.492)	(0.497)	(0.402)	(0.410)	(0.500)	(0.500)
Mother Uninvolved	0.118	0.137	0.087	0.134	0.107	0.172
	(0.322)	(0.343)	(0.282)	(0.341)	(0.309)	(0.377)
Mother Permissive	0.375	0.352	0.307	0.304	0.352	0.301
	(0.484)	(0.478)	(0.461)	(0.460)	(0.479)	(0.459)
Mother Authoritarian	0.112	0.142	0.128	0.154	0.127	0.173
	(0.315)	(0.349)	(0.333)	(0.361)	(0.332)	(0.378)
Mother Authoritative	0.395	0.369	0.478	0.408	0.414	0.354
	(0.489)	(0.482)	(0.499)	(0.492)	(0.492)	(0.479)

Table 2: Summary statistics (weighted): means and standard deviations for whites, blacks and Hispanics. Data: NLSY97

2.3 Cognitive and Non-Cognitive Proxies

An emerging literature in economics, mainly developed by Heckman and his coauthors[18]; has provided substantial evidence about the relevance of cognitive and non-cognitive skills in a wide number of outcomes. For instance, Neal and Johnson (1996) and Cawley, Heckman and Vytlacil (2001) show that cognitive abilities constitute an important predictor of educational attainment and labor outcomes. However, differences in cognitive skills are not the unique source of observed disparities in educational attainment. Heckman and Rubinstein (2001) point out that GED recipients have similar cognitive abilities to high school graduates with 12 years of schooling. They suggest that lower levels of non-cognitive skills likely explain the observed differences in high school completion and labor outcomes. This conjecture was later confirmed by Heckman, Stixrud and Urzua (2006), where they find out that GED recipients present a worse distribution of non-cognitive abilities than high school graduates.

[18] Cawley, Heckman and Vytlacil (2001), Heckman and Rubinstein (2001); Cunha, Heckman and Navarro (2005); Carneiro, Hansen and Heckman (2003); Heckman, Stixrud and Urzua (2006); Heckman, Lochner and Todd (2006); Cunha and Heckman (2008); and Heckman, Urzua and Veramendi (2010), among others.

In a similar vein, psychologists have also argued that children with low development of non-cognitive abilities are more likely to experience academic delays, enter school at risk of increasing behavior problems, be vulnerable to peer rejection, to dropout school and to be involved in criminal and risky activities[19]. Therefore, given the relevance of these skills in educational attainment, then it is suitable to analyze their possible contribution to the college gender gap.

The difficulty of finding accurate measures of non-cognitive skills[20] (used by psychologists) in massive surveys, makes it necessary to include proxies. In this sense, the following covariates (constructed until the age of 14) are considered as indirect and noisy measures of these abilities: suspensions at school, retention at school between grade 1 and 8[21], GPA at grade eight, involvement in fights and precocious sex. As it is shown in section 4, these variables are also considered as functions of cognitive skills. The main reason that explains why these proxies were constructed until the age of 14 is that students cannot make any educational decisions before this age; helping to avoid problems of endogeneity[22].

Table 3 shows means and standard deviations of this set of variables. Conditional on race, males are more likely to be suspended from school, to be involved in fights with the intention to hurt someone, to engage in precocious sex and to be retained in at least one grade. Moreover, African Americans show the highest differences between males and females in almost all of these variables, for example, 24.9% of black males were suspended from school at age 14 while only 13.4% black females were in that same situation. Furthermore, African American males are considerable more likely to engage in precocious sex than black females. This disparity is consistent with data from the Youth Risk Behavior Surveillance System (YRBSS), a cross-sectional, nationally representative survey of students in grades 9-12 established by the Centers for Disease Control and Prevention. Cavazos-Rehg et al. (2009) using this database show that by the 14th birthday, the likelihood of sexual debut is of 42% for African American males while it is of 17% for African American females. Performance in school indicates that females (conditional on race) obtained statistically significant higher GPA at grade 8 than males.

To sum up, the NLSY97 data shows that females (conditional on race) do better in all the described proxy measures. Robustness checks indicate that other national databases (as it is shown in the following section) can replicate the patterns in Table 3.

Finally, measures related to cognitive skills come from the Armed Forces Vocational Aptitude

[19] Domitrovich (2008), Nagin et al (2001), Shaw et al (2001), Payton et al, (2000), Brody et al. (2003), Ladd et al. (1999), Caspi et al. (1995) and White et al. (1990).

[20] Measures of non-cognitive skills are: antisocial behavior, conduct disorder, attention problems, anxiety among others.

[21] It has been argued that retention is based on teacher perceptions of a student's social maturity (Jacob, 2002)

[22] The implicit identifying assumption is that these proxies (which were constructed until age of 14) are not determined by a previous decision of not attending college. For example, the fact that a kid is suspended at school at age 11, cannot be driven by his/her decision of not attending college later in life.

	Non-cognitive Proxies: Means and Standard Deviations					
	White		Black		Hispanic	
	Males	Females	Males	Females	Males	Females
Suspensions at age 14	0.128	0.067	0.249	0.134	0.135	0.089
	(0.334)	(0.250)	(0.432)	(0.341)	(0.341)	(0.285)
Fights	0.210	0.101	0.274	0.174	0.202	0.136
	(0.407)	(0.302)	(0.446)	(0.379)	(0.402)	(0.343)
Precocious Sex	0.165	0.151	0.484	0.232	0.263	0.148
	(0.371)	(0.357)	(0.499)	(0.422)	(0.440)	(0.355)
Retention Grade 1 to 8	0.125	0.082	0.277	0.171	0.172	0.121
	(0.331)	(0.275)	(0.448)	(0.377)	(0.377)	(0.325)
GPA Grade 8 std.	-0.012	0.385	-0.448	0.004	-0.289	0.050
	(1.035)	(0.954)	(0.887)	(0.899)	(0.963)	(0.933)

Table 3: Summary statistics (weighted): means and standard deviations for whites, blacks and Hispanics. Note: "std." means standardized. Data: NLSY97

Battery (ASVAB) test scores (i.e. mathematical knowledge, arithmetic reasoning, word knowledge, paragraph comprehension, assembling objects and general sciences)[23]. ASVAB scores have been widely used in the economics literature as proxies for these skills[24]. Table 4 indicates that white and Hispanic males performed better than their female counterparts in general sciences, arithmetic reasoning and word knowledge, while the opposite occurred in math knowledge, assembling objects and paragraph comprehension. However, the picture is different among African Americans, where females performed better than males in all the categories.

The following section shows reduced form evidence that highlight the importance (if any) of family background characteristics, non-cognitive and cognitive skills to explain the gender gap in college enrollment.

3 Reduced Form Evidence

3.1 Family Background Characteristics

In order to analyze the role of family covariates in the gender gap, two OLS regressions were performed initially. The first one includes as dependent variable an indicator for college enrollment

[23] Coding Speed and Numerical Operations tests were not included because they were administered in a different format (i.e. non-adaptive, all respondents answer the same items in the same order).

[24] See references in footnote 17.

	Cognitive Proxies: Means and Standard Deviations					
	White		Black		Hispanic	
ASVAB Tests	Male	Female	Male	Female	Male	Female
General Sciences	0.483	0.300	-0.528	-0.518	-0.195	-0.357
	(0.977)	(0.850)	(0.889)	(0.842)	(0.930)	(0.868)
Arithmetic Reasoning	0.367	0.314	-0.567	-0.492	-0.140	-0.179
	(0.966)	(0.840)	(0.991)	(0.906)	(0.948)	(0.913)
Math Knowledge	0.225	0.372	-0.537	-0.309	-0.185	-0.138
	(0.962)	(0.920)	(0.948)	(0.938)	(0.965)	(0.923)
Assembling Objects	0.205	0.346	-0.578	-0.428	-0.028	0.014
	(1.102)	(0.916)	(0.861)	(0.892)	(0.958)	(0.928)
Word Knowledge	0.358	0.351	-0.531	-0.414	-0.223	-0.243
	(0.951)	(0.888)	(0.959)	(0.929)	(0.930)	(0.870)
Paragraph Comprehension	0.184	0.428	-0.605	-0.271	-0.230	-0.071
	(0.983)	(0.884)	(0.925)	(0.929)	(0.978)	(0.886)

Table 4: Summary statistics (standardized and weighted): means and standard deviations for whites, blacks and Hispanics. Data: NLSY97.

before age 25[25] and as independent variables gender, race and their interactions (this estimation works as the benchmark case). The second regression adds the following covariates: number of family members under age 18, mother education, parenting style and an indicator that denotes if a kid belonged to a broken family at age 12[26] [27]. Column 1 of Table 5 shows that the size of the gender gap is around -10% for whites and Hispanics, and -17% for African Americans. Moreover, this column also exhibits the presence of the well known disparities in educational attainment across races. The inclusion of family background covariates provides quite interesting results (column 3), while ethnic differences in college enrollment are fully explained, gender disproportions remain fairly constant for all racial groups. In this sense, the persistence of the gender gap after controlling for family covariates is consistent with the fact that males and females come (on average) from the same type of families; therefore no effect is expected.

[25] OLS regression results are presented because it is straightforward to interpret the coefficients. Logit specifications present similar outcomes.

[26] Measures of broken family at age 2 or 6 do not change the results.

[27] Family income was not included due to the extensive number of missing values and inconsistencies. For example, it was found that siblings that lived in the same house and with the same parents reported quite different amounts. In addition, OLS regressions in which income was included, show that results do not change.

OLS Regressions				
Dependent Variable: College Enrollment				
Variables	Coef. (1)	Std. Err. (2)	Coef. (3)	Std. Err. (4)
Constant	0.651***	0.012	0.557***	0.023
Male	-0.101***	0.016	-0.110***	0.015
Black	-0.123***	0.021	0.012	0.020
Black x Male	-0.074**	0.030	-0.078**	0.028
Hispanic	-0.164***	0.024	-0.027	0.023
Hispanic x Male	0.011	0.033	0.007	0.031
Broken Family	-	-	-0.182***	0.013
Mother Education	-	-	0.155***	0.006
Fam. Mem. under Age 18	-	-	-0.016***	0.005
Mother Uninvolved	-	-	-0.145***	0.021
Mother Permissive	-	-	-0.054***	0.014
Mother Authoritarian	-	-	-0.081***	0.019
R^2	0.034		0.199	
Observations	6643		6643	

Table 5: OLS regression results. *, ** and *** indicate that the coefficient is significant at 10%, 5% and 1% respectively; robust standard errors are reported. There are four (mutually exclusive) patterns of parenting styles: authoritarian, permissive, uninvolved and authoritative (see definition in subsection 2.2); the omitted category is authoritative style. The sample was restricted among those who completed at least eight grade; this implies to drop only 0.5% of the observations. Data: NLSY97.

However, it is still possible that family characteristics may have differential effects based on gender. For example, a broken family may produce more harmful effects on boys than on girls due to the lack of male role models at home. In order to test for this possibility, similar regressions to those presented in Table 5 are performed, but this time interactions between gender and family covariates are included (in order to keep the size of the table tractable, the African American sample is only considered[28]). Table 6 shows that in each of the three specifications black males are statistically less likely to attend college than black females, in addition the size of male's coefficient remains approximately stable across specifications. Column 5 indicates that family covariates do not have differential effects on males given the lack of statistical significance of the interactions. Moreover, a joint test of significance cannot reject the null hypothesis that the OLS coefficients on the interacted variables are equal to zero. Therefore, after considering together the results

[28] Similar results can be found for the other racial groups.

from Tables 5 and 6; then, it is possible to conclude that broad measures of family background covariates cannot explain the observed gender disparities in college enrollment but can explain the racial gap[29]. Moreover, these results are indicating that the channels that affect the gender gap in college enrollment are different from those that explain the racial gap. In this sense, analyzing differences in skills between males and females is suitable.

OLS Regressions (Black Sample)						
Dependent Variable: College Enrollment						
Variable	Coef. (1)	Std. Err. (2)	Coef. (3)	Std. Err. (4)	Coef. (5)	Std. Err. (6)
Constant	0.528***	0.018	0.451***	0.044	0.465***	0.062
Male	-0.175***	0.025	-0.181***	0.024	-0.210**	0.084
Broken Family	-	-	-0.088***	0.029	-0.092**	0.039
Mother Education	-	-	0.179***	0.013	0.193***	0.017
Family Mem. under Age 18	-	-	-0.021**	0.008	-0.023*	0.012
Mother Uninvolved	-	-	-0.066*	0.038	-0.120**	0.051
Mother Permissive	-	-	-0.009	0.027	-0.047	0.038
Mother Authoritarian	-	-	-0.033	0.038	-0.073	0.050
Broken Family x Male	-	-	-	-	0.010	0.057
Mother Education x Male	-	-	-	-	-0.026	0.025
Family Mem. under 18 x Male	-	-	-	-	0.003	0.016
Mother Uninvolved x Male	-	-	-	-	0.123	0.078
Mother Permissive x Male	-	-	-	-	0.072	0.054
Mother Authoritarian x Male	-	-	-	-	0.083	0.078
R^2	0.031		0.176		0.179	
Observations	1732		1732		1732	

Table 6: OLS regression results, black sample. *, ** and *** indicate that the coefficient is significant at 10%, 5% and 1% respectively; robust standard errors are reported. The sample was restricted among those who completed at least eight grade; this implies to drop only 0.5% of the observations. Data: NLSY97

3.2 Cognitive and Non-Cognitive Skills

Substantial gender disparities in behavior and school performance (i.e. proxy but noisy measures for non-cognitive and cognitive skills) can be found at the different schooling levels. For instance,

[29] However, these results do not imply that any type of parental investment in child development has the same effect on boys and girls.

boys show a greater prevalence of behavior problems than girls even at quite early stages of life. Lavigne et al. (1996) finds, based on a large sample study of preschool children, that while 6.6% of preschool females presented some kind of behavioral problem, 10% of males were in that same situation. In a similar vein, evidence from the National Prekindergarten Study[30] indicates that preschool boys are expelled at a rate of 4.5 times higher than girls [Gilliam, (2005)]. Moreover, many studies have shown an overrepresentation of boys suffering attention deficit disorders; depending on the type of setting (i.e. community or clinical) boy/girl ratios go from 3:1 to 5:1 respectively[31]. These gender disparities should not be disregarded, for example Currie and Stabile (2006) find that behavior problems have large negative effects on schooling attainment, regardless of income and maternal education. This problem may be magnified for African American children which are 1.92 times more likely to be labeled as emotionally disturbed[32].

Special education courses are dominated by males; according to the NCES[33], 12% of the students in kindergarten, first, or third grade received this type of education services; but boys are more likely than girls (16% vs. 8%) to participate in them. Consistently with these findings, the proportion of male students who were delayed in kindergarten through grade eight during the year 2007 was 11.7% while for females it was only 7.6% [34].

Evidence related to academic performance at elementary school[35] indicates that fourth grade females almost close the historical gap in math exams. For example, they performed (in large cities) as well as boys in the math National Assessment of Educational Progress (NAEP) tests of the year 2009. However, girls advantage in reading exams is still robust (especially among African Americans[36]). In addition, 38% of boys and 31% of girls in fourth grade could not achieve the basic level in the reading NAEP exam, while an identical proportion of males and females (19%) was below the basic level in math[37].

The percentage of public school male students in kindergarten through 12th grade who were suspended was 9.2% in the year 2000, while for females it was only 3.9% (NCES). These proportions

[30] A comprehensive data collection effort across each of the nation's 52 state-funded prekindergarten programs operating in the 40 states that fund prekindergarten.

[31] Diagnostic guidelines that were released in 2000 estimated the prevalence of ADHD to be between 4% and 12% of school-aged children [Schneider et al., 2006]

[32] Racial Inequity in Special Education (2002)

[33] Timing and Duration of Student Participation in Special Education in the Primary Grades, March 2007, NCES 2007–043

[34] The Condition of Education 2009, Indicator 18, Grade Retention. NCES.

[35] See Cornwell et al. (2012) for an analysis of the gender gap in test score performance in elementary school.

[36] Table B1 in appendix B shows mean scores and standard deviations of math and reading NAEP exams for fourth grade students open by race.

[37] Data extracted form the NCES website on June 2010, http://nces.ed.gov/nationsreportcard/naepdata/dataset.aspx Additional evidence for 8th grade students can be found in appendix B of this paper, table B2.

become much higher for African American men and women with a 17.4% and 9.1% respectively. In the NLSY97, 42% of boys has agreed with the statement *"When I was in school, I used to break rules quite regularly"*, while only 24% of females did so. Higher rates of suspensions for boys are not likely to be irrelevant in terms of the achievement gap given that prior research has confirmed the perception that students who have been suspended from school are at higher risk for other poor school outcomes, including dropping out of school. For example, Segal (2011) finds that those kids who misbehaved in the 8th grade are almost three times more likely to drop out high school and almost three times less likely to graduate from college. The negative correlation between 8th grade misbehavior and educational attainment remains even after controlling for test scores and family characteristics[38].

Gender disparities in behavior and school performance seem to persist in high school; for instance a large study of Minnesota adolescents [Harris, Blum and Resnick (1991)] found that a higher percentage of teenage boys reported frequent antisocial acts (e.g. vandalism) as compared to teenage girls did (10% vs. 6%). In addition, another set of studies in psychology points out that females are less often engaged in problem behaviors and are likely to terminate their involvement in such behaviors sooner than boys [Ensminger (1990), Petersen, Richmond and Leffert (1993), Lerner and Steinberg (2004)]. In terms of academic performance in high school, boys are more likely to drop out[39]; and among those who graduate, females performed better than males with mean GPA of 3.05 and 2.83 respectively in 2000[40]. In addition, girls are more likely than boys to enroll in college preparatory courses[41] and to participate in all types of after school activities except for athletics. For example, in 2001, 19.2% of high school females seniors and 11.8% of males reported participating in academic clubs[42].

Therefore, these preliminary facts from national samples suggest that gender differences in cognitive and non-cognitive/socio-emotional skills (measured as differences in behavior and academic performance during pre, elementary and high school) may explain the observed disparities in educational attainment.

Table 7 shows a set of regression results with the aim to highlight the likely importance of non-cognitive skills in educational attainment. First, it was regressed college enrollment before age 25[43] on gender, race, and gender interacting with race (this estimation will work as the benchmark

[38] See Bertarnd et al. (2011) for an analysis of the gender differences in non-cognitive skills.

[39] Although males comprise 51 percent of the population between 16 to 24 years old, they make up 58 percent of the dropouts in this age group. Source: U.S. Census Bureau, School Enrollment–Social and Economic Characteristics of Students: October 2005.

[40] The High School Transcript Study: A Decade of Change (2001). NCES.

[41] In 2005, 38% and 35% of high school graduates females and males respectively, completed college preparatory basic courses. NCES (2008).

[42] Trends in Educational Equity of Girls & Women. NCES (2004).

[43] OLS regression results are presented because it is straightforward to interpret the coefficients. Logit models

case). Then, proxies for non-cognitive and cognitive skills were included (i.e. suspensions from school[44], retention at school between grade 1 and 8, GPA at grade eight, involvement in fights and precocious sex). Finally, family covariates were added too. In order to alleviate possible problems of endogeneity, these variables were constructed considering early points in life (i.e. until age of 14). Column 3 shows that after including skills proxies, the gender gap is fully explained across all races. In addition, notice that racial disparities between African Americans and whites are also fully explained, but Hispanics are still less likely to enroll in college. However, column 5 indicates that after including family covariates (i.e. mother education), Hispanics are no longer less likely to attend college. To sum up, these regressions suggest that proxies measures for cognitive and non-cognitive skills are relevant factors that can explain the gender gap.

OLS Regressions						
Dependent Variable: College Enrollment						
Variable	Coef. (1)	Std.Err. (2)	Coef. (3)	Std.Err. (4)	Coef. (5)	Std.Err. (6)
Constant	0.651***	0.011	0.616***	0.011	0.447***	0.019
Male	-0.105***	0.016	-0.009	0.014	-0.022	0.014
Black	-0.138***	0.021	-0.021	0.019	0.018	0.019
Black x Male	-0.076***	0.029	-0.018	0.026	-0.016	0.026
Hispanic	-0.176***	0.023	-0.100***	0.021	-0.015	0.022
Hispanic x Male	0.047	0.032	0.048*	0.030	0.027	0.030
Grade Retention	-	-	-0.164***	0.017	-0.133***	0.017
GPA grade 8 std.	-	-	0.195***	0.006	0.174***	0.006
Suspensions	-	-	-0.035***	0.006	-0.030***	0.006
Fights	-	-	-0.029*	0.017	-0.024	0.016
Precocious Sex	-	-	-0.124***	0.016	-0.107***	0.016
R^2	0.036		0.277		0.329	
Family Covariates	No		No		Yes	
Observations	7061		7061		6560	

Table 7: OLS regression results. *, ** and *** indicate that the coefficient is significant at 10%, 5% and 1% respectively; robust standard errors are reported. The sample was restricted among those who completed at least eight grade; this implies to drop only 0.5% of the observations. Family covariates include, for example, mother education, broken family, among others. See Table C1 in appendix C for similar regressions as in tables 5 and 7, with the only difference that the sample size is kept constant across specifications (results do not change). Data: NLSY97.

provide similar results.

[44] This variable measures the number of academic years in which a student was suspended from school.

It may be possible to argue that these results are in fact driven just by the cognitive component of these variables. However, table 8 shows that similar regressions that include ASVAB[45] test scores (i.e. cognitive measures) cannot explain the gender gap.

OLS Regressions						
Dependent Variable: College Enrollment						
Variable	Coef. (1)	Std.Err. (2)	Coef. (3)	Std.Err. (4)	Coef. (5)	Std.Err. (6)
Constant	0.663***	0.012	0.554***	0.011	0.476***	0.021
Male	-0.104***	0.017	-0.070***	0.016	-0.076***	0.015
Black	-0.127***	0.022	0.073***	0.021	0.126***	0.021
Black x Male	-0.070***	0.033	-0.054*	0.029	-0.056**	0.028
Hispanic	-0.168***	0.026	-0.039	0.025	0.013	0.025
Hispanic x Male	0.029	0.037	0.012	0.035	0.015	0.034
General Sciences	-	-	0.025***	0.012	0.018	0.011
Arithmetic Reasoning	-	-	0.026***	0.012	0.024**	0.012
Mathematical Knowledge	-	-	0.117***	0.012	0.103***	0.011
Assembling Objects	-	-	0.020**	0.009	0.011	0.008
Word Knowledge	-	-	0.003	0.012	-0.015	0.011
Paragraph Comprehension	-	-	0.080***	0.009	0.069***	0.012
Family Covariates	No		No		Yes	
R^2	0.033		0.237		0.300	
Observations	5709		5709		5709	

Table 8: OLS regression results. *, ** and *** indicate that the coefficient is significant at 10%, 5% and 1% respectively; robust standard errors are reported. The sample was restricted among those who completed at least eight grade; this implies to drop only 0.5% of the observations. Family covariates include, for example, mother education, broken family, among others. See Table C1 in appendix C for similar regressions as in tables 5, 7 and 8, with the only difference that the sample size is kept constant across specifications (results do not change). Data: NLSY97

4 Factor Model of Educational Attainment

This section develops a sequential model of educational attainment coupled with a measurement system for cognitive and non-cognitive factors that provides several advantages relative to conven-

[45] Armed Forces Vocational Aptitude Battery (ASVAB) subtests: Mathematical Knowledge, Arithmetic Reasoning, Word Knowledge, Paragraph Comprehension, Assembling Objects and General Sciences.

tional OLS regression outcomes. For instance, possible problems of measurement error could be addressed. Moreover, this model will allow to distinguish the effects of non-cognitive skills from the cognitive ones (e.g. covariates such as GPA are most likely a function of cognitive and non-cognitive abilities) and to understand the effects of these skills at each stage of the schooling career. In addition, it will help to recover the disparities in the relative distribution of cognitive and non-cognitive skills between males and females, and to provide further evidence about the importance of skills in the gender gap size. The model follows the spirit of the factor model presented in Cameron and Heckman (2001). However, three main characteristics distinguish this work from theirs. First, the number of factors considered is higher. Second, the inclusion of a measurement system helps to give a specific interpretation to each factor. Third, the factors are allowed (by construction) to be correlated with a subset of the agents characteristics[46].

The key points of this strategy are based on two main ideas. First, incorporate into the analysis the fact that schooling attainment is the consequence of previous educational decisions [Cameron and Heckman (1998)]. Second, the inclusion of two latent factors associated with cognitive and non-cognitive skills is assumed; where these are known by each individual and fixed by the time agents start to make their schooling decisions[47]. In addition, these latent (for the econometrician) skills are assumed to be mutually independent where a measurement system will be used for their identification. The independence assumption may sound strong a priori; however, as it is described below, it provides (in the worst case) a lower bound for the effect of non-cognitive skills. Finally, the identification strategy follows Carneiro, Hansen, and Heckman (2003).

The remaining parts of this section are organized as follows: first, a description of the measurement system for the identification of latent factors (i.e. cognitive and non-cognitive skills) is provided; second, a sequential model of educational attainment is presented; and third, the complete likelihood is shown.

4.1 Measurement System

The following empirical strategy is just focused on the African American and white subsamples; therefore, the intention is to recover the relative distribution of skills among white females, white males, black females and black males; and their effects on educational attainment. Cognitive and non-cognitive skills of white females are assumed to be:

$$\begin{pmatrix} \theta^C \\ \theta^{NC} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right) \tag{1}$$

[46] In Cameron and Heckman (2001) and Heckman, Stixrud and Urzua (2006), among others, the factors need to be independent of all the covariates.

[47] Even though, factors are not allowed to change over time, they are able to fully explain the gender gap.

Factors have no scale associated with them, implying that it is not possible to know whether 1 or 100 is a substantial amount of the factor or not, and given that they are not observed, there is no way to know it. Therefore, normalizations of the means and the variances (without loss of generality) set the location and scale of the factors.

In order to recover the mean and standard deviation of the skills distributions for each subgroup of the population; cognitive (C) and non-cognitive (NC) abilities are expressed relative to white females' skills:

$$\begin{aligned} F_i^C &= \theta_i^C + \alpha_1^C male_i + \alpha_2^C male_i * \theta_i^C + \alpha_3^C black_i + \\ &\quad \alpha_4^C black_i * \theta_i^C + \alpha_5^C black_i * male_i + \alpha_6^C black_i * male_i * \theta_i^C \end{aligned} \tag{2}$$

$$\begin{aligned} F_i^{NC} &= \theta_i^{NC} + \alpha_1^{NC} male_i + \alpha_2^{NC} male_i * \theta_i^{NC} + \alpha_3^{NC} black_i + \\ &\quad \alpha_4^{NC} black_i * \theta_i^{NC} + \alpha_5^{NC} black_i * male_i + \alpha_6^{NC} black_i * male_i * \theta_i^{NC} \end{aligned} \tag{3}$$

Therefore, α_1^u and $(1 + \alpha_2^u)$ (where $u = C$ or NC) provide the mean and standard deviation of white males skills relative to white females, α_3^u and $(1 + \alpha_4^u)$ work similarly for black females, and finally $(\alpha_1^u + \alpha_3^u + \alpha_5^u)$ and $(1 + \alpha_2^u + \alpha_4^u + \alpha_6^u)$ for black males. Notice that the coefficient on θ_i^u is set to be equal one for identification purposes (i.e. sets the scale).

Linear measurement systems are assumed in order to identify the latent factors. In this sense, the Armed Forces Vocational Aptitude Battery (ASVAB) tests will be considered as noisy measures of cognitive skills that will help to identify the factor loadings:

$$\left\{ \begin{array}{c} G_{i1} = \gamma_{11}^C + \gamma_{21}^C F_i^C + \boldsymbol{\gamma}_{31}^C \mathbf{X}_i + \varepsilon_{i1}^G \\ \dots \\ G_{ij} = \gamma_{1j}^C + \gamma_{2j}^C F_i^C + \boldsymbol{\gamma}_{3j}^C \mathbf{X}_i + \varepsilon_{ij}^G \end{array} \right\} \tag{4}$$

where G_{ij} with $j = 1, ..., 6$ represents the result of agent i in the test j (i.e. paragraph comprehension, word knowledge, mathematical knowledge, arithmetic reasoning, general sciences and assembling objects), $F_i^C$the cognitive factor and $\mathbf{X}_i$ a vector of covariates (i.e. family background characteristics and age at the time the tests were taken[48]).

A relatively similar linear system is considered in order to identify the non-cognitive factor:

$$\left\{ \begin{array}{c} H_{i1} = \gamma_{11}^{CNC} + \gamma_{21}^{CNC} F_i^C + \gamma_{31}^{CNC} F_i^{NC} + \boldsymbol{\gamma}_{41}^{CNC} \mathbf{X}_i + \varepsilon_{i1}^H \\ \dots \\ H_{im} = \gamma_{1m}^{CNC} + \gamma_{2m}^{CNC} F_i^C + \gamma_{3m}^{CNC} F_i^{NC} + \boldsymbol{\gamma}_{4m}^{CNC} \mathbf{X}_i + \varepsilon_{im}^H \end{array} \right\} \tag{5}$$

[48] In order to provide robustness checks, controls for the schooling grade at the time of the exam were also included in other specifications and the results did not change.

where H_{im} with $m = 1, ..., 5$ represent GPA at grade 8, school retention between grade 1 to 8, suspensions from school (until age 14), involvement in fights with the intention to hurt (until age 14) and precocious sex[49]. In order to control for possible misreport bias, the measure precocious sex includes a dummy for male. F_i^C and F_i^{NC} denote the cognitive and non-cognitive factors and $\mathbf{X}_i$ represents a vector of family covariates. It is important to notice that there are no exclusive measures for non-cognitive skills. This implies a literal interpretation of the "non-cognitive" term, where F_i^{NC} will capture all the information in the $H's$ that cannot be explained by cognitive abilities[50].

The joint probability of the observed data is assumed to be independent over equations once conditioning on F_i^C, F_i^{NC} and $\mathbf{X}_i$. Therefore, this methodology can be characterized as a particular type of matching where the match variables creating conditional independence are not observed by the econometrician. A last normalization is required in order to completely identify the model; the sign of the factors effect needs to be established; hence, the coefficients γ_{21}^C and γ_{31}^{CNC} were set such that more of the factors is "good"[51]. Finally, the distribution of the errors are assumed normal[52] in the case of continuous variables and logit distributed for the binary ones[53].

4.2 Sequential Model of Educational Attainment

Agents make sequential decisions in order to define their final schooling level based on a set of family covariates and latent endowments[54]. In each period males and females have to decide whether to continue their studies, where their choice set is determined by their previous decisions. More specifically, students initially choose whether to finish grade 10. If the student drops out[55] no further decisions are made (i.e. dropping out constitutes an absorbing state), if the student finishes grade 10, then he/she has to make an additional decision, that is whether to finish grade 11. This

[49] Precocious sex denotes sexual debut before age 15. According to Armour et al. (2007) experiencing sexual debut earlier than one's peers is correlated with higher risks of engaging in delinquency compared to the risks experienced by adolescents debuting at the same time as their peers.

[50] Given that the aim of this paper is to explain the gender gap, the fact that this model is not accounting for constraints is not likely to be a concern (boys and girls on average face the same constraints given that they come from same type of families and attend same type of schools).

[51] An alternative normalization would be to set γ_{21}^C and γ_{31}^{CNC} equal to 1. This normalization would take care of the scale and sign of the effect of the factors. Notice that the normalization implemented in the paper is equally valid as this one.

[52] Mean 0 and variance 1.

[53] Basically, the different configuration of the error terms depending on the type of variable is due to computational issues. Assuming logit distributions speeds up considerably the time of estimation.

[54] As in Cameron and Heckman (2001) wages were not considered.

[55] If a student is not enrolled in school for two consecutive periods, then it is considered as a dropout. This definition avoids considering as a dropout a student who left school one period due to a quite specific reason (e.g. health problem).

process continues until the last year of college education or until they decide to stop their schooling career.

The latent utility of agent i from making educational choice s is defined as follows:

$$V_{is} = \beta_{0s} + \beta_{1s}F_i^C + \beta_{2s}F_i^{NC} + \boldsymbol{\beta}_{3s}\mathbf{Z}_i + \boldsymbol{\beta}_{4s}\mathbf{X}_i + \varepsilon_{is} \quad (6)$$

where $\mathbf{X}_i$, $F_i^C$and F_i^{NC} have the same definition as above, $\varepsilon_{i,s}$ is the error term which is logit distributed and independent of the regressors, and $\mathbf{Z}_i = \{male, black, black{*}male\}$[56]. The intention is to analyze the statistical significance and the sign of the coefficients on $\mathbf{Z}_i$ (at each educational level) once the latent factors are incorporated.

It is crucial to notice that the inclusion of factors controls for the dynamic selection process that occurs during the transitions from one grade to the next. Basically, selection occurs as low ability students leave school in early stages; and hence dropout from the sample. Therefore, it is expected that the distribution of abilities shifts to the right with later grades[57].

The binary outcome variable can be defined as:

$$D_{is} = \left\{ \begin{array}{c} 1 \text{ if } V_{is} \geq 0 \\ 0 \text{ otherwise} \end{array} \right\} \quad (7)$$

Therefore, the probability of finishing school level s can be expressed as a logit model:

$$\begin{aligned} \Pr(D_{i,s} &= 1|F_i^C, F_i^{NC}, \mathbf{Z}_i, \mathbf{X}_{i,} D_{i,s-1} = 1) = \Pr(V_{is} \geq 0 \mid F_i^C, F_i^{NC}, \mathbf{Z}_i, \mathbf{X}_{i,}, D_{i,s-1} = 1) \quad (8) \\ &= \frac{\exp\{\beta_{0s} + \beta_{1s}F_i^C + \beta_{2s}F_i^{NC} + \boldsymbol{\beta}_{3s}\mathbf{Z}_i + \boldsymbol{\beta}_{4s}\mathbf{X}_i\}}{\sum \exp\{\beta_{0s} + \beta_{1s}F_i^C + \beta_{2s}F_i^{NC} + \boldsymbol{\beta}_{3s}\mathbf{Z}_i + \boldsymbol{\beta}_{4s}\mathbf{X}_i\}} \end{aligned}$$

where $D_{i,s-1}$ is the past decision taken by agent i. Finally, the probability of any sequence of life cycle schooling histories can be written as:

$$\prod_{s=1}^{S} [\Pr(D_{i,s} = 1|F_i^C, F_i^{NC}, \mathbf{Z}_i, \mathbf{X}_{i,} D_{i,s-1} = 1)] * \Pr(D_{i,S+1} = 0|F_i^C, F_i^{NC}, \mathbf{Z}_i, \mathbf{X}_{i,} D_{i,S} = 1) \quad (9)$$

This expression implies that the likelihood of achieving educational level s is equal to the probability of reaching grade s times the probability of not continuing to $s+1$.

Notice that any dependence between $D_{i,s}$ and $D_{i,s-1}$ for the same person conditional on $\mathbf{Z}_i$ and $\mathbf{X}_i$, arises from F_i^C and F_i^{NC} (the only exception occurs when $D_{i,s-1} = 0$). Finally, the schooling

[56] $\boldsymbol{\beta}_{3s}$ and $\boldsymbol{\beta}_{4s}$represent vectors of parameters.

[57] Conditional on the initial schooling decision, θ (which is embedded in the $F's$) and the covariates are not independent. In order to understand why; consider the following example: it is expected that youths from very poor families tend to continue schooling only if they have high levels of cognitive and/or non-cognitive skills.

decisions are assumed to be independent from the measurement equations once F_i^C, F_i^{NC}, $\mathbf{X}_i$ and $\mathbf{Z}_i$ are included in the estimation.

4.3 Likelihood

Given that θ_i^C and θ_i^{NC} are not directly observed, then it is necessary to integrate them out. Therefore, the complete likelihood (after considering the independence assumptions) can be written as follows:

$$\prod_{i=1}^{N} \int \int \left[\prod_{s=1}^{S} \Pr(D_{is} = 1 | F_i^C, F_i^{NC}, \mathbf{Z}_i, \mathbf{X}_{i,} D_{i,s-1} = 1) \right] * \left[\prod_{j=1}^{J} \Pr(G_{ij} | \mathbf{X}_{i,} F_i^C) \right] * \tag{10}$$
$$* \left[\prod_{m=1}^{M} \Pr(H_{im} | \mathbf{X}_i, F_i^C, F_i^{NC}) \right] f(\theta^C) f(\theta^{NC}) d\theta^C d\theta^{NC}$$

Following, Aguirregabiria and Mira (2007), the distributions of the latent factors are considered as a discretized version of Normal$(0,1)$ distributions with $T = 21$ points of support, where θ_k^t is the expected value of a standard normal random variable between percentile $100((t-1)/T)$ and percentile $100(t/T)$ (k indicates cognitive or non-cognitive factor). If p_t is denoted as the percentile $100(t/T)$ of a standard normal such that $p_t = \Phi^{-1}(t/T)$, then $\theta_k^t = -(\phi(p_t) - \phi(p_{t-1}))T$. Alternative numerical integration methods such as the composite Simpson's rule provide similar results.

5 Results

The results presented in this section are focused on the African American and white subsamples of the NLSY97[58]. The remaining parts of this subsection are organized as follows: first, results of the educational attainment model are presented; and second, the mean and standard deviation of the estimated factors distributions are compared between genders.

5.1 Educational Attainment

Table 9 shows estimation results of the educational attainment model where agents make sequential decisions from grade 10 to (at most) the last year of college[59]. Three main conclusions can be extracted from this table. First, males are no longer less likely to finish high school or enrolled in postsecondary education after controlling for the latent factors. Indeed, men are shown to have "higher preferences" for educational attainment[60]. This finding is consistent with the empirical

[58] The estimation outcomes do not include family covariates due to their lack of effect on the gender gap size. Table D1 of appendix D shows that the gender gap results are similar when family covariates are included.

[59] Most students are 16 years old in grade 10, which is the age when they start to make their own schooling decisions.

[60] It is important to emphasize that given boys and girls come on average from the same type of families, then family background covariates cannot explain the gender gap (as it was shown in the previous section).

regularity that women still spend more time at home[61], and the fact that the expected benefits for education attainment continue being higher for males than females [see Becker et al. (2010)]. In a similar vein, Hubbard (2011) has shown that the college premium for women is not higher than the premium for men once topcoding biases in the CPS survey are corrected. finally, .

Second, African Americans show "higher preferences" for college. This result should not be surprising given that wage premium for college educated blacks, conditional on ability, is a distinctive characteristic of the US labor market [Arcidiacono et al. (2010)]. In a similar vein, Neal (2006) indicates that college educated blacks and whites have comparable wages at the time of initial entry into the labor market, which implies the presence of a substantial black wage premium given the racial differences in average AFQT scores.

Third, both latent factors are statistically significant different from zero at each stage of schooling career; however, the relative importance of one skill over the other varies across transitions. Cognitive abilities (conditional on reaching certain grade) have higher impact to complete an educational level than non-cognitive ones; especially after finishing high school. For example, Graphs 1 and 2 show that the probability of finishing grade 12 for white females (conditional on being enrolled in it) is more responsive to different values of non-cognitive skills than the probability of completing the fourth year of college (conditional on being enrolled on it). More precisely, Graph 1 indicates that young people with quite low levels of cognitive skills still show high probabilities of finishing high school if their levels of non-cognitive skills are high . On the contrary, Graph 2 shows that the probability of finishing the fourth year of college is considerably smaller (irrespective of non-cognitive levels) if cognitive skills are very low (see for example, coordinates -3 (cog), 5 (non-cog) in each graph). However, as it is shown in the following section, the substantial disparities in the distribution of non-cognitive skills between males and females make these abilities more relevant in terms of the gender gap size.

5.2 Skills Distributions

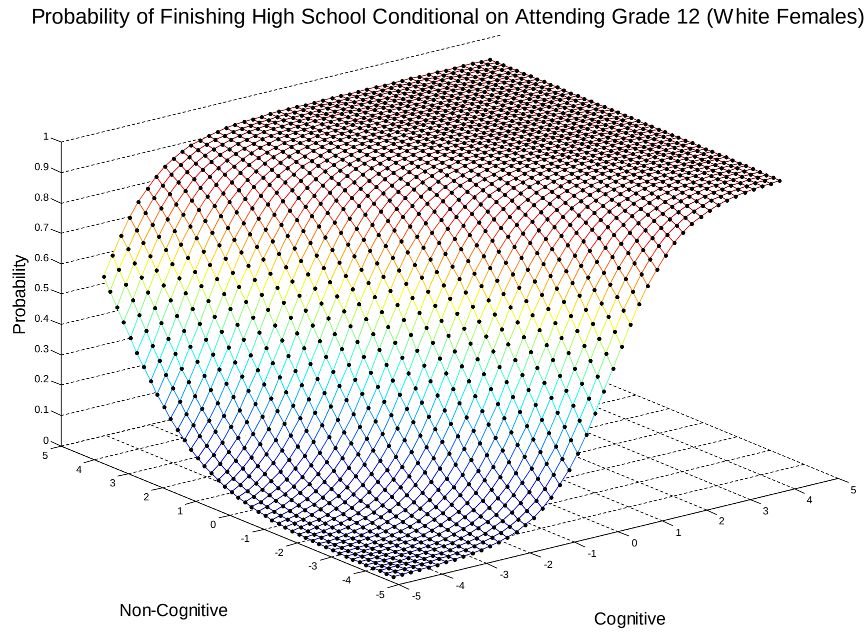
Table 10 presents the means and standard deviations of the estimated distributions of skills for both genders and racial groups[62]. White and black males show lower average skills than their females counterparts. However, gender differences in non-cognitive abilities are substantially higher than cognitive ones. In terms of variances, white males present higher dispersion than white females on both skills; this implies that the (right) intersection of their cognitive distributions occurs at the 93 percentile; while for the non-cognitive distribution it occurs at the 99 percentile.

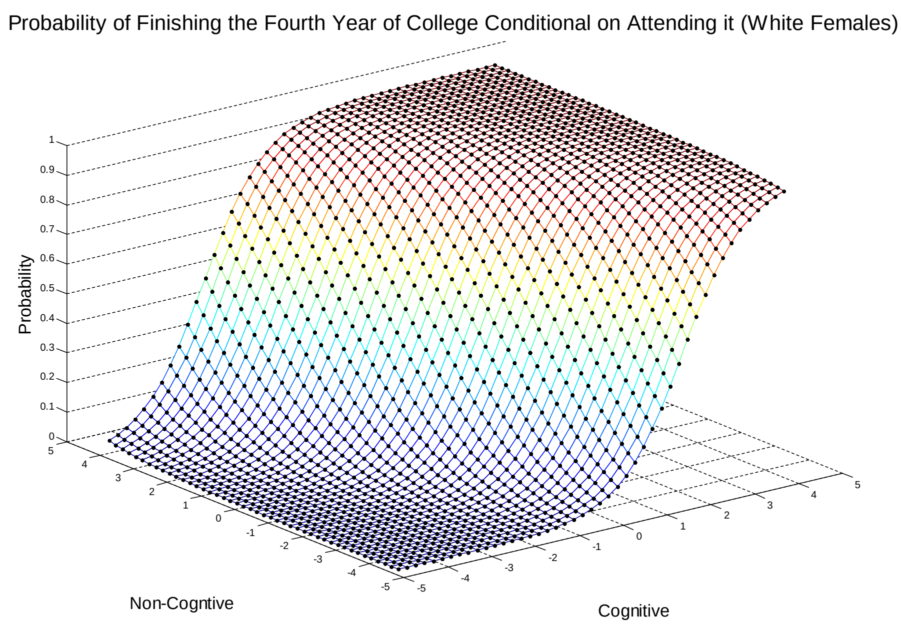
[61] While 96% and 89% of college educated and (just) high school graduates males participate in the labor force, 84% and 74% respectively of females do so. These proportions only include white and black subsamples between 24 and 50 years old during the period 2000 - 2009. Source: IPUMS-CPS.

[62] The values of the parameters α_1^u to α_6^u (where $u = C$ or NC) and their statistical significance can be found in Table D3 of appendix D. In addition, appendix E shows the values of the measurement system coefficients.

	Educational Progression. Full Sample															
	Finish Grade 10		Finish Grade 11		Finish High School		College 1		College 2		College 3		College 4		Finish	
	Coef.	Std.Err.	Coef.	Std.Err.	Coef.	Std.Err.	Coef.	Std.Err.	Coef.	Std.Err.	Coef.	Std.Err.	Coef.	Std.Err.	Coef.	Std.Err.
Constant	3.539***	0.123	3.423***	0.122	2.934***	0.106	0.582***	0.067	1.294***	0.088	0.788***	0.101	0.876***	0.134	0.699***	0.151
Male	0.792***	0.157	0.626***	0.169	0.813***	0.161	0.151	0.105	0.222	0.147	0.232	0.160	0.018	0.182	0.253	0.172
Black	1.337***	0.207	0.899***	0.220	0.907***	0.202	1.723***	0.169	0.852***	0.218	1.202***	0.258	0.909***	0.299	0.237	0.280
Black x Male	-0.106	0.247	-0.170	0.263	0.006	0.254	-0.005	0.205	0.554*	0.290	0.430	0.335	0.583	0.399	0.492	0.411
F^C	1.357***	0.080	1.099***	0.084	1.151***	0.082	1.688***	0.077	1.272***	0.103	1.444***	0.126	1.458***	0.159	0.636***	0.135
F^{NC}	0.829***	0.086	0.760***	0.093	0.773***	0.093	0.761***	0.077	0.667***	0.108	0.908***	0.138	0.731***	0.169	0.475***	0.166

Table 9: White and black sample model estimation results (logit coefficients). *, ** and *** indicate that the coefficient is significant at 10%, 5% and 1% respectively. Data: NLSY97





Graphs 1 and 2: White females probability of finishing grade 12 and fourth year of college (conditional on being enrolled in them) for different levels of cognitive and non-cognitive skills. Data: NLSY97

	Factors: Normal Distributions			
	Cognitive		Non-cognitive	
	Mean	Std. Dev.	Mean	Std. Dev.
White Female	0	1	0	1
White Male	-0.067	1.092	-1.003	1.353
Black Female	-1.203	0.957	-0.256	0.849
Black Male	-1.336	0.909	-1.672	1.125

Table 10: Factors distributions mean and standard deviation of black and white males and females. Data: NLSY97

Similarly, African American males show a higher variance in non-cognitive abilities than their females counterparts. This evidence is consistent with Hedges et al (1995), Arden et al (2006) and Deary (2007) et al. which evidence shows that males present higher variances in cognitive skills than females.

Table 10 also indicates the presence of higher disparities in skills average distributions between black males and females than between whites. For instance, while the difference in non-cognitive skills between whites is around one standard deviation (i.e. 1.003) of white females distribution, for blacks the difference (based on the same scale as whites) is 1.416[63]. In this sense, simulation exercises (as it is shown in the following section) indicate that if disparities in skills between black males and females mirrored (in size) the disparities between white men and women, then the size of the gender gap in college enrollment would be the same for both races.

6 Implications of the Model

The sequential model makes it possible to analyze the full profile of males and females schooling career from multiple perspectives. In order to provide the baseline picture, panel A of Table 11 shows the total sample proportion of girls and boys (open by racial group) that finish the different educational levels[64]. Women constitute the majority (conditional on race) in all the schooling years, and their overrepresentation is increasing every year. A substantial proportion of black males drop out high school, therefore gender disparities in college enrollment are also due to an important proportion of boys not even completing the necessary steps to attend college.

[63] Being -1.416 statistically significant higher than -1.003.

[64] The proportion of students enrolled in college is smaller than the one presented in table 1. This difference is given by the fact that in this section if a student drops out school for two consecutive periods it will be considered to be forever out of the schooling career. Notice, that the size of the gender gap does not change under this new configuration of the data. This assumption is made in order to identify the critical periods. Figure D1 of appendix D shows how well the model fits the data.

It has been shown (in Table 9) that males have "higher preferences" for educational attainment than females after controlling for skills differences. This implies that the gender gap in college enrollment would be bigger if boys had the same preferences as girls. In order to quantify how much bigger it would be, panel B of Table 11 displays the educational attainment of men after imposing women preferences on them. Results indicate that the percentage of black and white men enrolled in college would be only 22% and 39% respectively. Therefore, gender disparities for whites would increase from 10% to 15% and for African Americans from 17% to 22%.

The relative importance of one skill over the other is quite relevant in terms of policy recommendations. For instance, if a policy intends to close gender disparities in educational attainment; then establishing the importance of each skill matters, given that non-cognitive abilities are more malleable than cognitive ones [Cunha et al. (2005)]. Panel C of Table 11 shows the proportion of males that would finish each educational level if factors were increased by one standard deviation (one at a time). The results indicate that such an increase in cognitive skills has a higher impact on the probability of finishing certain grade than a similar one on non-cognitive skills. In addition, cognitive abilities become more relevant for college than non-cognitive ones. For instance, one standard deviation increase in males' non-cognitive abilities would improve the proportion enrolled in college from 0.44 to 0.60 for whites and from 0.27 to 0.41 for blacks. However, a similar increase in cognitive abilities would lead to proportions of 0.70 and 0.51 of white and black males respectively.

These results do not imply that non-cognitive skills are not important for the gender disparities in educational attainment; on the contrary, the fact that boys and girls have higher differences in these skills than in cognitive skills levels turns out to be more relevant in explaining the gender gap. In order to show this, panel D of Table 11 presents the white and black males unconditional probability of finishing each schooling level if it is assumed that they have the female skills distribution (one at a time). For instance, if black males had black females' non-cognitive distribution, 82% would graduate from high school and 45% would enroll in college. However, the percentage of black men finishing grade 12 and attending postsecondary education would only be 65% and 31%, if instead they had the cognitive distribution of black women. To sum up, the observed gap in college enrollment is not only a consequence of a significant number of boys deciding not to enroll in college after finishing high school; a substantial proportion leaves the system before graduating from high school, in part due to low levels of non-cognitive skills.

6.1 Cross-Racial Differences

It was described earlier that the size of the gender disparities in college enrollment are much bigger among blacks than among any other racial group. In this sense, it is suitable to analyze if this empirical regularity can be explained by the fact that African Americans show higher gender differences in average skills (see Table 10) than whites. Three simulation exercises have been

Educational Attainment								
	Grade 10	Grade 11	Grade 12	College 1	College 2	College 3	College 4	Finish
Panel A: Baseline Model								
Males and Females in Each Grade as Proportion of Total Demographic Subsample								
White Female	0.93	0.88	0.82	0.54	0.45	0.37	0.32	0.25
White Male	0.91	0.85	0.78	0.44	0.36	0.28	0.23	0.18
Black Female	0.91	0.83	0.74	0.44	0.33	0.24	0.18	0.13
Black Male	0.85	0.73	0.63	0.27	0.21	0.14	0.10	0.08
Percentage of Males in Each Grade Conditional on Race								
White Male	49.4%	49.0%	48.9%	45%	44.3%	43%	42%	42.2%
Black Male	48.2%	46.6%	45.9%	38.1%	38.4%	36.5%	35.8%	37.5%
Panel B: Males with Female Preferences for Educ. Attainment								
Males in Each Grade as Proportion of Total Demographic Subsample								
White Male	0.85	0.76	0.67	0.39	0.32	0.24	0.20	0.16
Black Male	0.77	0.63	0.49	0.22	0.15	0.09	0.06	0.04
Panel C: One Standard Deviation Increase in Skills								
Males in Each Grade as Proportion of Total Demographic Subsample								
White Male Cognitive	0.97	0.94	0.91	0.70	0.63	0.54	0.48	0.40
White Male Non-cognitive	0.96	0.93	0.89	0.60	0.53	0.45	0.39	0.33
Black Male Cognitive	0.94	0.87	0.80	0.51	0.43	0.34	0.28	0.22
Black Male Non-cognitive	0.92	0.85	0.77	0.41	0.34	0.26	0.21	0.16
Panel D: Males with Female Skills Distributions								
Males in Each Grade as Proportion of Total Demographic Subsample								
White Male Cognitive	0.92	0.86	0.80	0.45	0.37	0.29	0.23	0.18
White Male Non-cognitive	0.96	0.92	0.88	0.56	0.48	0.40	0.34	0.28
Black Male Cognitive	0.86	0.75	0.65	0.31	0.24	0.17	0.13	0.10
Black Male Non-cognitive	0.94	0.88	0.82	0.45	0.38	0.30	0.23	0.18

Table 11: Estimated educational attainment open by gender and race. The percentage of males in each grade conditional on race (see bottom lines of Panel A) refers to the number of males of race r (i.e. black or white) in grade g divided by the total number of race r youths (males plus females) in that grade. Data: NLSY97.

performed with this aim. First, college enrollment was simulated under the assumption that average differences in cognitive skills between black males and females are similar to white gender differences. More specifically, this means turning off the coefficient α_5^C from equation 2. The second simulation repeats this same procedure but this time with non-cognitive skills (i.e. turn off $\alpha_5^{NC}$from equation 3). Finally, the last simulation turns off both coefficients α_5^{NC} and α_5^C. Table 12 shows the differences between the white and black gender gap in each of the described scenarios (i.e. white gap - black gap). Results indicate that more than 70% of the "additional" gap observed among African Americans can be explained by higher gender differences in non-cognitive skills as compared to whites; while less than 30% is explained by higher differences in cognitive skills. Therefore, non-cognitive skills are crucial to explain why African Americans show considerable gender disparities in educational attainment.

College Enrollment Gender Gap Across Races	
	White Gap - Black Gap
Actual Difference	-0.07
Simulation Cognitive Skills ($\alpha_5^C = 0$)	-0.05
Simulation Non-cognitive Skills ($\alpha_5^{NC} = 0$)	-0.02
Simulation Both Skills ($\alpha_5^{NC} = \alpha_5^C = 0$)	0

Table 12: Changes in the difference between the white and black gap if differences in gender skill distributions among blacks mirror the white ones. Data: NLSY97.

6.2 Oaxaca's Decompositions

An alternative approach to quantify the gender gap in educational attainment is to perform a sequence of Oaxaca's decompositions for each level of schooling career. These decompositions will show in more detail how differences in skills levels and preferences contribute to explain the gaps. For example, results in panel B1 of Table 13 indicate that if gender differences in preferences were eliminated, then equalizing the cognitive and non-cognitive levels of African American males and females would close a gap of 0.221. This means that the sign of the gender gap would change (i.e. more males than females would attend college). This table also points out that males have "higher preferences" for educational attainment than females. For instance in the absence of differences in skills between genders, the proportion of white males finishing the first year of college would be 4.7% higher than the proportion of white females.

Panels A2 and B2 of Table 13 indicate that most of the *gap due to skills differences* is mainly explained by disparities in non-cognitive abilities. If preferences for educational attainment were the same across genders, then more than 80% of the *gap due to skills differences* would be explained by differences in non-cognitive abilities (for both races)[65].

6.3 Dynamic Selection Process: Factors Distribution

Finally, Table 14 shows the evolution in cognitive and non-cognitive skills due to the dynamic selection process that occurs at each schooling level (i.e. low skill students left school in early stages and hence dropout from the sample). The intention is to analyze if the selection process is mainly driven by a type of skill. Results indicate a substantial increase in the mean of cognitive skills distribution between the end of high school and the last year of college. In addition, white males show a higher mean of cognitive abilities than females by the end of schooling career (despite they started behind). In terms of non-cognitive skills, black males show important shifts between grade 9 and the end of college. These changes are of much higher magnitude than the ones experienced by white or black females (i.e. more than twice). To sum up, this table indicates that the dynamic selection process is operating intensively in both factors, though the selection process in terms of cognitive skills is more aggressive after high school.

Selection Process: Evolution of Mean Factors				
	Grade 9	Grade 12	College 1	College 4
	Cognitive Skills			
White Female	0	0.182	0.499	0.796
White Male	-0.067	0.170	0.603	0.953
Black Female	-1.203	-0.966	-0.622	-0.213
Black Male	-1.336	-1.039	-0.615	-0.238
	Non-Cognitive Skills			
White Female	0	0.117	0.245	0.424
White Male	-1.003	-0.772	-0.512	-0.187
Black Female	-0.256	-0.131	-0.026	0.168
Black Male	-1.672	-1.375	-1.135	-0.782

Table 14: Cognitive and non-cognitive average skills at selected grades open by gender and race. Data: NLSY97.

To conclude, the previous set of results have shown that once controls for cognitive and non-cognitive skills are included, males are not less likely to attend college than females, being the disparities in the distribution of non-cognitive skills quite relevant to explain the gender gap in educational attainment within and across races.

[65] Appendix F shows Oaxaca decompositions using males covariates as base.

	Oaxaca's Decompositions: Unconditional Probability (Base: Females Covariates)							
	Grade 10	Grade 11	Finish HS	College 1	College 2	College 3	College 4	Finish
	Panel A1: Whites							
Actual Gap	-0.023	-0.035	-0.036	-0.097	-0.093	-0.091	-0.087	-0.067
Gap due to Preferences	0.033	0.05	0.077	0.047	0.047	0.045	0.034	0.037
Gap due to Skills Diff.	-0.056	-0.085	-0.113	-0.144	-0.140	-0.136	-0.121	-0.104
	Panel A2: Contributions to *Gap due to Skills Differences*							
Cognitive	13%	12%	12%	15%	13%	10%	9%	7%
Non-cognitive	87%	88%	88%	86%	87%	90%	91%	93%
	Panel B1: Blacks							
Gender Gap	-0.062	-0.106	-0.113	-0.169	-0.125	-0.104	-0.082	-0.051
Gap due to Preferences	0.039	0.057	0.095	0.052	0.084	0.089	0.083	0.088
Gap due to Skills Diff.	-0.101	-0.163	-0.208	-0.221	-0.209	-0.193	-0.165	-0.139
	Panel B2: Contributions to *Gap due to Skills Differences*							
Cognitive	6%	7%	8%	17%	19%	20%	22%	22%
Non-cognitive	94%	93%	92%	83%	81%	80%	77%	77%

Table 13: Oaxaca's decompositions for unconditional probability of finishing different stages of schooling career. Black and whites separately were considered, using females covariates as base. Data: NLSY97.

7 Conclusions

The sizable gender gap in college enrollment, especially among African Americans, constitutes an empirical regularity that may have serious consequences on marriage markets, males labor force participation and college campuses diversity. Regressions results indicate that family background covariates have no impact on the gender gap. However, the estimation of a sequential model for educational attainment indicates that disparities in cognitive and non-cognitive skills more than explain the gender differences. Indeed, males are shown to have "higher preferences" for schooling than females after controlling for the latent factors.

Cognitive skills exhibit a higher effect (conditional on reaching certain grade) to transition from one schooling level to the next than non-cognitive ones, especially for college enrollment. However, the substantial disparities in the distribution of non-cognitive skills between males and females make these abilities crucial to explain the gender gap size within races. Moreover, the observed gap in college enrollment is not only a consequence of a significant number of boys deciding not to enroll in college after finishing high school; a substantial proportion leaves school at early stages, in part due to low levels of non-cognitive skills.

The puzzling gap size between African American males and females is mainly explained by the substantial gender differences in non-cognitive skills distribution. Simulation exercises show that if black gender disparities in skills mirrored the white ones, then the size of the gap would be the same for both races.

Finally, the fact that the difference in the total population proportion of white and black females attending college (13%) is smaller than the proportion between black females and black males (17%) indicates that it may be necessary to develop public policies that target specific subgroups of the population (e.g. black males).

References

[1] Aguirregabiria, Victor; and Mira, Pedro (2007). "Sequential Estimation of Dynamic Discrete Games." *Econometrica*, 75(1), pp. 1-53.

[2] Almlund Mathilde, Duckworth Angela, Heckman James and Tim Kautz (2011). "Personality Psychology and Economics". IZA DP No. 5500

[3] Arcidiacono, Peter; Bayer, Patrick; and Hizmo, Aurel (2010). "Beyond Signaling and Human Capital: Education and the Revelation of Ability." Forthcoming in *AEJ: Applied Microeconomics.*

[4] Arden, Rosalind and Plomin, Robert (2006). "Sex differences in variance of intelligence across childhood." *Personality and Individual Differences,* 41, 39–48

[5] Armour, Stacy and Dana Haynie. "Sexual Debut and Delinquency." *Journal of Youth and Adolescence*, 36:141-152.

[6] Averett, Susan L., Mark L. Burton (1996). "College Attendance and the College Wage Premium: Differences by Gender" *Economics of Education Review*, 15(1): 37–49.

[7] Baumrind, Diana (1991). Effective parenting during the early adolescent transition. In P.A. Cowan & E. M. Hetherington (Eds.), Advances in family research (Vol. 2). Hillsdale, NJ: Erlbaum.

[8] Becker Gary, Hubbard William, Murphy Kevin (2010). "Explaining The Worldwide Boom in Higher Education of Women" *Journal of Human Capital* 4(3):203-241.

[9] Bertrand, M. (2010). "New Perspectives on Gender" *Handbook of Labor Economics,* Volume 4b, Chapter 17, Elsevier Ltd.

[10] Bertrand, M. and Pan, Jessica (2011). "The Trouble with Boys: Social Influences and the Gender Gap in Disruptive Behavior" NBER Working Paper No. 17541

[11] Brody, L.M., Nagin, D.E., Tremblay, R.E., Bates, J.E., Brame, B., Dodge, K.A. (2003). "Developing trajectories of childhood disruptive behaviors and adolescent delinquency: A six-site, cross-national study." *Developmental Psychology*, Vol. 39(2), Pages 222-245.

[12] Cameron Stephen, Heckman James (2001). "The Dynamics of Educational Attainment for Black, Hispanic, and White Males." *Journal of Political Economy*, vol. 109(3), pages 455-499, June.

[13] Cameron Stephen, Heckman James (1998). "Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males" *Journal of Political Economy*, vol. 106(2), pages 262-333.

[14] Carneiro Pedro, Hansen Karsten, Heckman James (2003). "Estimating Distributions of Treatment Effects with an Application to the Returns to Schooling and Measurement of the Effects of Uncertainty on College Choice." Working Paper Series 2003:9, IFAU - Institute for Labour Market Policy Evaluation.

[15] Caspi Avshalom, Shiner Rebecca (2006). "Personality Development." In: W. Damon, & R. Lerner (Series Eds.) & N. Eisenberg (Vol. Ed.), *Handbook of Child Psychology: Vol.3.* Social, Emotional, and Personality Development (6th ed.). New York: Wiley.

[16] Caspi, A.; Henry, B.; McGee, R.; Moffit, T; Silva, P. (1995). "Temperamental Origins of Child and Adolescent Behavior Problems: From Age Three to Age Fifteen." *Child Development*, Vol(66), Pages: 55-68.

[17] Cavazos-Rehg PA, Krauss MJ, Spitznagel EL, Schootman M, Bucholz KK, Peipert JF, Sanders-Thompson V, Cottler LB, Bierut LJ (2009). "Age of sexual debut among US adolescents." *Contraception* 80, Pages: 158–162

[18] Cawley John, Heckman James, Vytlacil Edward (2001). "Three Observations on Wages and Measured Cognitive Ability." *Labour Economics* 8(4), 419-442.

[19] Cornwell, C., Mustard, D. and Van Parys, J. (2012). "Non-cognitive Skills and Gender Disparities in Test Scores and Teacher Assessments: Evidence from Primary School" Forthcoming *Journal of Human Resources*

[20] Cho, Donghun (2007). "The Role of High School Performance in Explaining Women's Rising College Enrollment." *Economics of Education Review*, Volume 26, Issue 4, Pages 450-462.

[21] Cunha Flavio, Heckman James (2008). "Formulating, Identifying and Estimating the Technology of Cognitive and Non-cognitive Skill Formation." *Journal of Human Resources*, vol. 43(4).

[22] Cunha Flavio, Heckman James, Lance Lochner, Dimitriy V. Masterov (2005). "Interpreting the Evidence on Life Cycle Skill Formation." Discussion Paper Series, IZA DP No. 1675.

[23] Cunha Flavio, Heckman James, Navarro Salvador (2005). "Separating Uncertainty from Heterogeneity in Life Cycle Earnings." *Oxford Economic Papers*, Oxford University Press, vol. 57(2), pages 191-261, April.

[24] Currie Janet, Stabile Mark (2004). "Child Mental Health and Human Capital Accumulation the Case of ADHD." NBER, Working Paper 10435.

[25] Deary, I.; Irwing, P.; Der, G.; Bates, T. (2007). "Brother–sister differences in the g factor in intelligence: Analysis of full, opposite-sex siblings from the NLSY1979". *Intelligence*, vol. 35, no5, pp. 451-456

[26] Digman John (1990). "Personality Structure: Emergence of the Five Factor Model." *Annual Review of Psychology*, 41, 417-440.

[27] Domitrovich, Celene (2008). "Social-Emotional Skills in Early Childhood: Implications for Educational Standards." Penn State Prevention Research Center, Texas Pre-Kindergarten Standards Summit.

[28] Ensminger Margaret (1990). "Sexual Activity and Problem Behaviors Among Black, Urban Adolescents." *Child Development*. Vol 61 pp. 2032–2046

[29] Ferguson Ronald (2005). "Why America's Black-White School Achievement Gap Persists." *Ethnicity, Social Mobility, and Public Policy.* eds. Glenn C. Loury, Tariq Modood and Steven Teles. New York, NY: Cambridge University Press, 2005.

[30] Fortin, M., Oreopoulos, P. and Phipps, S. (2012). "Leaving Boys Behind: Gender Disparities in High Academic Achievement", Working Paper University of British Columbia

[31] Freeman, Catherine (2004). "Trends in Educational Equity of Girls & Women: 2004." U.S. Department of Education, Institute of Education Sciences, (NCES 2005-016).

[32] Frierson, Herny; Pearson Willie; and Wyche, James. (2009). Black American Males in Higher Education: Diminishing Proportions. Diversity in Higher Education. Volume 6. Emerald Group Publishing Limited. First Edition.

[33] Fryer, Roland (2010). "Racial Inequality in the 21st Century: The Declining Significance of Discrimination" NBER Working Paper No. 16256.

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

[35] Gilliam Walter (2005). "Prekindergarteners Left Behind: Expulsion Rates in State Prekindergarten Programs." Foundation for Child Development, http://www.fcd-us.org/PDFs/NationalPreKExpulsionPaper03.02_new.pdf

[36] Gullone Eleonora, Moore Susan (2000). "Adolescent Risk Taking and the Five Factor Model of Personality." *Journal of Adolescence*, 23(4), 393-407.

[37] Harris Linda, Blum Robert, Resnick Michael (1991). "Teen Females in Minnesota: A portrait of Quiet Disturbance." *Women and Therapy*, 11 (3-4), 119-135.

[38] Hauser, Robert (1999). "Should We End Social Promotion? Truth and Consequences." Center for Demography and Ecology, University of Wisconsin-Madison, Working Paper No. 99-06

[39] Hausmann Ricardo, Tyson Laura, Zhidi Saadia (2009). "The Global Gender Gap Report." World Economic Forum.

[40] Heckman James, Rubinstein Yona (2001). "The Importance of Non-cognitive Skills: Lessons from the GED Testing Program." *American Economic Review*, 91, 145-149

[41] Heckman James, Lochner Lance, Todd Petra (2006). "Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond." *Handbook of the Economics of Education*, Elsevier.

[42] Heckman James, Stixrud Jora, Urzua Sergio (2006). "The Effects of Cognitive and Non-cognitive Abilities on Labor Market Outcomes and Social Behavior." *Journal of Labor Economics*, vol. 24(3), pages 411-482, July.

[43] Heckman James, Rodrigo Pinto, Hongbo Wang (2008). "Determinants of Risk Behaviors in Early Adulthood: Interplay of Cognitive Ability, Non-cognitive Skills, and Health." Working Paper, University of Chicago.

[44] Heckman James, Urzua Sergio, Veramendi Gregory (2010). "The Effects of Schooling on Labor Market and Health Outcomes." Working Paper, University of Chicago.

[45] Hedges, LV; Nowell, A (1995). "Sex differences in mental test scores, variability, and numbers of high-scoring individuals". *Science*, 269, 41–5

[46] Hoff Sommers, Christina (2000). The War Against Boys: How Misguided Feminism is Harming Our Young Boys. Simon and Shuster

[47] Holt Emily, McGrath Daniel, Herring Willian (2007). "Timing and Duration of Student Participation in Special Education in the Primary Grades." U.S. Department of Education, (NCES 2007–043).

[48] Hubbard, William (2011). "The Phantom Gender Difference in the College Wage Premium." *Journal of Human Resources* 46(3): 568–586

[49] Jacob, Brian (2002). "Where the Boys Aren't: Non-cognitive Skills, Returns to School and the Gender Gap in Higher Education." *Economics of Education Review*, December, 21:6, pp. 589-98.

[50] Jencks Christopher, Philips Meredith (1998). *The Black-Whites Test Score Gap.* Washington, D.C. Brookings Institution Press, 1998.

[51] Judge Timothy, Higgins Chad, Thoresen, Carl, Barrick Murray (1999). "The Big Five Personality Traits, General Mental Ability, and Career Success Across the Life Span." *Personnel Psychology*, 52, 621–652.

[52] Kerwin Charles , Luoh Ming-Ching (2003). "Gender Differences in Completed Schooling." *The Review of Economics and Statistics*, Vol. 85, No. 3, pp. 559-577

[53] Kewal Ramani, Gilbertson, Fox, Provasnik (2007). "Status and Trends in the Education of Racial an Ethnic Minorities." U.S. Department of Education, Institute of Education Sciences, (NCES 2007-039).

[54] King, Jacqueline (2010). "Gender Equity in Higher Education: 2010." American Council on Education Center for Policy Analysis.

[55] Judith, Kleinfeld (2006). "Five Powerful Strategies for Connecting Boys to Schools." Paper for White House Conference on Helping America's Youth Indianapolis, Indiana

[56] Ladd, G.W., & Burgess, K.B. (1999). "Charting the relationship trajectories of aggressive, withdrawn, and aggressive-withdrawn children during early grade school." *Child Development*, Vol. 70(45), Pages 910-929.

[57] Lavigne,J., Gibbons, R., Christofel, K; Arend, R. et al. (1996). "Prevalence Rates and Correlates of Psychiatric Disorders among Preschool Children." *Journal of the American Academy of Child and Adolescent Psychiatry*, Vol. 35(2), 204-214.

[58] Lerner Richard, Steinberg Laurence (2004). *Handbook of Adolescent Psychology,* (Second Edition). Hoboken, N.J.: John Wiley & Sons.

[59] Losen, Dan and Orfield, Gary (2002). *Racial Inequity in Special Education.* Harvard Education Publishing Group.

[60] McDaniel, A., DiPrete, T., Buchmann. C., and Shwed, U. (2011). "The Black Gender Gap in Educational Attainment: Historical Trends and Racial Comparisons" *Demography*, 48:889–914

[61] Nagin, D.S., & Tremblay, R.E. (2001). "Parental and early childhood predictors of persistent physical aggression in boys from kindergarten to high school." *Archives of General Psychiatry*, 58(4), 389-394.

[62] Neal Derek, Johnson William (1996). "The Role of Premarket Factors in Black-White Wage Differences." *Journal of Political Economy*, 104(5), pp. 869-95

[63] Neal Derek (2004). "The Measured Black-White Wage Gap among Women Is Too Small." *Journal of Political Economy*, 2004, 112(1), pp. S1-28.

[64] Neal, Derek. 2006. "Why Has Black-White Skill Convergence Stopped?". *Handbook of Economics of Education I*, eds. Eric Hanushek and Finis Welch, North Holland.

[65] NCES, Digest of Education Statistics (2005). U.S. Department of Education, National Center for Education Statistics.

[66] NCES, Integrated Postsecondary Education Data System (2003). U.S. Department of Education, National Center for Education Statistics.

[67] Planty, M., Hussar, W., Snyder, T., Kena, G., KewalRamani, A., Kemp, J., Bianco, K., Dinkes, R. (2009). "The Condition of Education 2009" (NCES 2009-081). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

[68] Owings, W. and Magliaro, S. (1998). "Grade Retention: A History of Failure." ERIC Clearing House. Washington, DC: U.S. Department of Education.

[69] Payton, J.; Wardlaw, D.; Graczyk, P., Blooddworth, M.; Tompsett, C.; Weissberg, R. (2000). "Social and Emotional Learning: a framework for mental health and reducing risk behavior in children and youth." *Journal of School Health*, Vol. 70, Pages 179-185.

[70] Peter Katharin, Horn Laura. (2005). "Gender Differences in Participation and Completion of Undergraduate Education and How They Have Changed Over Time." U.S. Department of Education, (NCES 2005–169).

[71] Petersen Anne, Richmond Julius, Leffert Nancy (1993). "Social Changes Among Youth: The United States Experience." *Journal of Adolescent Health*, Volume 14, Issue 8, Pages 632-637.

[72] Perkins Robert, Brian Kleiner, Stephen Roey (2004). "The High School Transcript Study: A Decade of Change in Curricula and Achievement, 1990-2000." U.S. Department of Education, Institute of Education Sciences, NCES 2004-455.

[73] Planty, Hussar, Snyder, Kena, Kewal Ramani, Kemp, Bianco, Dinkes (2009). "The Condition of Education 2009." U.S. Department of Education, Institute of Education Sciences, (NCES 2009-081).

[74] Rampey, Dion, Donahue (2009). "NAEP 2008 Trends in Academic Progress." U.S. Department of Education, Institute of Education Sciences, (NCES 2009–479).

[75] Reynolds, John, and Jennifer Pemberton (2001). "Rising College Expectations among Youth in the United States: A Comparison of the 1979 and 1997 NLSY." *The Journal of Human Resources*, Vol. 36, No. 4, Special Issue on Early Results from the National Longitudinal Survey of Youth, 1997 Cohort, pp. 703-726.

[76] Sandberg Seija (1996). Hyperactivity and Attention Disorders of Childhood. Cambridge University Press.

[77] Segal Carmit (2011). "Misbehavior, Education, and Labor Market Outcomes." *The Journal of the European Economic Association*, Forthcoming..

[78] Schneider Helen, Eisenberg Daniel (2006). "Who Receives a Diagnosis of Attention-Deficit/Hyperactivity Disorder in the United States Elementary School Population?" *Pediatrics*, Vol 117(4), Pages 601-609.

[79] Shaw, D., Owens, E., Giovannelli, J. & Winslow, E. (2001). "Infant and toddler pathways leading to early externalizing disorders." *Journal of the American Academy of Child & Adolescent Psychiatry*, Vol. 40, 36-43.

[80] Singh Ilina (2008). "Beyond Polemics: Science and Ethics of ADHD." *Nature Reviews, Neuroscience*, Vol 9, Pages 957–964.

[81] Wansbeek, Tom and Erik Meijer. "Measurement Error and Latent Variables." A Companion to Theoretical Econometrics. Baltagi, Badi H. Blackwell Publishing, 2003. Blackwell Reference Online. 29 July 2010 <http://www.blackwellreference.com/subscriber/tocnode?id=g9781405106764_chunk_g978140510676413>

[82] White, J.L., Moffitt, T.E., Earls, F, Robins, L., & Silva, P.A. (1990). "How early can we tell? Predictors of childhood conduct disorder and adolescent delinquency." *Criminology*, Vol. 28, Pages 507-533.

[83] Whitemire, Richard (2010). Why Boys Fail: Saving Our Sons from and Educational System That's Leaving Them Behind. Amacon, Broadway, New York.

8 Appendix A

Gender Gap in Undergraduate Fall Enrollment				
	Race			
Year	White	Black	Hispanic	Asian
2000	-10.7%	-25.5%	-13.8%	-4.9%
2001	-10.8%	-26.2%	-14.4%	-5.6%
2002	-11.2%	-27.2%	-15.3%	-6.1%
2003	-11.8%	-28.1%	-16.9%	-7.3%
2004	-11.8%	-28.6%	-17.1%	-7.5%

Table A1: Difference in the percentage of undergraduate fall enrollment in degree-granting institutions between males and females (conditional on race). For instance, the percentages of white males and white females enrolled in degree granting institutions conditional on total white enrollment in 2004 were 44.1% and 55.9% respectively; then, the difference is -11.8%. The percentages in Tables 2, 3 and 4 follow the same interpretation. Source: NCES

Gender Gap in Associate's degrees. Year 2002 - 2003				
	White	Black	Hispanic	Asian
Gap	-18%	-32%	-21%	-14%
Selected Majors				
Business	-26%	-28%	-26%	-44%
Engineering	25%	10%	22%	37%
Health Professions	-65%	-35%	-30%	-43%
Liberal Arts and Sciences	-47%	-30%	-60%	-55%

Table A2: Difference in the percentage of Associate's degrees and majors obtained in the academic year 2002-2003 between males and females conditional on race. Source: NCES

Gender Gap in Bachelor's degrees. Year 2002 - 2003				
	White	Black	Hispanic	Asian
Gap	-13%	-33%	-21%	-9%
Selected Majors				
Business	7%	-22%	-10%	-31%
Computer and Information Sci.	15%	1%	7%	44%
Health Professions	-29%	-14%	-13%	-29%
Liberal Arts and Sciences	-7%	-5%	-13%	-9%
Psychology	-24%	-12%	-19%	-27%

Table A3: Difference in the percentage of Bachelor's degrees and majors obtained in the academic year 2002-2003 between males and females conditional on race. Source: NCES

9 Appendix B

NAEP Reading and Math Average Scores				
	White		Black	
	Math			
	Male	Female	Male	Female
National	249	247	222	223
	(0.3)	(0.2)	(0.4)	(0.5)
Large Cities	250	250	219	220
	(1.4)	(0.9)	(0.8)	(0.7)
	Reading			
National	227	233	200	209
	(0.3)	(0.3)	(0.6)	(0.6)
Large Cities	230	236	198	205
	(1.7)	(1.4)	(1.1)	(1.0)

Table B1: National Assessment of Educational Progress, average scores results of fourth grades students in reading and math at national level and large cities. Source: U.S. Department of Education

NAEP Reading and Math				
Proportion Performing Below Basic Level				
	National Level			
	Math		Reading	
	Male	Female	Male	Female
Grade 4	19%	19%	38%	31%
Grade 8	28%	29%	30%	22%

Table B2: Proportion of students in fourth and eight grade performing below the basic level in the National Assessment of Educational Progress tests at the national level. Source: The Nation Report Card. Reading 2009 and Mathematics 2009, National Assessment of Educational Progress at grades four and eight.

10 Appendix C

OLS Regressions (Constant Sample)								
Dependent Variable: College Enrollment								
Variable	Coef. (1)	Std.Err. (2)	Coef. (3)	Std.Err. (4)	Coef. (5)	Std.Err. (6)	Coef. (7)	Std.Err. (8)
Constant	0.689***	0.013	0.591***	0.025	0.639***	0.013	0.575***	0.012
Male	-0.107***	0.018	-0.115***	0.016	-0.013	0.016	-0.076***	0.016
Black	-0.132***	0.024	0.009***	0.023	-0.008	0.021	0.065***	0.022
Black x Male	-0.087**	0.034	-0.092***	0.032	-0.022	0.031	-0.066**	0.031
Hispanic	-0.184***	0.027	-0.048*	0.027	-0.112***	0.025	-0.054**	0.027
Hispanic x Male	0.031	0.039	0.018	0.037	0.054	0.036	0.018	0.037
Family Covariates	No		Yes		No		No	
Non-Cognitive Proxies	No		No		Yes		No	
Cognitive Proxies	No		No		No		Yes	
R^2	0.038		0.199		0.279		0.240	
Observations	5109		5109		5109		5109	

Table C1: OLS regressions similar to those presented in Tables 5, 7 and 8 with the only difference that the sample is kept constant. Tables 5, 7, and 8 intend to maximize the size of the sample, while these tables show that the results are similar when the sample is kept constant. Family covariates, non-cognitive proxies and cognitive proxies are the same to those included in tables 5, 7 and 8 respectively. Data: NLSY97

11 Appendix D

Table D1 presents the estimation results of the educational attainment model (as in table 9) but with the difference that the following family background covariates were included in the estimation: mother education, number of household members with age less than 18, and an indicator for broken family. As it was mentioned earlier, the relevant results do not change: the gender gap in college enrollment is no longer present after the inclusion of the factors and both skills are statistically significant in each stage of the schooling career.

Educational Progression. Full Sample (With Family Covariates)																
	Finish Grade 10		Finish Grade 11		Finish High School		College 1		College 2		College 3		College 4		Finish	
	Coef.	Std.Err.	Coef.	Std.Err.	Coef.	Std.Err.	Coef.	Std.Err.	Coef.	Std.Err.	Coef.	Std.Err.	Coef.	Std.Err.	Coef.	Std.Err.
Constant	3.429***	0.158	3.197***	0.147	2.702***	0.135	-0.010	0.109	0.840***	0.129	0.227	0.157	0.232	0.209	0.362*	0.209
Male	0.940***	0.176	0.817***	0.187	0.961***	0.179	0.202*	0.118	0.329***	0.163	0.269	0.176	0.255	0.218	0.425***	0.197
Black	1.326***	0.230	1.090***	0.241	1.147***	0.228	1.855***	0.185	0.938***	0.235	1.347***	0.279	1.193***	0.347	0.333	0.303
Black x Male	-0.009	0.276	-0.277***	0.289	0.031	0.287	0.173	0.225	0.670*	0.311	0.594	0.361	0.784*	0.451	0.679	0.422
F^C	1.182***	0.079	0.994***	0.079	0.981***	0.076	1.489***	0.072	1.104***	0.093	1.300***	0.115	1.349***	0.157	0.590***	0.123
F^{NC}	0.928***	0.096	0.807***	0.100	0.856***	0.102	0.820***	0.084	0.767***	0.116	0.977***	0.147	1.061***	0.205	0.679***	0.185

Table D1: White and black sample model estimation results (logit coefficients). *, ** and *** indicate that the coefficient is significant at 10%, 5% and 1% respectively. Family covariates were included in the estimation. Data: NLSY97

Table D2 shows the mean and standard deviation of factors distributions after including family covariates in the estimation. The results indicate that differences between blacks and whites on cognitive and non-cognitive skills are reduced in relation to the results in table 10; however, racial differences persist, with the exception of non-cognitive skills for black females.

	Factors: Normal Distributions			
	Cognitive		Non-cognitive	
	Mean	Std. Dev.	Mean	Std. Dev.
White Female	0	1	0	1
White Male	-0.086	1.151	-1.102	1.329
Black Female	-1.035	0.959	-0.091	0.860
Black Male	-1.228	0.969	-1.613	1.071

Table D2: Factor means and standard deviations of black and white males and females, after including family background covariates. Data: NLSY97.

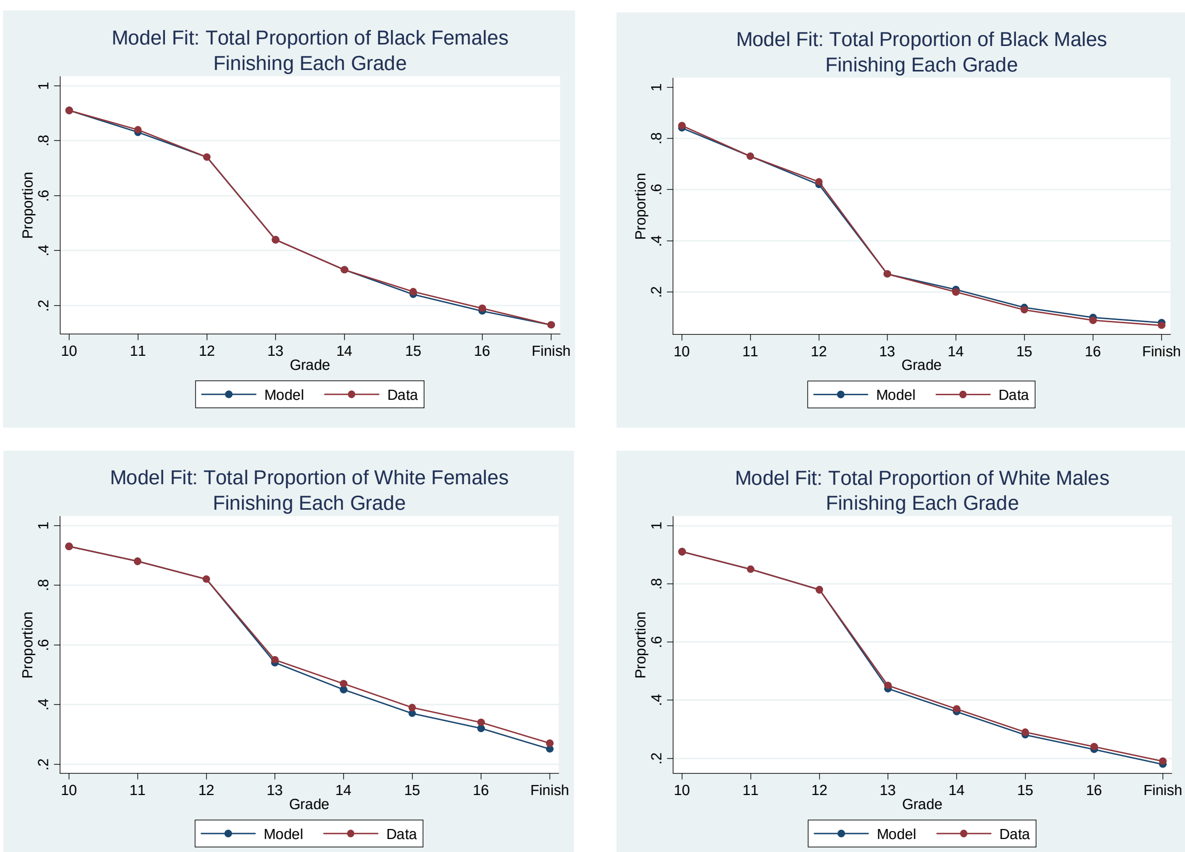


Figure D1: Educational attainment open by race and gender: data and model fit.

Parameters of Factors Distributions (α_s^u)				
	Cognitive		Non-cognitive	
Variable	Coef. (1)	Std.Err. (2)	Coef. (3)	Std.Err. (4)
$male_i$	-0.067*	0.035	-1.003***	0.074
$male_i * \theta_i^u$	0.092***	0.027	0.355***	0.076
$black_i$	-1.203***	0.060	-0.256***	0.097
$black_i * \theta_i^u$	-0.043	0.040	-0.151	0.128
$black_i * male_i$	-0.067	0.078	-0.413***	0.133
$black_i * male_i * \theta_i^u$	-0.140***	0.058	-0.079	0.075

Table D3: Coefficients and standard errors of the factors parameters (i.e α_1^u to α_6^u). Data: NLSY97

12 Appendix E

	Measurement System Parameters (Cognitive Variables)					
	Arithmetic Reasoning	Assembling Objects	Word Knowledge	Paragraph Comprehension	Math Knowledge	General Sciences
	Coef.	Coef.	Coef.	Coef.	Coef.	Coef.
Constant	-2.022***	-1.777***	-3.000***	-2.357***	-3.547***	-2.540***
F^C	0.641***	0.567***	0.596***	0.637***	0.625***	0.609***
Age_Asvab	0.156***	0.140***	0.221***	0.177***	0.256***	0.192***

Table E1: Coefficients and standard errors of the cognitive measurement system. "Age_Asvab" denotes the age at the time the exam was taken. Data: NLSY97

	Measurement System Parameters (Cognitive and Non-Cognitive Variables)				
	Fights	Grade Retention	Precocious Sex	GPA Grade Eight	Suspensions
	Coef.	Coef.	Coef.	Coef.	Coef.
Constant	-2.534***	-3.144***	-2.255***	6.189***	0.238***
F^C	-0.354***	-1.208***	-0.699***	0.996***	-0.267***
F^{NC}	-0.804***	-0.323***	-1.042***	0.541***	-0.323***
Male	-	-	-0.933***	-	-
Black	-	-	-	0.866***	-

Table E2: Coefficients and standard errors of the cognitive/non-cognitive measurement system. Binary variables such as grade retention, fights and precocious sex present logit coefficients; therefore, they cannot be interpreted directly. Given the substantial differences between males and females in reporting sexual behavior, a dummy for male was included to control for misreport bias. Similarly, due to the fact that whites and blacks attend on average different types of schools a dummy for race was included in the measure for GPA at grade eight. Data: NLSY97

13 Appendix F

	Oaxaca's Decompositions: Unconditional Probability (Base: Males Covariates)							
	Grade 10	Grade 11	Finish HS	College 1	College 2	College 3	College 4	Finish
	Panel A1: Whites							
Actual Gap	0.023	0.035	0.036	0.097	0.093	0.091	0.087	0.067
Gap due to Preferences	-0.061	-0.082	-0.111	-0.049	-0.044	-0.036	-0.026	-0.027
Gap due to Skills Diff.	0.084	0.117	0.147	0.146	0.137	0.127	0.113	0.094
	Panel A2: Contributions to *Gap due to Skills Differences*							
Cognitive	18%	15%	13%	9%	7%	4%	3%	1%
Non-cognitive	82%	86%	88%	91%	93%	96%	97%	99%
	Panel B1: Blacks							
Gender Gap	0.062	0.106	0.113	0.169	0.125	0.104	0.082	0.051
Gap due to Preferences	-0.079	-0.097	-0.134	-0.047	-0.06	-0.051	-0.042	-0.039
Gap due to Skills Diff.	0.141	0.203	0.247	0.216	0.185	0.155	0.124	0.090
	Panel B2: Contributions to *Gap due to Skills Differences*							
Cognitive	14%	13%	13%	15%	15%	14%	14%	13%
Non-cognitive	86%	87%	87%	84%	85%	86%	86%	87%

Table F1: Oaxaca's decompositions for unconditional probability of finishing different stages of schooling career. Black and whites separately were considered, using males covariates as base. Data: NLSY97.

CENTRE FOR ECONOMIC PERFORMANCE
Recent Discussion Papers

1219	Peter Arcidiacono Esteban Aucejo Andrew Hussey Kenneth Spenner	Racial Segregation Patterns in Selective Universities
1218	Silvana Tenreyro Gregory Thwaites	Pushing On a String: US Monetary Policy is Less Powerful in Recessions
1217	Gianluca Benigno Luca Fornaro	The Financial Resource Curse
1216	Daron Acemoglu Ufuk Akcigit Nicholas Bloom William R. Kerr	Innovation, Reallocation and Growth
1215	Michael J. Boehm	Has Job Polarization Squeezed the Middle Class? Evidence from the Allocation of Talents
1214	Nattavudh Powdthavee Warn N. Lekfuangfu Mark Wooden	The Marginal Income Effect of Education on Happiness: Estimating the Direct and Indirect Effects of Compulsory Schooling on Well-Being in Australia
1213	Richard Layard	Mental Health: The New Frontier for Labour Economics
1212	Francesco Caselli Massimo Morelli Dominic Rohner	The Geography of Inter-State Resource Wars
1211	Stephen Hansen Michael McMahon	Estimating Bayesian Decision Problems with Heterogeneous Priors
1210	Christopher A. Pissarides	Unemployment in the Great Recession
1209	Kevin D. Sheedy	Debt and Incomplete Financial Markets: A Case for Nominal GDP Targeting
1208	Jordi Blanes i Vidal Marc Möller	Decision-Making and Implementation in Teams
1207	Michael J. Boehm	Concentration versus Re-Matching? Evidence About the Locational Effects of Commuting Costs
1206	Antonella Nocco Gianmarco I. P. Ottaviano Matteo Salto	Monopolistic Competition and Optimum Product Selection: Why and How Heterogeneity Matters

1205	Alberto Galasso Mark Schankerman	Patents and Cumulative Innovation: Causal Evidence from the Courts
1204	L Rachel Ngai Barbara Petrongolo	Gender Gaps and the Rise of the Service Economy
1203	Luis Garicano Luis Rayo	Relational Knowledge Transfers
1202	Abel Brodeur	Smoking, Income and Subjective Well-Being: Evidence from Smoking Bans
1201	Peter Boone Ila Fazzio Kameshwari Jandhyala Chitra Jayanty Gangadhar Jayanty Simon Johnson Vimala Ramachandrin Filipa Silva Zhaoguo Zhan	The Surprisingly Dire Situation of Children's Education in Rural West Africa: Results from the CREO Study in Guinea-Bissau
1200	Marc J. Melitz Stephen J. Redding	Firm Heterogeneity and Aggregate Welfare
1199	Giuseppe Berlingieri	Outsourcing and the Rise in Services
1198	Sushil Wadhwani	The Great Stagnation: What Can Policymakers Do?
1197	Antoine Dechezleprêtre	Fast-Tracking 'Green' Patent Applications: An Empirical Analysis
1196	Abel Brodeur Sarah Flèche	Where the Streets Have a Name: Income Comparisons in the US
1195	Nicholas Bloom Max Floetotto Nir Jaimovich Itay Saporta-Eksten Stephen Terry	Really Uncertain Business Cycles
1194	Nicholas Bloom James Liang John Roberts Zhichun Jenny Ying	Does Working from Home Work? Evidence from a Chinese Experiment

The Centre for Economic Performance Publications Unit
Tel 020 7955 7673 Fax 020 7404 0612
Email info@cep.lse.ac.uk Web site http://cep.lse.ac.uk