Catching Up to Girls: Understanding the Gender Imbalance in Educational Attainment Within Race[*]

Esteban M. Aucejo[†]
Arizona State University

Jonathan James[‡]
California Polytechnic State University

March 8, 2017

Abstract

Black females are 17 percentage points more likely to attend college than black males, making the gender gap among black youth larger than the black-white racial gap in college enrollment (14.7 pp). We estimate a sequential model of schooling and arrests to assess the major contributing factors to the gender imbalance in educational attainment within racial groups. First, we find that differences between males and females in measures of early behavior account for the majority of the gender gap for each racial group. Second, despite the fact that 50% of black males were arrested at least once before age 25, we find little evidence that arrest outcomes influence educational attainment, and that the negative correlation of educational attainment and arrests is entirely attributable to the same behavioral factors that explain the gender gap in education. Finally, we find that black males have the largest response to improvements in family background characteristics, such that equalizing the distribution of family background characteristics for black and white youths reduces the gender gap in college enrollment among black youth by 50% and completely eliminates the black-white racial gap in college enrollment.

Keywords: Gender Gap, Educational Attainment, Behavior, Factors, Race

JEL Classification Codes: I2, J15, J16

[*]We thank Domenico Ferraro, Tyler Ransom, Robert Topel, and Gregory Veramendi for valuable comments. We also thank seminar participants at Arizona State University, and California Polytechnic State University. All errors are our own.

[†]Department of Economics, W.P. Carey School of Business, Arizona State University & CEP. Email: Esteban.Aucejo@asu.edu

[‡]Department of Economics, California Polytechnic State University. Email: jjames04@calpoly.edu. Homepage: `http://www.calpoly.edu/~jjames04`

1 Introduction

The disparities in educational outcomes between black and white Americans have received substantial attention in social science research for decades. While the racial gap in educational attainment is large, gender disparities in educational attainment within race are even starker, yet have received far less attention. For example, overall, white youths are 35% more likely to enroll in college than black youths. However, among black youths, females are 50% more likely to enroll in college than their male counterparts, making the gender gap within black youth significantly larger than the black-white racial gap.[1] These gender imbalances have considerable implications from a public policy standpoint, affecting a range of issues from male employment and labor market outcomes (Acemoglu and Autor, 2011), to family formation and the economic mobility of future generations (McDaniel et al., 2011; Autor and Wasserman, 2013; Lundberg and Pollak, 2014), as well as diversity on college campuses.[2]

Research studying the aggregate gender gap in educational attainment (Goldin et al., 2006; Jacob, 2002; Buchmann and DiPrete, 2006; DiPrete and Jennings, 2012; Fortin et al., 2015; Chetty et al., 2016) has highlighted two important mechanisms: first, the role of underlying group differences in characteristics between males and females, and second, differential group responses to characteristics. With respect to these two mechanisms, Jacob (2002) finds that the aggregate gender gap in college enrollment is largely explained by differences in non-cognitive skills and that differences in cognitive skills and family background characteristics as well as differential responses to these skills have little bearing on the gender gap in college enrollment. Recently, Autor et al.

[1] In the National Longitudinal Survey of Youth 1997, among black youth, females are 17 percentage points more likely to attend college than males, while the black-white racial gap in college enrollment is 14.7 percentage points. Similar gaps can be found in the American Community Survey (ACS). In the 2009 ACS, the racial gap in college enrollment was 14.1 percentage points and the gender gap for black youth was 16.5 percentage points.

[2] Acemoglu and Autor (2011) show that earnings and employment prospects of less educated workers have declined sharply since the early 1980's such that recent cohorts of less-educated males, and in particular black males, have fewer opportunities in the labor market than their predecessors. Autor and Wasserman (2013) document that less-educated males have lower marriage rates but father children at rates equal to more-educated males. Consequently, the children of less-educated men are more likely to to come from broken families. As discussed in McDaniel et al. (2011); Autor and Wasserman (2013); Lundberg and Pollak (2014); Heckman (2011) these children receive fewer parental investments, facing larger risks of academic achievement deficits that may perpetuate current inequalities. Finally, the shortage of black males in post-secondary education may also weaken efforts to increase college campus diversity, which is "a compelling state interest," as Justice Sandra O'Connor wrote for the majority in Grutter v. Bollinger (2003).

(2016), studying the gender gap in high school completion, find that while differences in family background characteristics are small between males and females, family socio-economic status has a much stronger influence on the high school completion decisions for males compared to females. Autor et al. (2016) show that since black youth are more likely to come from families with lower socio-economic status, this sensitivity to lower socio-economic backgrounds for males is one explanation for the larger gender gap in high school completion among black youth compared to white youth. Finally, studying the aggregate gender gap in college enrollment in the United Kingdom, Aucejo and James (2015) find that differences in verbal skills between males and females is the primary explanation of the gender gap, although the size of the gap in the UK is smaller than the US.

This paper investigates the major factors contributing to the differential schooling decisions of males and females within racial groups.[3] Our focus on gender gaps within race is noteworthy as the aggregate gap may overlook important channels that are specific within race and may overstate the importance of particular channels that are irrelevant for some races. We study the gender gap in educational attainment using an empirical framework that combines a latent factor model with a sequential discrete choice schooling and arrest model starting at grade 10, which we estimate using data from the National Longitudinal Survey of Youth 1997 (NLSY97).

Our methodology builds on the previous literature in three important ways. First, we use a latent factor approach relying on 59 measures of early student information to recover three key student level characteristics, which encompass the student's family background characteristics (family), the student's math and verbal skills (math/verbal), and the student's disposition towards normative and externalizing behavior (behavior).[4] These three explanatory variables have previously been identified as influential on the aggregate gender gap in education outcomes (Autor et al., 2016; Aucejo and James, 2015; Jacob, 2002), but their relative importance for the gender gaps across race is unclear. The main benefit of the factor approach is that it allows us to efficiently extract information from conflated and mis-measured data. For example, while it is clear in Jacob (2002)

[3]In this paper race refers to race/ethnicity, which we differentiate as black, white, and Hispanic.

[4]Normative behavior has been identified with obedience, conformity and compliance. Externalizing behavior denotes negative behaviors that are directed toward the external environment.

that middle school grades have a significant affect on college enrollment; however, if grades proxy for many skills, a simple regression framework cannot distinguish which skill is most relevant. A factor model, on the other hand, is capable of disentangling and identifying multiple skills from this type of data.

Second, while previous studies have primarily focused on the gender gap at a single transition —for example Autor et al. (2016) study the gender gap in high school graduation and Jacob (2002) studies the gap in college enrollment for high school graduates— our analysis studies choices across the entire path of schooling decisions beginning in the tenth grade. This approach allows us to study which characteristics (family, math/verbal, behavior) are most influential at various stages of the schooling career and how their influence changes with different levels of education. Furthermore, understanding choices at the grade level provides a clearer picture of at which exact grade level gender gaps materialize, and it allows us to separate first and second-order policy effects. For example, improvements in behavioral outcomes may have a direct effect on college enrollment but may also increase college enrollment through increased high school graduation rates. These effects cannot be distinguished when educational attainment is studied as a single outcome.

The final distinguishing feature of our empirical approach is that, in addition to our model of grade-to-grade schooling transitions, we also incorporate a simultaneous model of yearly arrests, which allows us to study the role of social constraints on education decisions.[5] Given that 50% of black males in our sample were arrested at least once before age 25, it is critical to assess whether —conditional on family background, math/verbal skills, and behavior— these higher arrest rates for black males and males in general "push" youth to drop out of the education system.[6] In whole, our framework allows us to jointly study how different mechanisms —which include differential student characteristics, differential responses to characteristics, and social constraints— operate on education decisions over the schooling career and importantly how individual groups respond differentially to each mechanism. In this sense, we are able to distinguish which mechanisms are most important in explaining the gender gap, highlighting the mechanisms that are common across

[5]More specifically, we aim to analyze to what extent a proxy for social constraints (i.e. experiencing higher arrests rates after conditioning on background characteristics) could contribute to explain the large gender gap in educational attainment among African Americans.

[6]Brame et al. (2014) also find similar rates of arrests for minorities.

racial groups and the mechanisms that are uniquely active on the gender gap within racial groups.

This paper also introduces a new method for estimating factor models with discrete measures. We adopt the minorization-maximization (MM) algorithm in James (2016), which demonstrates how the likelihood function can be bounded by a quadratic function. This lower bound quadratic function is easily optimized in closed form, only requiring least squares operations, which makes the iterative procedure very fast and easy to implement even with the complexities of the factor structure.

Our analysis produces three main results. First, there are large differences in our behavior factor between males and females that are prevalent across all racial groups. The differences in behavior alone explain a substantial amount of the gender gap in schooling outcomes within race, which broadens the findings of Jacob (2002).[7] Counterfactual simulations indicate that holding fixed the family and the math/verbal factor and increasing only the mean in the behavioral factor of males in each racial group relative to their female peers fully explains the gender gap in high school graduation for all racial groups. These findings are stronger than Autor et al. (2016), whose analysis, based on the role of family background characteristics, is able to account for 44% of the total gender gap in high school completion among black youth. With respect to college enrollment, differences in behavior explain 50% of the gender gap for black youth, 70% of the gender gap for white youth, and all of the gender gap for Hispanic youth.[8] To a lesser degree than differences in behavior, differences in the math/verbal factor between males and females within race also contribute to the gap. For example, equalizing the mean of black and white males' math/verbal skills to their females counterparts, holding fixed the other factors, reduces the gender gap by about three percentage points for both races.

Second, we find that, for grade transitions during high school, the behavior factor has a stronger effect on decisions than the family and math/verbal factors. However, for post-secondary grade transitions, the math/verbal factor shows the strongest effect. Thus, the aggregate effect masks two

[7] Jacob (2002) shows that behavioral characteristics explain the aggregate gender gap in the transition from high school to college. Our findings extend these results by demonstrating that behavioral characteristics are crucial in explaining much of the gender gap within race and across multiple grade level transitions.

[8] College enrollment is defined as completing 13 years of education or more. This definition applies for the rest of the paper.

important within-gender and within-race differences. First, while all groups are more responsive to the behavior factor than the math/verbal and family factor in high school, males across all races are even more responsive to behavior than females. Therefore, policies designed to improve behavioral outcomes will have a larger effect on male education decisions at the high school level than females, even if such policies are targeted at both gender groups. Second, for high school grade transitions and the transition to college, black males are substantially more responsive to family background than black females. This finding is particularly salient given that we do not observe a similar phenomenon among white or Hispanic youth. In particular, we show that giving black males and females the mean family background characteristics of white females would reduce the black gender gap in college enrollment by 56%. In summary, our results indicate that gender differences in the behavior and math/verbal factors, and differential responses to family background characteristics can fully close the gender gap in high school graduation and post-secondary participation across all racial groups.

Finally, even though males are more likely to be arrested than females across all racial groups, we do not find an important role for arrests on educational gender gaps. This is primarily due to the fact that, once we condition on the family, math/verbal, and behavior factors, arrests have a near zero effect on schooling across all grade transitions.[9] This finding indicates that the negative empirical correlation of education and arrest is not a causal relationship but driven entirely by latent variables.

The rest of this paper is organized as follows: Section 2 describes the data from the NLSY97, which we use in our analysis, and summarizes education and arrest outcomes by gender and race. Section 3 outlines our main empirical model and estimation strategy. Section 4 presents the main results. Sections 5 and 6 perform counterfactual simulations to understand the sources of the gender gaps in educational attainment and arrests within race. Finally, Section 7 concludes.

[9] We do find that arrest has a small, but significant effect on completing the 12th grade.

2 Data Description

Our analysis uses the National Longitudinal Survey of Youth 1997 (NLSY97), which is a nationally representative sample of youths who were 12 to 17 years old when they were first surveyed in 1997. The NLSY97 collects extensive information on family background characteristics, educational experiences and labor market outcomes through time, with the aim to document the transition of the survey participants from school to work and into adulthood.[10] This section describes the NLSY97 data that are relevant for our research question in three parts. First, we summarize gender differences in educational attainment and arrests by race that serve as the baseline stylized facts we aim to study in our main analysis. Second, we document important differences between males and females across a range of student-level characteristics measured prior to middle school completion that we later use to explain the gender disparities in educational attainment within race. Finally, as a precursor to our main analysis that investigates the sources of the educational gender gap within race using a sequential schooling model, we conclude this section by presenting simple OLS regressions that explore how the gender gap in college enrollment for each race changes after controlling for different dimensions of student-level characteristics. This preliminary analysis offers some evidence as to which student factors may be most influential on the gap.

2.1 Gender Differences in Educational Attainment and Arrests By Race

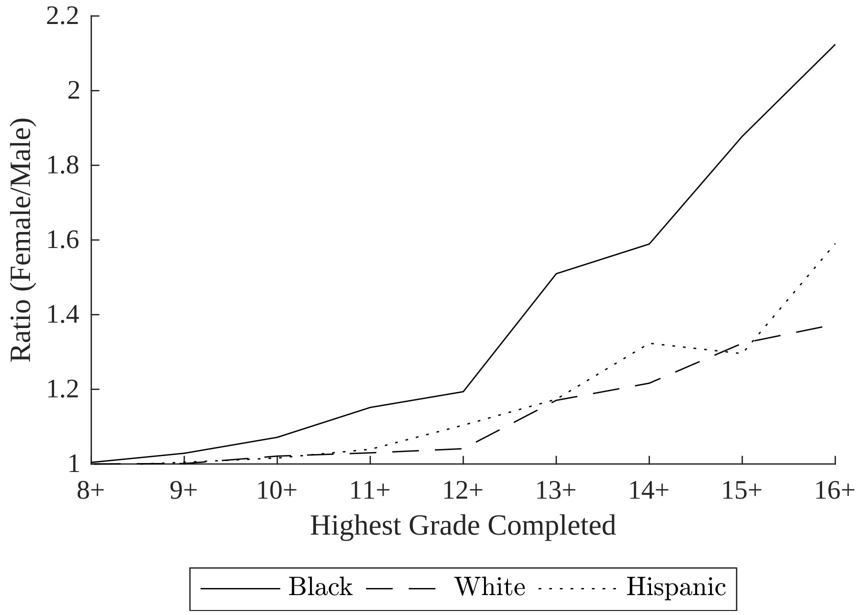
Figure 1 plots the ratio of females to males by race for different levels of educational attainment prior to age 25. Beginning at the lowest levels of education, the ratio of females to males with an eighth grade education or higher is at parity for all races. However, moving to higher levels of educational attainment, large gender imbalances emerge.[11] This figure highlights two important facts. First, there is significant heterogeneity in the gender gap across race, where the gender gap is largest for black youth. For example, for every 25 year old black male with a college degree there are two black females. Second, gender gaps appear to compound over the schooling career.

Table 1 offers more detail about the distribution of educational attainment prior to age 25 by

[10]The data used in the analysis also include the supplemental sample of black and Hispanic youth.

[11]Our analysis uses the set of respondents who at least completed the ninth grade. Figure 1 shows that at ninth grade gender ratios are near unity across races.

Figure 1: Gender Ratio By Years of Education



Note: Ratio of females to males by race for different levels of educational attainment prior to age 25.

race and gender. Prior to age 25, more than one-third of black males lack a high school degree. Of the two-thirds of black males that did finish high school, only half enrolled in college. In comparison, nearly 80% of black females completed high school and of that population 65% went on to enroll in college. Combined, prior to age 25, 50.5% of all black females had enrolled in college, compared to only 33.4% of all black males. This 17 percentage point gender gap in college enrollment among black youth is, surprisingly, larger than the black-white racial gap in college enrollment, which is 14.7 percentage points. Similar, but smaller, gender patterns in educational choices are observed for other races, where the gender gap in college enrollment is 8.9 percentage points for white youth and 6.6 percentage points for Hispanic youth. The gender disparities in educational attainment in the NLSY97, summarized in Figure 1 and Table 1, are representative of the disparities in educational decisions across gender and race observed in other national datasets.[12]

[12]For example, in the 2009 American Community Survey (ACS), the gender gap in college enrollment among 24 year old black respondents is 16.5 percentage points, while the black-white gap is 14.1 percentage points, which are both close to the gaps in Table 1. Similarly, the National Center of Educational Statistics (NCES) reports information on all undergraduates in the United States. In 2014, 55.7% of all white enrollees where female, while 44.3% of white enrollees were male, a ratio of 1.25. This ratio is similar to the ratio for white youths with 13 years of schooling or more in Figure 1. The NCES reports that among all black enrollees, 63% were female, while 37% were male, a ratio

Table 1: Educational Attainment as % of Total Demographic Group (Before age 25)

	Black		White		Hispanic	
	Male	Female	Male	Female	Male	Female
	(1)	(2)	(3)	(4)	(5)	(6)
High School Dropout	35.3	22.8	18.6	15.3	29.6	22.3
High School Degree Only	31.3	26.7	29.4	23.9	31.8	32.4
College Enrollment	33.4	50.5	52.0	60.9	38.6	45.2
Observations	1082	1096	2037	1897	890	854

Note: Columns sum to 100 percent.

The goal of this paper is to identify the main factors that give rise to the gender disparities in educational attainment observed in Table 1. Given the large and negative correlation between arrests and educational attainment (Bernburg and Krohn, 2003; Sweeten, 2006; Brame et al., 2014), one possible explanation is that the disproportionately large arrest rate for males may act as a constraint for males in educational promotion, which motivates our simultaneous model of arrest and education decisions for studying education gender gaps. The NLSY97 documents large racial and gender disparities in arrests. The top panel of Table 2 shows (average) annual rate of arrests and the fraction of respondents who were ever arrested over the age range studied in our analysis. Males in all racial groups are substantially more likely to be arrested than their female counterparts.[13] Black males have the highest arrest rates, where 50% had been arrested prior to age 25.[14] Table 2 also reproduces the statistics in Table 1 separately for those respondents that had been arrested prior to age 25 and those that had never been arrested. Separating the respondents in this way substantially reduces the observed gender gaps in educational outcomes. For example, in Table 1, the gender gap in college enrollment for black youth overall is 17 percentage points. However, this gender gap shrinks to 8.4 percentage points if we look only at black youth who had never

of 1.7, which is also consistent with the ratios for black youth with 13 years of schooling or more in Figure 1.

[13]While the reported *cumulative* rate of arrests seems large, most of the well-known available evidence corresponds to *annual* arrests rates reported by governmental agencies (e.g. Bureau of Justice Statistics). According to Brame et al. (2014) very little is known about race and gender variation in the *cumulative* prevalence of arrest. Finally, it is worth to emphasize that the annual and cumulative arrest rates reported in Table 2 are consistent with data obtained from the Bureau of Justice Statistics and the Add Health database, respectively.

[14]Disparities in the criminal justice system for black males have been well documented. For example, Fryer Jr (2016) find that blacks and Hispanics where significantly more likely to be targeted by New York City's Stop and Frisk program, even after accounting for a large set of controls.

Table 2: Arrests and Educational Outcomes

	Black		White		Hispanic	
	Male	Female	Male	Female	Male	Female
	(1)	(2)	(3)	(4)	(5)	(6)
Annual Arrest Rate	9.9	2.3	6.7	2.6	6.8	2.0
Percent Ever Arrested Before Age 25	52.9	22.1	39.6	20.8	43.0	17.3
Age 25 Educational Attainment, Never Arrested						
High School Dropout	18.2	16.7	8.1	9.9	17.3	19.6
High School Degree Only	35.0	28.0	28.0	22.9	30.7	31.4
College Enrollment	46.9	55.3	63.9	67.2	52.1	49.0
Age 25 Educational Attainment, Arrested At Least Once						
High School Dropout	50.5	44.7	34.7	35.3	45.9	36.1
High School Degree Only	28.0	22.1	31.5	27.4	33.3	37.6
College Enrollment	21.6	33.2	33.9	37.3	20.8	26.3

Note: Top panel displays (average) annual arrest rates and proportion ever arrested. The last two panels show proportion ever arrested by race, gender, and educational attainment of individuals who have never been arrested and those who were arrested at least once.

been arrested. Similarly, for white youth that had never been arrested, the gender gap in college enrollment is only 3.3 percentage points, down from 8.9 percentage points in the aggregate. In this regard, it appears that conditional on arrest, males and females (within race) experience more similar educational outcomes, so one possible explanation for the large aggregate difference may be due to the fact that males are substantially more likely to be arrested, which we explore further in our main analysis.

2.2 Gender Differences in Student-Level Characteristics

In this section, we document differences between males and females across a range of variables that are likely to be relevant for education decisions. In particular, we focus on three aspects of students' characteristics: family background, math/verbal skills, and normative/externalizing behavior.[15] These variables were constructed using only information up until the age of 14 to avoid

[15]We use these variables later in a factor model to extract three latent student-level characteristics: the student's family background characteristics, math/verbal skills, and normative/externalizing behavior.

possible endogeneity due to simultaneity with education choices. More specifically, we assume that gender differences in math/verbal skills and behavior during middle school are not driven by differences in educational expectations.[16] Evidence from the NLSY97 suggests that this assumption is reasonable given that educational attainment expectations between males and females at age 15 and 16 are very similar. For example, black males and females between the age of 15 and 16 show small differences (2.1 percentage points) in their reported expected chances of having a four year college degree by the time they turn 30.[17]

The first group of variables included in our analysis are proxies family background characteristics. Cameron and Heckman (2001) show that differences in family background characteristics account for much of the difference in education outcomes between black and white youth. However, given that males and females of the same race belong to similar families on average, it is questionable whether family socioeconomic status can explain observed gender gaps within race. One possibility is that although males and females come from similar families, family background characteristics may have differential effects on male and female decisions. For example, a family headed by a mother may have a larger negative effect on boys than on girls due to the lack of a male role model at home.

Panel A of Table 3 shows the sample average for a selection of the covariates on family background characteristics by race and gender.[18] The four family measures reported include: whether the youth's mother was a college graduate, the number of members in the youth's household under the age of 18, whether the youth was raised in a broken family, and the youth's family income. As expected, differences in family background characteristics between males and females within race are (mostly) not statistically significant, but there are large differences across races. For example,

[16]For example, our identification assumes that middle schoolers do not get involved in fights because they have already decided not to attend college.

[17]Specifically, black females a reported 76.5 percent chance of having a four-year college degree by age 30, while black males reported a 74.4 percent chance. These expectations seem to be quite overly optimistic. White and Hispanic youth report similarly large magnitudes. Carneiro et al. (2005) also find overly optimistic expectations in educational attainment when using data from the CNLSY. However, as mentioned in Carneiro et al. (2005), expectation formation models are complex, and difficult to test empirically because they often lead to multiple equilibria. Nevertheless, the reported evidence seems to support the claim that proxy measures are not likely to be biased by differences in expectations across genders.

[18]A complete description of the family measures used in our main analysis is in Table 17.

77% of black youth belonged to a broken family, compared to less than 45% for white youth.[19] Similarly, the data indicate that white students have mothers with more years of education and belong to smaller and higher-income families than black and Hispanic students.

The second group of variables that we consider are related to math/verbal skills. Neal and Johnson (1996), Cawley et al. (2001), and Aucejo and James (2015) have all documented that these skills have a significant impact on educational attainment and labor market outcomes. Panel B of Table 3 reports the average age-normalized score for four of the subject-specific test scores on the Armed Services Vocational Aptitude Battery (ASVAB) for respondents in the NLSY97 by race and gender.[20] These four subject tests —Arithmetic Reasoning, Mathematical Knowledge, Word Knowledge, and Paragraph Comprehension —are the four components of the Armed Forces Qualification Test (AFQT) score and have been widely used in the economics literature to proxy math/verbal skills.[21] For black youth, females perform much better than males across all four subjects, with the difference being largest for Paragraph Comprehension, where black females score on average 1/3 of a standard deviation higher than black males. Similarly, white females perform better on average than white males on most of these tests, where white males perform slightly better on Arithmetic Reasoning. Finally, on these tests Hispanic males have similar scores to Hispanic females, with the exception of Paragraph Comprehension, where Hispanic females score on average 0.18 standard deviations higher than Hispanic males. The fact that females in each race group appear to perform better on these tests compared to their males counterparts suggests that differences in math/verbal skills may provide one explanation for the educational gender gaps observed in Table 1.

The third group of student-level characteristics proxy early behavioral problems, which we la-

[19] Broken family is defined as any household composition that does not include both biological parents during school grades 6 to 8. This empirical regularity is consistent with the sharp rise (over the past few decades) in out-of-wedlock births and single parent family formation among African Americans.

[20] Many studies have shown that the ASVAB scores are not likely to be racial or gender biased. For example, Ree and Carretta (1995) conclude that the predictiveness of AFQT is consistent across racial and gender groups. Moreover, Neal and Johnson (1996) summarize the results of a National Academy of Sciences study (for the Department of Defense) that found that AFQT predicts performance in tasks required for military occupations about equally well across races. In this regard, they indicate that the AFQT score provides an unbiased measure of pre-market job preparation. However, whether these results can be generalized to jobs outside of the military is unknown. Finally, it is worth mentioning that gender differences in average performance in the ASVAB exams in our sample are not driven by differences in educational attainment at the time of the test.

[21] See for example, Heckman et al. (2006).

Table 3: Select Variable Means

	Black		White		Hispanic	
	Female	Gender Diff.	Female	Gender Diff.	Female	Gender Diff.
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Family Background						
Mother College Degree	0.11	-0.00	0.23	0.01	0.09	-0.00
Num. in Family Under 18	2.46	-0.04	2.10	-0.00	2.45	0.01
Broken Family	0.77	0.01	0.45	-0.03**	0.51	-0.01
Family Income	30.87	0.80	59.99	0.31	35.18	2.59
Panel B: Math/Verbal						
ASVAB Arithmetic Reas.	-0.68	-0.07	0.16	0.07**	-0.35	0.03
ASVAB Math Knowledge	-0.49	-0.24***	0.24	-0.14***	-0.30	-0.08
ASVAB Paragraph Comp.	-0.44	-0.34***	0.30	-0.24***	-0.22	-0.18***
ASVAB Word Knowledge	-0.63	-0.11**	0.21	0.02	-0.44	0.00
Panel C: Normative/Externalizing Behavior						
Any Illegal Substance Use	0.34	0.07***	0.49	0.01	0.42	0.05*
Attack	0.14	0.07***	0.07	0.07***	0.09	0.05***
Precocious Sex	0.26	0.21***	0.17	0.02*	0.19	0.10***
Suspensions	0.30	0.18***	0.09	0.15***	0.16	0.16***
GPA in Grade 8	2.80	-0.38***	3.11	-0.33***	2.84	-0.26***
School Retention	0.06	0.07***	0.03	0.01**	0.04	0.03**

Note: Mother college degree is an indicator for whether the mother has graduated from college. Broken family denotes whether at least one of the biological parents does not live in the house. Family income is measured in thousands of dollars. ASVAB tests have been normalized by age. Table 17 provides a detailed description of the variables that we use in our empirical analysis.

bel normative/externalizing behavior and broadly covers rule-breaking, anti-social, aggressive, and risky behavior. Several studies have shown the presence of large disparities in behavior between males and females across all schooling levels (Lavigne et al., 1996; Bertrand and Pan, 2013; Deater-Deckard et al., 1998; Lerner and Steinberg, 2009; Thompson et al., 2010). An emerging literature in economics documents the importance of these types of behavior in predicting many life outcomes. For example, Heckman et al. (2013) point out that aggressive, antisocial, and rule-breaking behavior can explain many labor market, health and criminal outcomes. Similarly, Segal (2013) shows that misbehavior in middle school is negatively correlated with educational attainment, and more broadly the psychology literature has related externalizing behavior with lower educational attainment (Vitaro et al., 2005) and adolescent and adult delinquency (Broidy et al., 2003; Huesmann et al., 2002).

The NLSY97 does not contain direct measures of externalizing behavior, self-control, interpersonal skills, or other personality traits, so instead we rely on a set of observed behavior measures that have been shown in the economics and psychology literature to be closely related to our concept of normative/externalizing behavior.[22] To begin, Thompson et al. (2010) show that children with high or moderate levels of externalizing behavior are more likely to engage in illegal substance use and be involved in fights, so we consider whether the youth engaged in any illegal substance use during middle school, which includes marijuana, alcohol, or tobacco, or was ever involved in a fight while in middle school.[23] Second, Schofield et al. (2008) show that children with high rates of externalizing behavior are at greater risk of early sexual activity, so we also consider whether the youth was sexually active while in middle school. Next we look at whether the youth had any suspensions while in middle school. Bertrand and Pan (2013) show that suspensions are mostly predicted by measures of externalizing behavior, interpersonal skills, and self-control. One challenge with using in-school measures, like suspensions, is that, in addition to measuring normative/externalizing behavior, these variables may also be partially determined by math/verbal skills. This issue is especially pertinent for the final two in-school behavior measures that we consider, which are the

[22]Ideally, we would like to use information similar to what is found in the Strengths and Difficulties Questionnaire (Goodman, 1997). However, no such information is available in the NLSY97.

[23]Middle school is defined as sixth, seventh, and eighth grade.

student's self-reported eighth grade GPA (grade point average) and whether the student was ever retained a grade in middle school. The degree to which grades measure cognitive skills versus behavior has become a matter of interest in its own right. Recently, Borghans et al. (2015) show that while IQ is a better predictor of achievement tests like the AQFT, measures of personality are relatively more important than IQ for predicting grades.[24] Similarly, Cooley et al. (2011) document that grade retention is correlated with both lower cognitive skills and lower social and interpersonal skills. The fact that these in-school measures reflect a combination of cognitive skills and behavior poses an empirical challenge for studying which student-level characteristics are most important in explaining the gender gap within race. This complication is one of the main motivations for the factor approach outlined in Section 3, which allows us to directly estimate the relative importance of normative/externalizing behavior and math/verbal skills on these in-school measures.

Panel C of Table 3 reports the gender differences in the behavior variables by race. Theses results shows that, for each race, relative to females, males are more likely to engage in each of these types of measured behaviors, where in many instances the difference is substantial. Consistent with the striking gender gap in educational attainment among black youth, black males in middle school are significantly more likely than black females to have used illegal substances, been involved in fights, been sexually active, been suspended, have lower academic performance, and been retained a grade.[25] For white youth, males and females have similar rates of illegal substance use. However, compared to white females, white males are twice as likely to be involved in a fight, more likely to be sexually active at an early age, more likely to be suspended, more likely to have lower academic performance, and more likely to be retained a grade. In summary, the NLSY97 data show that for each race, females on average show a better profile of behavior than males in each of the behavior measures.

[24] Almlund et al. (2011) document similar findings.

[25] The large gender disparities in precocious sex engagement are 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. Using this database, Cavazos-Rehg et al. (2009) show that by the 14th birthday, the likelihood of sexual debut is 42% for African American males while it is 17% for African American females.

2.3 Preliminary Evidence

We conclude this section by performing a set of OLS regressions using the variables in Table 3. The dependent variable in these regressions is an indicator for college enrollment before age 25. While insufficient on its own, this analysis provides some perspective on the potential magnitude of the observed differences in Table 3 on the gender gap in college enrollment. Columns (1) and (2) of Table 4 show the impact of the gender gap in college enrollment once we account for student-level family background characteristics. Since some individuals have missing values for the family background covariates, Column (1) provides a baseline of the regression for those with non-missing family background characteristics. The baseline regression shows that the size of the gender gap in college enrollment is almost 17 percentage points for black youth, and around 9 percentage points for white and Hispanic youth. This regression also exhibits the well known disparities in educational attainment across races, where white youth are substantially more likely to attend college than black and Hispanic youth. Column (2) adds to the baseline regression controls for family background characteristics, i.e. mother's education level, number of household members under age 18, and broken family.[26] Similar to Cameron and Heckman (2001), the second regression shows that controlling for family background characteristic has a large effect on racial education gaps. For example, conditional on family background characteristics, black females are more likely to enroll in college compared to white females. However, since family background characteristics are so similar between males and females of the same race, column (2) demonstrates that controlling for these variables (without including gender-race-family interactions) has little impact on the gender gaps within race.[27]

Columns (3) and (4) in Table 4 show the influence of our math/verbal measures on the college enrollment gaps. Column (3) presents a baseline regression that only controls for race and gender for the population who had non-missing ASVAB scores. Column (4) shows the residual gap after we account for differences in the four AFQT subject tests: Mathematical Knowledge, Arith-

[26]Family income was not included in the regression due to the number of missing values. Specifications that include this covariate provide similar results.

[27]Including all the gender-race-family interactions in this regression would make difficult the interpretations of the results. Therefore, we postpone this analysis when estimating the sequential model of scholling.

Table 4: Probability of College Enrollment (OLS)

	Family Controls		Math/Verbal Controls		Behavior Controls	
	Baseline	w/ Controls†	Baseline	w/ Controls††	Baseline	w/ Controls†††
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	0.622***	0.546***	0.638***	0.520***	0.634***	0.530***
	(0.012)	(0.011)	(0.012)	(0.011)	(0.012)	(0.010)
Black	-0.099***	0.042**	-0.109***	0.118***	-0.127***	-0.014
	(0.021)	(0.020)	(0.022)	(0.020)	(0.021)	(0.019)
Hispanic	-0.143***	-0.004	-0.163***	-0.002	-0.173***	-0.103***
	(0.024)	(0.023)	(0.025)	(0.023)	(0.023)	(0.021)
Black×Male	-0.169***	-0.168***	-0.180***	-0.118***	-0.164***	-0.025
	(0.025)	(0.023)	(0.026)	(0.023)	(0.025)	(0.024)
White×Male	-0.087***	-0.092***	-0.094***	-0.059***	-0.075***	0.008
	(0.017)	(0.015)	(0.017)	(0.015)	(0.017)	(0.014)
Hispanic×Male	-0.090***	-0.090***	-0.054*	-0.032	-0.049*	0.047*
	(0.029)	(0.028)	(0.030)	(0.028)	(0.028)	(0.027)
Observations	6775	6775	6290	6290	6598	6598

† Family Background controls include broken family, number of household members under age 18, and mother college graduate.
††Math/verbal controls include performance in the AFQT subtests - mathematical knowledge, arithmetic reasoning, word Knowledge, and paragraph comprehension.
††† Behavior controls include any illegal substance use, attack, precocious sex, ever suspended, ever retained at school, and GPA in grade 8.
***, ** and * indicate significance at 1%, 5%, and 10% level, respectively.

metic Reasoning, Word Knowledge, and Paragraph Comprehension. Again, similar to Cameron and Heckman (2001), controlling for the ASVAB tests completely reverses the race gap in college enrollment. However, even after controlling for differences in math/verbal skills, the gender gap in college enrollment for black and white youth remains large and statistically significant. In total, differences in math/verbal skills alone can account for at most 35% of the gender gap.[28]

The final two columns of Table 4 present regression results of the residual gender and race gap for college enrollment after we control for the early behavior measures from Panel C of Table 3, which include controls for illegal substance use, being involved in fights, early sexual activity, middle school suspensions, eighth grade GPA, and middle school retentions. Column (5) shows the baseline gaps for the population with non-missing values in these behavior measures. Column (6) shows the corresponding gaps after controlling for the behavior measures, which nearly eliminates the gender gaps across all races, even reverting the gender gap for Hispanics. As mentioned previously, some of the behavior measures like GPA and retention also capture math/verbal skill, so the dramatic elimination of the gender gap may not be entirely due to the normative/externalizing behavior component. However, given that Column (4) shows that math/verbal skills on their own are not nearly enough to account for the entirety of the gender gaps, the results in Column (6) suggests that normative/externalizing behavior is the main contributing factor to the observed gaps.

The preliminary evidence in Table 4 is insufficient on a number of important dimensions. First, differential responses to skills by race and gender are not taken into account. For example, while differences in family background characteristics do not explain gender gaps within race, these regressions do not address how differential responses to family background characteristics may contribute to observed education gaps, for example the type of effects documented in Autor et al. (2016). Second, in order to determine which of these three types of variables, family background, math/verbal, or normative/externalizing behavior is most important in explaining the gaps, all three types of variables must be incorporated into a competing analysis rather than analyzed separately. Third, Figure 1 demonstrates that the observed gender gaps at any given level of education

[28] The reduction of the gender gap after including the ASVAB tests is not driven by differences in educational attainment between boys and girls at the time of the test (i.e. specifications that control for years of schooling at the time of the ASVAB test show similar results).

are the outcome of a cumulative sequence of decisions. Therefore the college enrollment gap is just one of potentially many educational gaps of interest. Fourth, given that many variables used in our analysis, like GPA, are proxying both math/verbal and normative/externalizing behavior factors, the reduced form analysis does not offer any method to isolate the two effects. Finally, these regressions do not address the potential importance of differential arrest rates between males and females, which Table 2 suggests may be a contributing factor to the gender gaps in educational attainment. Addressing these five issues motivates our empirical strategy that relies on a structural equation modeling approach outlined in the next section.

3 Empirical Model and Estimation

Our empirical strategy implements a factor model to condense a large set of measures (a selection of which are summarized in Table 3), into three latent student characteristics: a student's (1) family background characteristics, (2) math/verbal skills, and (3) normative/externalizing behavior. We embed the estimation of these factors into a structural sequential model of schooling and arrests to study the relative importance of these factors for education decisions and arrest outcomes, and ultimately analyze how differences in these variables between males and females contribute to the observed educational gender gaps for each race. In this section, we first outline the identification of the factors and then present the schooling decision model.

3.1 Identification of Factors

We use a total of 59 measures from the NLSY97 to identify three unobserved factors: a student's family background characteristics, math/verbal skills, and normative/externalizing behavior. Each measure is constructed from information obtained before the student completed middle school. While in Section 2 we provide a brief characterization of some of the variables that we use, a more detailed description of the measures is provided in Appendix B. Notably, we use as much detail as possible in constructing the measures. For example, rather than looking at only whether a student was ever suspended while in middle school, we construct six measures on suspensions that provide much more detail. Specifically, the measures are (1) whether the student was ever suspended in the 6th grade, (2) if suspended in the 6th grade, the number of days suspended, (3) whether the student was ever suspended in the 7th grade, (4) if suspended in the 7th grade, the number of days suspended, (5) whether the student was ever suspended in the 8th grade, and (6) if suspended in the 8th grade, the number of days suspended. By organizing the data in this way allows the factors to have a non-linear effect on the measures, differentially affecting the measures at the extensive and intensive margin. Furthermore, this richer level of detail generates greater variation between students, which allows us to better identify individual student profiles.

The cornerstone of a factor model is that each of the observed measures in the data is a function of the latent variables plus independent measurement error. Each measure differs in how it loads onto each of the factors and its amount of measurement error. By making mild restrictions on the loadings, for example, only allowing parent's income to load on the family background factor and not allowing it to load on the other two factors, the estimated model will produce a unique and interpretable set of factors. Of the 59 measures that we use, 19 are continuous, 29 are binary and 11 are multinomial.[29]

Each individual is characterized by a 3×1 unobserved factor vector, θ_i, where i indexes the individual. The elements of this vector are individual i's family background characteristics, math/verbal skills, and normative/externalizing behavior. There are two types of measures, continuous and discrete. We model the continuous measures as a linear function of the factors. If measure m is a continuous variable then it is described by,

$$w_{im} = \mu_m + \lambda_m' \theta_i + \eta_{im}$$

where μ_m is the mean of measure m and λ_m are the corresponding factor loadings. The final term, η_{im} is the residual component of the measure that is not explained by the factors, and is assumed to be independent of the factors, measures, schooling decisions, and arrest outcomes. Assuming that the measure residual is normally distributed with mean zero and variance σ_m^2, the probability

[29] Appendix B shows the restrictions on the loadings for each variable.

of observing the continuous measurement w_{im} conditional on the unobserved factors is

$$\Pr(w_{im}|\theta_i) = \frac{1}{\sqrt{2\sigma_m^2\pi}} \exp\left(-\frac{(w_{im} - \mu_m - \lambda_m' \theta_i)^2}{2\sigma_m^2}\right) \tag{1}$$

If measure m is a discrete variable, then w_{im} takes observed values $1, 2, \ldots, J_m$. These outcomes are determined by the J_m latent variables $w^*_{im1}, w^*_{im2}, \ldots, w^*_{imJ_m}$ and

$$w_{im} = \underset{j=1,2,\ldots,J_m}{\text{argmax}} \{w^*_{imj}\}$$

For all of the discrete measures, $w^*_{im1} = 0$ by normalization, the remaining latent variables are defined by,

$$w^*_{imj} = \mu_{mj} + \lambda'_{mj}\theta_i + \eta_{imj} \qquad \text{for } j = 2, \ldots, J_m$$

where $\eta_{im} = \{\eta_{im2}, \ldots, \eta_{imJ_m}\}$ is a $J - 1$ vector of unobserved residuals to the latent variables, which are assumed to be independent of all of the other data and logistically distributed. The probability of any observed outcome has a multinomial logit form

$$\begin{aligned}
\Pr(w_{im} = 1|\theta_i) =& \frac{1}{1 + \sum_{k=2}^{J_m} \exp\left(\mu_{mk} + \lambda'_{mk}\theta_i\right)} \\
\Pr(w_{im} = j|\theta_i) =& \frac{\exp\left(\mu_{mj} + \lambda'_{mj}\theta_i\right)}{1 + \sum_{k=2}^{J_m} \exp\left(\mu_{mk} + \lambda'_{mk}\theta_i\right)} \qquad \text{for } j = 2, \ldots, J_m
\end{aligned}$$

And the probability of observing the discrete outcome w_{im} is

$$\Pr(w_{im}|\theta_i) = \prod_{j=1}^{J_m} \Pr(w_{im} = j|\theta_i)^{\mathbb{1}(w_{im}=j)} \tag{2}$$

Letting $w_i = \{w_{i1}, w_{i2}, \ldots, w_{i59}\}$ denote the vector of observed measures and using the probability in Eq. (1) for the continuous measures and the probability in Eq. (2) for the discrete measures, then the probability of the entire vector of observed measures conditional on θ_i is

$$L(w_i|\theta_i) = \prod_{m=1}^{59} \Pr(w_{im}|\theta_i) \tag{3}$$

3.2 Schooling Decisions

In each period, individuals make a decision to increase their level of education. Let h_{it} represent individual i's current level of education in years at the beginning of period t. We assume individuals begin making education decisions after they complete the 9th grade, so $h_{i1} = 9$ for all individuals.[30] The schooling decision in period t is $d_{it} \in \{0, 1\}$ where one indicates they choose to increase their years of education and zero they do not. Given the education decision, the following period's education level is $h_{it+1} = h_{it} + d_{it}$.

Individuals increase their level of education if the value of increasing education is larger than the value of remaining at the current level. Let V_{it} denote the net benefit of an additional year of schooling, so $d_{it} = 1$ if $V_{it} > 0$ and $d_{it} = 0$ otherwise. The value function is defined as:

$$V_{it} = \alpha_{hg} + \beta'_{hg}\theta_i + \omega'_h \begin{bmatrix} t & \mathbb{1}(d_{it-1} = 0) & a_{it} \end{bmatrix}' + \varepsilon_{it} \tag{4}$$

The value function depends on h_{it}, the individual's current education level, and $g_i \in \{1, 2, \ldots, 6\}$, which indicates the race/gender group of the individual and maps numerically as 1:black male, 2:black female, 3:white male, 4:white female, 5:Hispanic male, and 6:Hispanic female. In this equation, α is the intercept and β is the coefficient on the factors, which represents how each factor affects schooling decisions. Both the intercept and the factor responses depend on years of schooling h, and the individual's race/gender group g, allowing groups to respond differently to the factors and allowing the factors to have different effects over the schooling career.

The third term in Eq. (4) incorporates three additional covariates into the value function. The first is time, t, which represents years since completing the 9th grade. This accounts for the fact that completing additional years of schooling may become more difficult as the individual gets older. The second covariate captures the impact of schooling disruption on the value function. The

[30] Students can drop out of school at age 16 or later, which is approximately when they have completed 9th grade.

indicator function $\mathbb{1}(d_{it-1} = 0)$ takes a value of 1 if the individual did not complete a grade in the previous period. Since the first period of the decision model is defined as the year after the individual completes the ninth grade, then $d_{i0} = 1$ for all individuals.[31] The final covariate, a_{it}, is an indicator for whether the youth is arrested in period t. The vector of coefficients on these three variables, ω_h, depends on the individual's current level of schooling so that arrests, for example, can have a differential impact on educational progression for a youth with a 10th grade education, versus a youth with a 12th grade education.

The final term in Eq. (4), ε, is an unobserved taste shock to the value of increasing educational attainment, which we assume to be independent of the factors, θ_i. The taste shock is assumed to be drawn from a logistic distribution. Therefore, conditional on the unobserved factors, the probability that an individual completes an additional year of schooling is:

$$\Pr(d_{it} = 1|\theta_i) = \frac{\exp\left(\alpha_{hg} + \beta'_{hg}\theta_i + \omega'_h \begin{bmatrix} t & \mathbb{1}(x_{it} \geq 2) & a_{it} \end{bmatrix}'\right)}{1 + \exp\left(\alpha_{hg} + \beta'_{hg}\theta_i + \omega'_h \begin{bmatrix} t & \mathbb{1}(x_{it} \geq 2) & a_{it} \end{bmatrix}'\right)}$$

We model education decisions from the time individuals complete the 9th grade ($t = 1$) to T_i^d where either the individual graduates college (completing the 16th year of schooling), reaches age 25, or attrits from the sample. Letting $d_i = \{d_{i1}, d_{i2}, \ldots, d_{iT_i^d}\}$ represent the full set of observed schooling decisions for individual i, then assuming the taste shocks are independent across time periods, the likelihood of observing this sequence of choices conditional on the unobserved factors is:[32]

$$L(d_i|\theta_i) = \prod_{t=1}^{T_i^d} \left\{[1 - \Pr(d_{it} = 1|\theta_i)]^{\mathbb{1}(d_{it}=0)} \times \Pr(d_{it} = 1|\theta_i)^{\mathbb{1}(d_{it}=1)}\right\} \tag{5}$$

[31]These variables were included to improve the precision of the model. Notice that they do not play any role on explaining the gender gap in educational attainment (see Appendix Table 16).

[32]Permanent unobserved heterogeneity is accounted for through the factors, θ_i, therefore serial independence of ε is a reasonable assumption.

3.3 Arrests

To assess the role that arrests may have on educational attainment and its implications for the observed gender gaps, we model arrests simultaneously with education decisions. We assume that the observed arrest outcome, $a_{it} = 1$, occurs when the latent variable index $A_{it} > 0$, where

$$A_{it} = \kappa_g + \tau'\theta_i + \rho' \begin{bmatrix} t & t^2 \end{bmatrix}' + \nu_{it}$$

The latent arrest index has a race/gender group specific mean, κ_g, and depends on the unobserved factors through parameters τ. t and t^2, which represent years since completing 9th grade, where included to capture the age-arrest profile. The final term ν is an arrest shock which is assumed to be independent of the factors and the schooling shock, ε. This framework allows for correlation in the schooling decision and arrest outcome through the unobserved factors and the direct effect of being arrested on completing an additional year of schooling.

Assuming the arrest shock is also drawn from a logistic distribution, the probability of arrest conditional on the unobserved factors is:

$$\Pr(a_{it} = 1|\theta_i) = \frac{\exp\left(\kappa_g + \tau'\theta_i + \rho' \begin{bmatrix} t & t^2 \end{bmatrix}'\right)}{1 + \exp\left(\kappa_g + \tau'\theta_i + \rho' \begin{bmatrix} t & t^2 \end{bmatrix}'\right)}$$

Arrest outcomes are modeled from the time the individual completes the 9th grade ($t = 1$) to T_i^a where either the individual reaches age 25, or attrits from the sample. Letting $a_i = \{a_{i1}, a_{i2}, \ldots, a_{iT_i^a}\}$ represent the full set of observed arrest outcomes, the likelihood of observing these outcomes conditional on the unobserved factors is:

$$L(a_i|\theta_i) = \prod_{t=1}^{T_i^a} \left\{ [1 - \Pr(a_{it} = 1|\theta_i)]^{\mathbb{1}(a_{it}=0)} \times \Pr(a_{it} = 1|\theta_i)^{\mathbb{1}(a_{it}=1)} \right\} \tag{6}$$

3.4 Factor Distributions

To complete the description of the model, we assume that the underlying distribution of factors can be modeled as a mixture of normals. We use eight components of the mixture and allow the mixing probabilities to be race/gender specific.[33] This approach is more flexible than simply assuming factor specific means for each of the six race/gender groups and in fact includes as a special case the model with race/gender specific means. Let γ_c denote the mean of component c in the mixture. We assume all components share the same covariance matrix Δ. An individual who is in race/gender group g belongs to component c with probability π_{cg}. Therefore, the probability density function of the factor for an individual in group g_i is $p(\theta|g_i) = \Pr(\theta = \theta_i|g_i) = \sum_{c=1}^{8} \pi_{cg} f(\theta|\gamma_c, \Delta)$, where $f(\cdot|\gamma, \Delta)$ is the probability density function of the normal distribution with mean γ and covariance Δ.

The location and scale of the factors are not identified without normalizations. The location of the factors is identified by setting the measurement mean to zero for three of the measures, so we fix the mean of the measures (μ) for family income measure, GPA in the eighth grade, and suspensions in eighth grade to zero. The scale of the factors are identified by normalizing the factor loading. We normalize the loading on the family factor for family income to one, the loading on the math/verbal factor in GPA in the eighth grade to one, and the loading on the behavior factor on days of suspension in eighth grade to negative one.[34] GPA and suspensions load on both the math/verbal factor and the behavior factor, where the loading that is not subject to a normalization is estimated.

Finally, we assume some of the loadings in the model are zero. For example, the family income equation only loads on the family background factor and does not load on the math/verbal or behavior factors. Likewise, GPA loads on both the math/verbal and behavior factors but does not load on the family background factor. These assumptions do not imply that family income and GPA are independent. Instead, this specification only implies that family background has no direct effect on GPA. However, family income and GPA will be correlated because the underlying factors are correlated, i.e., people with higher values of the family background factor will tend to have higher values of the math/verbal factor; thus students from high income families will tend to have higher GPA outcomes.

[33]Specifications with 9 and 10 components did not show an appreciable improvement in the log-likelihood.

[34]The loading on the behavior factor in suspensions is fixed to negative one, so that lower levels of normative/externalizing behavior imply more days suspended.

3.5 Estimation

We estimate the parameters of the model using maximum likelihood. Let Ψ denote the full set of parameters to be estimated, which we partition into

$$\Psi = \left\{ \begin{array}{ll} \Psi_d = \left\{\{\{\alpha_{hg}, \beta_{hg}\}_{g=1}^{6}\}_{h=9}^{15}, \{\omega_h\}_{h=9}^{15}\right\} & \text{, Schooling decision parameters} \\ \Psi_a = \left\{\{\kappa_g\}_{g=1}^{6}, \tau, \rho\right\} & \text{, Arrest outcome parameters} \\ \Psi_w = \{\mu_m, \lambda_m, \sigma_m^2\}_{m=1}^{59} & \text{, Measurement outcome parameters} \\ \Psi_\theta = \left\{\{\{\pi_{cg}\}_{g=1}^{6}\}_{c=1}^{8}, \{\gamma_c\}_{c=1}^{8}, \Delta\right\} & \text{, Factor distribution parameters} \end{array} \right\}$$

The full model log-likelihood function is

$$LL(\Psi) = \sum_{i=1}^{n} \ln \left(\int L(d_i|\theta, \Psi_d) L(a_i|\theta, \Psi_a) L(w_i|\theta, \Psi_w) p(\theta|g_i, \Psi_\theta) d\theta \right) \tag{7}$$

Finding the more than 400 parameters that maximize Eq. (7) presents a number of challenges. The main issue is that the integral does not have a closed form solution and must be evaluated numerically. For example, Heckman et al. (2015) estimate a factor model with discrete measures and use Gauss-Hermite quadrature to numerically evaluate the integral. If the measures and outcomes consisted of only continuous data, the integral would have a closed form solution, so the requirement of numerical integration only surfaces because of the presence of the discrete measures and discrete outcomes. In maximum simulated likelihood (MSL), all methods to numerically simulate the integral will require a change of variables so that the sampling points are taken from a parameter-free distribution and then transformed to the desired distribution inside of the likelihood for the measures and outcomes.[35] This change of variables introduces many new parameters that are persistent throughout the likelihood, which makes optimizing the log-likelihood function difficult.[36] These difficulties worsen as the dimension of the integration increases or when the factors are correlated.

Given the difficulties associated with directly maximizing the simulate likelihood function, we

[35] For example, the covariance matrix must be converted to its cholesky components and inserted into the likelihood of observed measures and outcomes.

[36] Specifically, it makes finding the analytical gradient of the log-likelihood more challenging.

pursue an alternative approach, which is to maximize Eq. (7) through the use of simple surrogate functions. The estimation strategy developed in James (2016) shows that for a given value of the parameters Ψ_o, $LL(\Psi)$ can be bounded by a quadratic function $Q(\Psi|\Psi_o)$. The maximum of the surrogate function, $Q(\Psi|\Psi_o)$, is closed form and has a least squares representation $(X'X)^{-1}X'Y$. Finding $\Psi' = \operatorname{argmax}_\Psi Q(\Psi|\Psi_o)$ and then replacing $\Psi_o = \Psi'$, the process is simply iterated until the parameters converge. This approach has two primary strengths over other methods. First, it avoids deriving and coding the gradient of $LL(\Psi)$, making it much easier to implement. Each iteration only entails a single evaluation of the simulated likelihood and then one least squares computation for each of the continuous and discrete measures and outcomes. Second, unlike MSL, it requires no change of variables for the function of integration, which offers more flexibility in specifying the distribution of the latent factors and more flexibility in how the factors influence the outcome equations, for example allowing the factor responses in the schooling equation to differ by race and gender. This second feature allows us to easily estimate many versions of our factor model with very little modification of the computer code. The full details of the estimation algorithm are presented in Appendix C.

4 Main Results

This section presents the estimation results from the factor and sequential schooling and arrest model described in Section 3. First, we summarize the estimated factor distribution for each race/gender group. Next, we document the role of the factors in schooling decisions, emphasizing the overall relative importance of the factors at each grade-level transition, as well as their differential effects for males and females by race. We conclude the description of the main results by discussing the role of the factors on explaining arrest outcomes and discussing the link between education, arrests, and the factors.

4.1 Factor Distributions

Figure 2 shows the distributions for the family, math/verbal, and behavior factors by race and gender. The plotted distributions are highly non-normal, which underscores the importance of

Table 5: Factor Means (standardized) By Race and Gender

	Black		White		Hispanic	
	Female	Gender Diff.	Female	Gender Diff.	Female	Gender Diff.
	(1)	(2)	(3)	(4)	(5)	(6)
Family	-0.618***	-0.016	0.252***	0.031	-0.630***	-0.012
	(0.036)	(0.050)	(0.019)	(0.034)	(0.041)	(0.057)
Math/Verbal	-0.426***	-0.296***	0.339***	-0.218***	-0.397***	-0.169**
	(0.049)	(0.071)	(0.029)	(0.049)	(0.058)	(0.071)
Behavior	0.029	-0.646***	0.266***	-0.405***	0.335***	-0.528***
	(0.042)	(0.057)	(0.031)	(0.054)	(0.055)	(0.070)

Note: Non-standardized factor moments are reported in Appendix Table 19. Standard errors are reported in parentheses. *** and ** indicate significance at 1% and 5% level, respectively.

using a flexible distributional assumption, like a mixture of normals, when estimating latent factors. In particular, the behavioral factor is nearly bimodal for all groups.[37]

Table 5 reports the standardized factor means by race and gender which quantifies the differences between males and females for each race in the plotted distributions.[38] Notably, gender differences in the mean of the family factor are not statistically different from zero. However, males in each race have a much lower average math/verbal factor compared to females, with the largest gap being among black youth, where the difference in the average math/verbal factor is 0.3 standard deviations.[39] The factor with the greatest gender differences overall is the behavior factor, where the difference is at least twice as large as the difference in the math/verbal factor for all races. Again, the greatest gender disparities in the behavior factor is among black youth, where the difference of the average is 0.65 standard deviations. However, gender differences for Hispanic youth and white youth are large as well as 0.53 and 0.41 standard deviations respectively.

Another key aspect of the factor distribution is the correlation of the three factors, which is pre-

[37]The shape of the factor distributions are somewhat consistent with the findings in Heckman et al. (2006). In particular, they reject the null hypothesis that their factors (i.e. cognitive and non-cognitive) are normally distributed. Appendix Table 18 reports the mean and shares of the factor type estimates.

[38]Non-standardized factor moments are reported in Appendix Table 19.

[39]The female advantage in math/verbal is consistent with the findings in Aucejo and James (2015), which show that females present higher verbal skills than males, while there are no differences in math skills. Unfortunately, we cannot separately identify math and verbal factors due to the low number of measurements.

Figure 2: Factor Distribution By Race and Gender

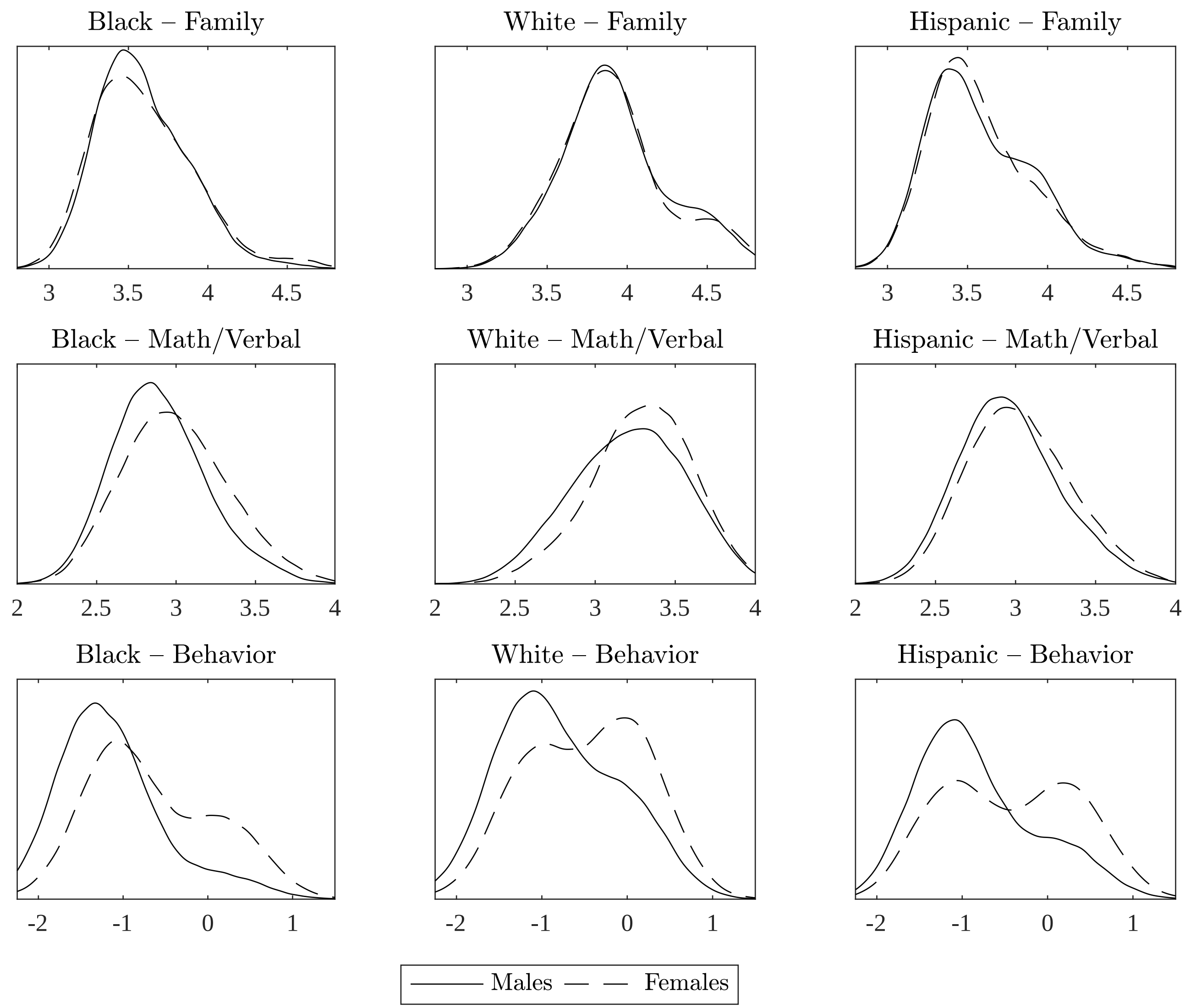


Table 6: Factor Correlations, Overall

	Family	Math/ Verbal	Behavior
Family	1	–	–
Math/Verbal	0.598*** (0.019)	1	–
Behavior	0.363*** (0.027)	0.262*** (0.048)	1

Note: Correlations by gender and race are reported in Appendix Table 19. Standard errors are reported in parenthesis. *** indicate significance at 1% level.

sented in Table 6. This table reports the overall correlation, which is a function of the race/gender specific mixture component probabilities, the mixture component means, and the shared covariance.[40] Not surprisingly, all of the correlations between the factors are positive and statistically significant different from zero. The correlation between the family and math/verbal factors is 0.598, which is larger than the correlation between the family factor and the behavior factor (0.363). The smallest correlation is between the behavior factor and the math/verbal factor, which is 0.262. This relatively small correlation suggests that the math/verbal factor and the behavior factor are capturing different aspects of student characteristics.[41]

Next, Table 7 provides information on the identification of the factors from the measures. Instead of reporting the factor loadings for the measures this table shows the average marginal effect of the factors on the measurements when each factor is increased one standard deviation, holding fixed the other factors. Only six of the 59 measures are summarized in the table. Similar information is reported for all of the measures in Appendix Table 20. Mother's degree and broken home are variables that only load on the family factor, so the average marginal effect for the math/verbal factor and the behavior factor is assumed to be zero. Mother's degree takes one of five values. While 18.5% of the population have mothers with a college degree, increasing the

[40] The race/gender specific correlations are reported in Appendix Table 19 and for the most part mimic the overall correlation.

[41] Heckman et al. (2013) find that the correlations between the externalizing behavior and cognitive factors are 0.254 for females and 0.099 for males, which are not very different from the ones reported in this paper in light of the fact that the identification of the factors is obtained from different sets of measurements.

Table 7: Summary of Measurements: Average Marginal Effects of Factors

Measure	Categorical Values	Average Value	Family	Math/ Verbal	Behavior
10. GPA in Grade 8	Continuous	2.866	0	0.362*** (0.019)	0.331*** (0.017)
18. Mother Degree	Missing	0.064	-0.043*** (0.002)	0	0
	HS Dropout	0.168	-0.128*** (0.004)	0	0
	HS Deg.	0.351	-0.093*** (0.008)	0	0
	Some Col.	0.232	0.030*** (0.009)	0	0
	Col. Deg.	0.185	0.234*** (0.009)	0	0
31. Attack 8	Yes	0.042	0	0	-0.031*** (0.002)
32. Broken Household	Yes	0.504	-0.227*** (0.006)	0	0
35. Any Days Suspended 8	Yes	0.113	0	-0.033*** (0.006)	-0.085*** (0.005)
53. Retention 8	Yes	0.019	0	-0.008*** (0.002)	-0.017*** (0.002)

Note: Average marginal effects correspond to the loadings in the measurement equations described in Section 3. Standard errors are reported in parentheses. Appendix Table 17 provides a detailed description of all these measurements. Measure numbers allow to match variables across tables. *** indicate significance at 1% level.

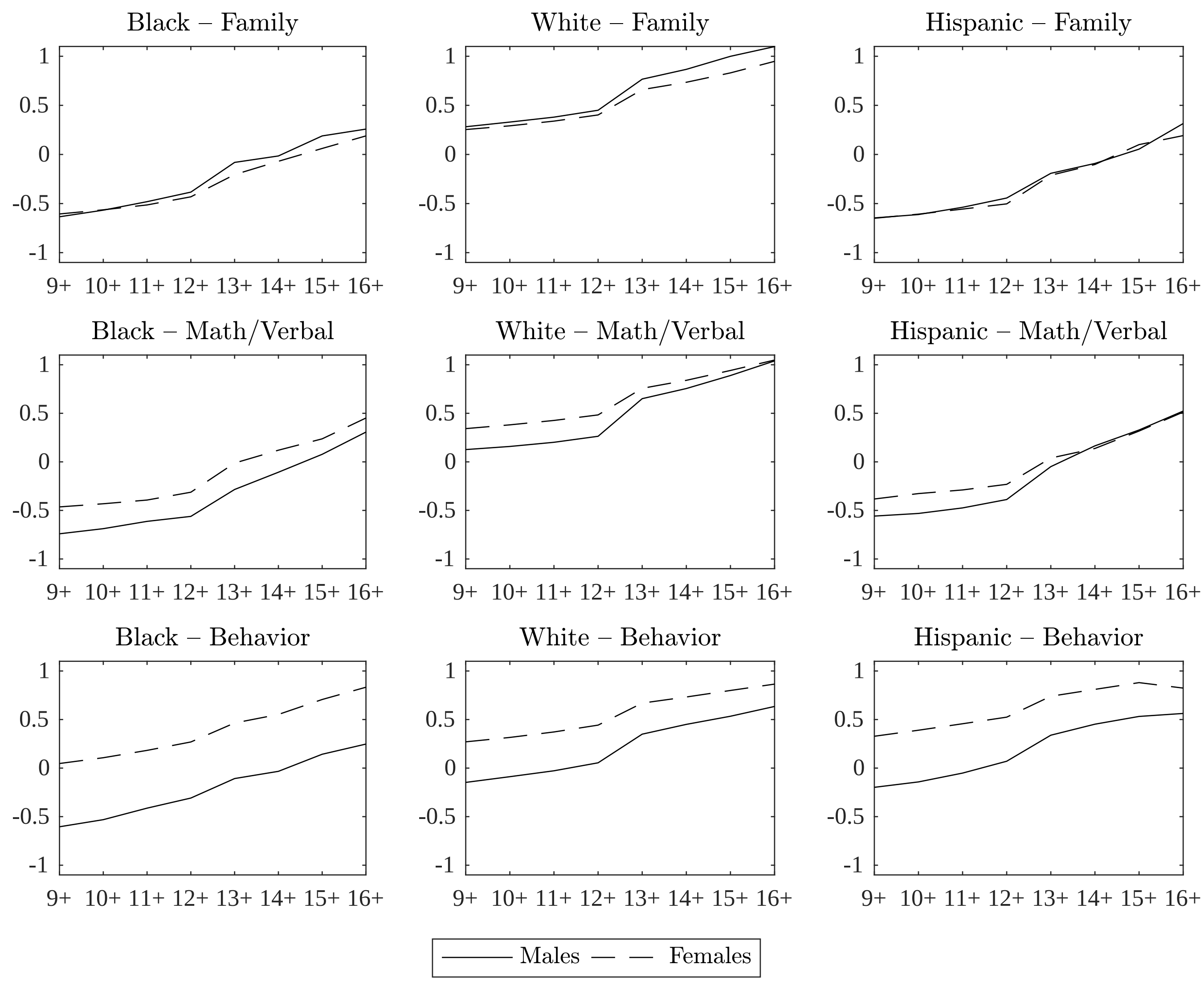
family factor by one standard deviation increases the probability of having a mother with a college degree by 23.4 percentage points. Increasing the family factor also reduces the probability that the mother's education level is missing, supporting the notion that youth who fail to report their mother's education are more likely to come from a disadvantaged background. One of the variables that only loads on the behavior factor is whether the youth was involved in a fight in the eighth grade. This measure is highly correlated with the behavior factor, such that increasing the behavior factor (positive is better) by one standard deviations decreases the probability that an individual would be involved in a fight by 75%.

Of the measures that load on both the math/verbal factor and the behavior factor, GPA in the eighth grade is the measure that loads most evenly on these two factors. For example, a one standard deviation increase in the math/verbal factor would increase GPA on average by 0.36 grade points, which is very close to the 0.33 average increase in GPA if the behavior factor is increased one standard deviation. This finding is similar to Almlund et al. (2011) who claim that the association between school grades and personality traits is almost as large as the association between cognitive skills and grades. School suspensions and retention, on the other hand, load more heavily on the behavior factor than the math/verbal factor, although both are statistically significant. Specifically, while about 11% of the sample is suspended in the eighth grade, increasing the behavior factor one standard deviation decreases the probability of being suspended on average by 8.5 percentage points, compared to only a 3.3 percentage point reduction if the math/verbal factor is increased one standard deviation.

To conclude the description of the factor estimates, while Table 5 summarizes the overall factor mean by race and gender, Figure 3 characterizes the dynamic selection process by plotting the average of the standardized factors by educational attainment. The left side of each plot in the figure shows the average of each factor for those with nine or more years of schooling. Since nearly all of the people in our analysis completed at least the ninth grade, these numbers are similar to the overall means in Table 5. By looking at higher levels of education, these plots illustrate how the mean composition of the factors change as segments of students leave the education system. For example, those completing college have values for the family factor that are typically about 0.8 standard deviations higher than their race/gender group mean.

Furthermore, from Figure 3 we can see that even across race/gender, many of the factors share a common profile, indicating a structural break at the first year of college. Not only is there a discrete jump in the factor mean from 12 to 13 years of schooling but in general, the selection slope is much steeper for the college years compared to the high school years, which points to the factors playing a greater role for post-secondary education decisions. Lastly, we can see from these plots that, although females in each race have an overall advantage in the math/verbal factor and behavior factor, these skill gaps shrink through selection as males and female progress through the education system. In fact, by the end of college, the gender gap in the math/verbal factor for white and Hispanic youth is completely eliminated.

Figure 3: Factor Means (standardized) By Years of Education at Age 24 By Race



While Figure 3 offers a useful characterization of educational sorting on the factors, these plots do not provide clarity on the underlying mechanisms driving educational decisions. For example because the family factor and the math/verbal factor are so highly correlated, selection on the family factor will cause changes in the average math/verbal factor, even if individuals are not selecting directly on math/verbal. This makes drawing inferences from these plots difficult. To get at the main mechanisms driving the selection process we need to look at the coefficient estimates of the structural choice model, which we do in the next subsection.

4.2 Schooling Decisions

Our schooling decision model has seven grade transitions: tenth grade to college completion. Since we allow the three factors to have differential effects by race and gender across these seven grade transitions, we are estimating 126 ($7 \times 6 \times 3$) marginal effects. We summarize these effects in two ways. First, since we estimate the underlying race/gender effects for the factors, we can construct an aggregate population effect of each factor at each grade transition, which allows us to study, on the whole, the overall importance of each factor at different stages of the schooling career. Second, to summarize the differential responses to the factors by race/gender, for each race/gender group we compute the cumulative effect of each factor on three key schooling milestones: high school completion, college enrollment, and college completion; and compare the differential effects across gender for each race.

4.2.1 Overall Impact of the Factors

To characterize the relative importance of the factors at different stages of the schooling career, Table 8 constructs the population average marginal effects for each factor at each grade level. The population effect is constructed by averaging over each of the race/gender specific effects. Since the schooling equations differ at the grade level, each column is reporting an average marginal effect on a conditional probability. The bottom row of the table shows the baseline conditional probabilities, which shows for example that among those that completed 11 years of schooling, 92.3% will complete the 12th year. To allow us to aggregate the marginal effects over first-time and repeat grade

Table 8: Population Average Marginal Effects of Factors on Grade Completion

	10th	11th	12th	13th	14th	15th	16th
Family	0.017***	0.024***	0.037***	0.107***	0.041***	0.058***	0.024
	(0.002)	(0.004)	(0.006)	(0.011)	(0.011)	(0.014)	(0.015)
Math/Verbal	0.011***	0.017***	0.028***	0.164***	0.090***	0.103***	0.127***
	(0.003)	(0.004)	(0.005)	(0.013)	(0.011)	(0.014)	(0.013)
Behavior	0.024***	0.034***	0.045***	0.130***	0.070***	0.073***	0.071***
	(0.002)	(0.003)	(0.004)	(0.012)	(0.009)	(0.014)	(0.014)
Base Conditional Probability	0.970	0.955	0.923	0.626	0.837	0.793	0.792

Note: The effect of the factors are calculated over all race and gender groups. Each column conditions on the sample of individuals that have completed the previous educational level. ***, ** and * indicate significance at 1%, 5% and 10% level, respectively.

attempts, the completion probability is constructed as the cumulative probability of completing the grade in three attempts or less.[42] The average marginal effect is obtained by increasing each factor by one standard deviation, holding the other factors fixed, and then computing the change in the average (conditional) probability of being promoted to the next schooling level.

Table 8 shows that the baseline probability of completing each secondary education grade is much higher than the probabilities of completing each post-secondary education grade. Consequently, while the factors are all statistically significant, during high school the magnitude of the effects is smaller than during college. In terms of importance, during high school, the behavior factor has the strongest effect on grade completion. For example, increasing the behavior factor by one standard deviation for 12th graders increases the average probability of completing the 12th grade by 4.5 percentage points. The family factor consistently has the second largest effect in high school, and the math/verbal factor is the least impactful during high school.

The effect of the factors are significantly amplified in the transition from high school to college, and at this point, the math/verbal factor becomes slightly dominant over the other factors. For example, the baseline probability of completing the first year of college for those who graduate

[42] Since we do not model choices beyond age 25, we choose three attempts or less so that the marginal effects computed for grade 16 completion would be consistent with the other grade levels.

high school is about 62%. Increasing the math/verbal factor by one standard deviation increases the college enrollment probability by 16.4 percentage points, a substantial 26.1% increase over the baseline probability. Similar increases in the behavior and family factors have comparable effects, increasing the college enrollment probability by 13 and 10.7 percentage points, respectively. Beyond the first year of college, the factors continue to play a more appreciable role than in high school, with math/verbal being the leading factor followed by the behavior factor and the family factor.

4.2.2 Differential Factor Responses by Race and Gender

Table 9 summarizes the results pertaining to differential responses to the factors by race/gender. While we estimate a model of individual grade transitions, to make the table tractable we aggregate the average marginal effects of the factors across three education milestones: high school completion, college enrollment and college graduation. Column (1) shows the average marginal effect, for females in each race, of a one standard deviation increase in each of the factors (holding the other factors fixed) on the probability of completing high school (conditional of completing the ninth grade).[43] Column (2) shows the differential response of males relative to females for each race. Interestingly, while Table 8 shows that the behavior factor is the dominant factor during the high school years, Column (2) of Table 9 indicates that males are significantly more responsive to the behavioral factor than females across all races. The biggest difference is among black youth, where a one standard deviation increase in the behavior factor leads on average to a 16.5 percentage point increase in the probability of completing high school for black males, compared to 10.3 percentage points for black females. In the high school years, black males are also substantially more responsive to family background characteristics than black females, where the average marginal effect of the family factor for black males is 70% larger than the effect for females.

Columns (3) and (4) look at the average marginal effect of the factors on the probability of completing the first year of college conditional on the population graduating high school. These columns are directly analogous to the grade 13 column in Table 8 except now the effects are disaggregated to the race/gender group. Table 8 showed that the dominant overall factor for

[43] Conditional on completing ninth grade, the gender gaps in high school graduation are smaller than the unconditional gaps reported in Table 1.

Table 9: Average Marginal Effects of Factors on Schooling Level, By Race and Gender

	Graduating High School (conditional on completing 9th grade)		Enrolling in College (conditional on completing high school)		Graduating College (conditional on enrolling in college)	
	Female	Gender Diff.	Female	Gender Diff.	Female	Gender Diff.
	(1)	(2)	(3)	(4)	(5)	(6)
Black						
Family	0.088***	0.061**	0.069**	0.130***	0.051	-0.015
	(0.019)	(0.026)	(0.035)	(0.049)	(0.055)	(0.062)
Math/Verbal	0.056**	0.011	0.160***	-0.049	0.176***	0.003
	(0.024)	(0.036)	(0.025)	(0.048)	(0.048)	(0.074)
Behavior	0.103***	0.062**	0.090***	0.008	0.137***	-0.043
	(0.015)	(0.030)	(0.020)	(0.044)	(0.036)	(0.060)
Base Conditional Probability	0.830	-0.088	0.699	-0.134	0.471	-0.110
White						
Family	0.051***	0.014	0.086***	0.030	0.069***	0.015
	(0.011)	(0.016)	(0.020)	(0.029)	(0.023)	(0.035)
Math/Verbal	0.054***	-0.030**	0.170***	-0.029	0.220***	-0.038
	(0.008)	(0.015)	(0.013)	(0.020)	(0.017)	(0.030)
Behavior	0.067***	0.026**	0.131***	-0.004	0.151***	-0.035
	(0.007)	(0.010)	(0.013)	(0.020)	(0.023)	(0.029)
Base Conditional Probability	0.896	-0.018	0.750	-0.071	0.682	-0.074
Hispanic						
Family	0.050*	0.036	0.113***	-0.060	0.058	-0.003
	(0.028)	(0.035)	(0.039)	(0.049)	(0.053)	(0.084)
Math/Verbal	0.082***	0.003	0.134***	0.043	0.181***	-0.020
	(0.025)	(0.034)	(0.047)	(0.060)	(0.044)	(0.085)
Behavior	0.083***	0.063***	0.106***	0.022	0.074**	-0.026
	(0.015)	(0.022)	(0.029)	(0.037)	(0.037)	(0.060)
Base Conditional Probability	0.824	-0.047	0.633	-0.057	0.459	-0.103

Note: Gender difference shows the heterogeneous effects by denoting the differential effect of the factor on males. Notice that the gender gaps reported in the last row of each panel are conditional on a given educational attainment. ***, ** and * indicate significance at 1%, 5% and 10% level, respectively.

college enrollment was the math/verbal factor. Columns (3) and (4) of Table 9 illustrate a similar pattern for each of the race/gender groups except for black males. For black males, the most influential factor on the college enrollment decision is the family factor, whose effect is nearly 80% larger than the effect of the math/verbal factor. This result suggests a substantial gender gap in the the response to the family factor for black youth, where a one standard deviation increase in the family factor increases the conditional probability of completing the first year of college by nearly 20 percentage points for black males compared to a response of only 6.9 percentage points for black females.

The final two columns of Table 9 show the average marginal effect of the factors on the probability of graduating college, conditional on completing the first year of college. Column (5) shows the effects of the factors for females in each race. These effects are consistent with the overall patterns in Table 8, where math/verbal is the dominate factor for college completion, followed by the behavior factor and the family factor. Column (6) shows the differential effects of the factors for males relative to females for each race. This column shows no significant gender differences in the factor responses, which suggests that, for the those students reaching higher levels of education, the effect of the factors is more homogeneous.

Altogether, our analysis of the differential responses to the factors by race/gender documents that, for the high school years, males are much more responsive to the behavioral factor than females across all races. In addition, for high school completion and college enrollment, black males are substantially more responsive to the family factor than black females, while gender differences in the response to the family factor for white and Hispanic youth are not statistically significant. Each of these findings will play a key role in our analysis of the gender gap in educational attainment in Section 5. Our finding that black males are more responsive to family characteristics and environment than black females is partially consistent with previous findings in the literature, e.g., Chetty et al. (2016) and Autor et al. (2016). In particular, Autor et al. (2016) find that, on the high school completion margin, males across all races are more sensitive to changes in socio-economic status relative to females within race, while our results suggest that the differences in the effects are localized to only black males, with no significant gender differences for the family factor for white or Hispanic youth.[44] Overall, however, our results are in agreement with the main conclusion in Autor et al. (2016) which indicates that one explanation for the excessive gender gap in high school completion for black youth, compared to the gender gap for white youth, stems from the fact that black males are more sensitive to family background characteristics, couple with the additional fact that black males and females are more likely to come from families with lower socio-economic status. However, since Autor et al. (2016) only focus on the high school graduation gap, our results extend their findings by showing that the differential responses for black males and females to the family factor not only persist but are also magnified in higher education, even after conditioning on high school completion.

4.3 The Role of Arrests

Before proceeding to our analysis of the gender gaps in educational attainment and arrests in Sections 5 and 6, we conclude this section by summarizing the role of the factors in determining arrest outcomes and then, the role of arrests on schooling progression once we control for the factors.

First, Table 10 summarizes the effect of each factor on the contemporaneous probability of being arrested, i.e. being arrested in period t. The average marginal effect is calculated by increasing each factor one standard deviation, holding the other factors fixed, and comparing the change in the fraction arrested with the baseline probability. In the data, about 4.7% of individuals are arrested in each academic year. Not surprisingly, our results indicate that the behavior factor has the largest effect on arrests such that increasing the behavior factor by one standard deviation reduces the average probability of being arrested in any period by 75% in the population. Once we control for the behavior factor, the family and math/verbal factor do not have statistically significant effects on arrests.[45]

Next, as expressed in Eq. (4), our model allows arrest outcomes to have a direct effect on schooling decisions, which, given the significantly higher arrest rates for males, in particular black

[44] The main findings of Autor et al. (2016) only allow for heterogeneous effects of family background characteristics on gender, but not on gender and race. They argue that triple interactions (i.e. gender-race-family background) are not statistically significant for all racial groups.

[45] In specifications that do not control for the behavior factor, math/verbal shows a large and negative effect on arrests.

Table 10: Average Marginal Effects of Factors on Arrest in Period t

Family	-0.008
	(0.007)
Math/Verbal	0.014
	(0.009)
Behavior	-0.035***
	(0.002)
Base Probability	0.047

Note: ***, ** and * indicate significance at 1%, 5% and 10% level, respectively.

males, may contribute to the differences in the eductional decisions of males and females. Table 11 shows the average marginal effect of arrest on completion of each grade level. The bottom part of the table reports the baseline fraction of those completing each grade, conditional on completing the previous grade. Similar to Table 8, the probability of completing a grade is constructed as the cumulative probability of completing the grade in three attempts or less, which allows us to aggregate the marginal effects over first-time and repeat grade attempts. The average marginal effect of arrest is calculated by computing the probability of grade completion when the arrest variable is set to one for all individuals and comparing this to the completion probability when the arrest variable is set to zero for all individuals. Panel A presents the effect of arrest from a reduced form model that does not include controls for the factors. Without controlling for the factors, arrests appear to have a large and significant effect on grade completion throughout the schooling career, with the largest effect for completing the 12th grade.

Panel B of the table shows that once we include controls for the factors in these regressions, the effect of arrests on educational progression becomes statistically insignificant for almost all schooling levels. The only grade level where arrests remain statistically significant is the 12th grade, where being arrested reduces the probability of completing the 12th grade by 6.7 percentage points. However, this effect is nearly 70% smaller than the reduced form effect at the 12th grade. Overall, these findings indicate that, after accounting for the factors, arrests play a significantly muted role in shaping schooling progression. To provide a final assessment on how arrests affect

Table 11: Average Marginal Effects of Arrest on Grade Completion

	10th	11th	12th	13th	14th	15th	16th
Panel A: Reduced Form, No Factor Controls							
Arrest	-0.122***	-0.128***	-0.224***	-0.150***	-0.093**	-0.047	-0.155*
	(0.018)	(0.020)	(0.027)	(0.032)	(0.039)	(0.058)	(0.089)
Panel B: Controlling for the Factors							
Arrest	-0.007	-0.012	-0.067***	-0.023	-0.016	0.009	-0.109
	(0.006)	(0.009)	(0.017)	(0.023)	(0.029)	(0.044)	(0.083)
Base Conditional Probability	0.970	0.955	0.923	0.626	0.837	0.793	0.792

Note: Logistic regressions. Dependent variable is an indicator determining completion of a given school level. Arrest refers to being arrested in the current school year. Coefficients correspond to average marginal effects. Each regression conditions on the sample of individuals who completed the previous school level. Additional controls include race/gender group, number of years since completed 9th grade, and whether the student completed a year of schooling in the pevious period. ***, ** and * indicate significance at 1%, 5% and 10% level, respectively.

educational outcomes, the following section performs counterfactual simulations where we study how the schooling path of those who have been arrested at least once would have changed in the absence of their arrest.

5 Understanding the Race/Gender Gaps in Education

Using the estimates from our model, we are now able to investigate the main drivers of the observed gender gaps in educational attainment documented in Table 1. To understand the relative importance of each feature of our model, we use our estimates to simulate a series of counterfactual scenarios that elucidate the importance of gender differences in the factors, gender differences in the response to the factors, and gender differences in arrest outcomes in determining each of the observed educational gender gaps across race.

To study the role of gender differences in the factor means, which are documented in Table 5, Table 12 shows the counterfactual education gaps when we equalize each of the factor means for males to the female factor mean in each race. We study counterfactual changes in the gap at three levels of educational attainment: high school graduation or higher, college enrollment or higher, and college graduation or higher. Column (1) of Table 12 shows the unconditional fraction of females in each race having completed each of these education milestones prior to age 25. Column (2) reports the baseline race-specific gender gap for each level of education. For example, while 51.2% of all black females enrolled in college, black males were 17 percentage points less likely to enroll in college than black females. As Column (2) indicates, males across all races are less likely to complete each education level, with all differences being statistically different from zero.

Columns (3) through (6) of Table 12 present counterfactual gender gaps at the different schooling levels when we equalize the means of certain factors for males and females within racial group. Importantly, because we are looking at the unconditional gender gap, this analysis accounts for the cumulative effect of the counterfactual changes in the factors. For example, improving the math/verbal factor will have a direct effect on college enrollment and will indirectly affect college enrollment through increased high school graduation. This analysis reports the total combined effect.[46] Column (3) shows the counterfactual gender gap for each level of education when we set the mean of the family factor for males equal to the mean of the family factor of females for each race, holding the rest of the factors fixed. Not surprisingly, since the averages of the family factor between males and females are so close to begin with, equalizing the family factor means has no effect on the gender gaps, with each gap in Column (3) being nearly identical to the one reported in the baseline column. Column (4) shows the gender gap when we repeat the same analysis for the math/verbal factor, holding the other factors fixed. Equalizing the math/verbal factor has a significant effect on the gender gap: it reduces the college enrollment gap for black youth from 17 percentage points to 14.3 percentage points, a 15% reduction in the gap. Although there are significant differences in the math/verbal factor between males and females and this factor plays a significant role on education decisions, equalizing the math/verbal factor mean across genders is only able to account for between 15% to 35% of the respective gender gaps across race.

Column (5) of Table 12 shows counterfactual gender gaps when we equalize the behavior factor for males and females within race, holding the other factors fixed. Equalizing the behavior factor between males and females has a substantial effect on each of the gaps, in particular gender dif-

[46] Table 9 reports only the conditional effects.

Table 12: Counterfactual Educational Attainment at Age 24

	Female Fraction	Baseline Gap	Counterfactual Gender Gap Equalizing Factor Mean Male to Factor Mean Female			
			Family Only	Math/ Verbal Only	Behavior Only	All Factors
	(1)	(2)	(3)	(4)	(5)	(6)
Black						
HS Grad. and Higher	0.790	-0.116*** (0.027)	-0.113*** (0.024)	-0.096*** (0.025)	0.022 (0.029)	0.040 (0.028)
College Enroll and Higher	0.512	-0.170*** (0.020)	-0.167*** (0.018)	-0.143*** (0.019)	-0.084*** (0.032)	-0.050 (0.030)
College Grad. and Higher	0.170	-0.089*** (0.014)	-0.089*** (0.014)	-0.069*** (0.016)	-0.059*** (0.017)	-0.031* (0.019)
White						
HS Grad. and Higher	0.866	-0.026** (0.011)	-0.028*** (0.011)	-0.020* (0.011)	0.033*** (0.012)	0.035*** (0.013)
College Enroll and Higher	0.615	-0.090*** (0.016)	-0.094*** (0.014)	-0.060*** (0.016)	-0.025 (0.019)	0.002 (0.017)
College Grad. and Higher	0.342	-0.100*** (0.014)	-0.102*** (0.013)	-0.066*** (0.014)	-0.053*** (0.016)	-0.020 (0.014)
Hispanic						
HS Grad. and Higher	0.794	-0.064*** (0.020)	-0.063*** (0.021)	-0.047** (0.021)	0.038* (0.022)	0.052** (0.020)
College Enroll and Higher	0.460	-0.077*** (0.023)	-0.076*** (0.023)	-0.049* (0.027)	0.011 (0.026)	0.040 (0.025)
College Grad. and Higher	0.144	-0.051*** (0.013)	-0.050*** (0.013)	-0.037** (0.015)	-0.027 (0.019)	-0.011 (0.019)

Note: Female fraction denotes the proportion of females completing a given schooling level. Baseline gap reports the gender gap at that specific educational level. Columns (3) to (6) report how the gender gap in educational attainment would change if males of a given racial group were given the factor means of their female counterparts. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

ferences in high school graduation and college enrollment. For example, the gender gap in high school completion is fully explained for all racial groups once behavior differences between males and females are eliminated. In a similar vein, 51% of the gap in college enrollment among black youth is eliminated when the mean of the behavior factor for black males is equalized to the mean of the behavior factor for black females. Moreover, the gender gap in college enrollment for white youth and Hispanic youth are no longer statistically significant once we control for mean differences in the behavior factor. Finally, the last column of the table shows the counterfactual gender gap when all of the factor means for males are equalized to the female factor means. We find that, for almost all levels of the educational career, gender differences in educational attainment are no longer statistically significant or in some cases even reverted. The only exception is college completion for black males, who are still 3.1 percentage points less likely to graduate from college than black females. These counterfactuals show that nearly all of the gender differences in education can be explained by gender differences in the factors, with differences in the behavior factor being the primary component.

In the next set of counterfactuals, we focus exclusively on the gender and racial gaps in college enrollment as we set the factor means for each racial/gender group equal to the factor means of white females, who have the highest endowments for each of the factors.[47] Because we are equalizing all of the factor means for each race/gender group to a common value, this analysis allows us to study how differences in the factor means contribute to both gender and racial differences in educational decisions. Column (1) of the table focuses on the gender gap in college enrollment for black youth and shows how the gender gap changes as the factor means of both black males and black females are equalized to the means of white females. The bottom row of the table reports the original baseline gap, which is 17 percentage points. Improving the family factor of both black males and females, holding the other factors fixed, has a substantial effect on the gender gap, reducing it by 55%. This result highlights one of our key findings: black males are much more responsive to improvements in family/environment than black females.

Column (3) turns to an analysis of the black-white racial gap in college enrollment among

[47]Although white males have on average higher values for the family factor, the difference is not statistically significant.

Table 13: Counterfactual Gender/Race Gap in College Enrollment

	Equalizing Factor Mean of Each Race/Gender Group to Factor Mean of White Females					
	Male v. Female		Male		Female	
	Black	Hispanic	Black v. White	Hisp. v. White	Black v. White	Hisp. v. White
	(1)	(2)	(3)	(4)	(5)	(6)
Family Only	-0.075*	-0.118**	0.005	-0.074**	-0.014	-0.050
	(0.043)	(0.046)	(0.038)	(0.035)	(0.033)	(0.042)
Math/Verbal Only	-0.200***	-0.036	-0.112***	-0.018	0.029	-0.042
	(0.042)	(0.062)	(0.037)	(0.042)	(0.033)	(0.040)
Behavior Only	-0.086**	0.008	-0.129***	-0.130***	-0.069***	-0.163***
	(0.038)	(0.024)	(0.038)	(0.024)	(0.022)	(0.019)
All Factors	-0.010	0.016	0.136***	0.066***	0.148***	0.052**
	(0.033)	(0.036)	(0.029)	(0.025)	(0.021)	(0.024)
Baseline Gap	-0.170***	-0.077***	-0.183***	-0.141***	-0.103***	-0.155***
	(0.020)	(0.023)	(0.020)	(0.019)	(0.020)	(0.017)

Note: Columns (1) and (2) report the counterfactual gender gap in college enrollment if males of a given racial group were given the factor means of white females. Columns (3) and (4) report the counterfactual racial gap if males in each racial group were given the factor means of white females. Columns (5) and (6) report the counterfactual racial gap if females in each racial group were given the factor means of white females. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

males. Black males are 18.3 percentage points less likely to enroll in college compared to white males. However, this difference is eliminated entirely when the mean of the family factor for both populations is equalized to the mean of the white female family factor, holding the levels of the other factors fixed. Column (5) focuses on the black-white racial gap among females and also demonstrates that the gap is eliminated when the family factor is equalized. In sum, Table 12 shows that equalizing the factor means for all race/gender groups to a common set of values, i.e. the factor means of white females, can account for all race and gender gaps in college enrollment.

Finally, to summarize the role of differential arrest rates on the gender gap in educational attainment, we study how educational outcomes would change in a counterfactual scenario where arrests have no impact on educational decisions, i.e. we set the coefficient on arrest in the schooling equation to zero. To implement this analysis, we first simulate data under the full model. Columns (1), (3) and (5) in Table 14 show the distribution of education outcomes by race and gender for those experiencing an arrest in the simulation. These education shares are nearly identical with the actual averages in Table 2. Columns (2), (4) and (6) show the change in the education shares when we re-simulate the data for these individuals and set the coefficient on arrests in the schooling equation to zero. These results show that arrests have little impact on educational outcomes. For example, while only 52.3% of black males who were arrested prior to age 25 will complete high school, this figure only increases 2.2 percentage points when the effect of arrests are removed from their decisions. From this analysis we can conclude that the below-average education profile of those being arrested can be entirely accounted for by the prevalence of the factors in both the schooling and arrest outcome equations and cannot be explained by the material act of being arrested. Furthermore, differences in arrest rates can account for little, if any, of the gender gap in educational attainment.

In summary, the counterfactual simulations shown in this section highlight three main findings. First, differences in the factor means between males and females account for most of the gender gaps in educational attainment, which is primarily explained by differences in the behavior factor and the remainder by differences in the math/verbal factor. Second, black males have a substantially stronger response to the family factor compared to black females and the population in general,

Table 14: Counterfactual Age 24 Educational Attainment With No Arrests

	Black		White		Hispanic	
	Arrested Prior to Age 25	Increase if Never Arrested	Arrested Prior to Age 25	Increase if Never Arrested	Arrested Prior to Age 25	Increase if Never Arrested
	(1)	(2)	(3)	(4)	(5)	(6)
Males						
HS Grad. and Higher	0.523	0.022	0.678	0.020	0.563	0.023
College Enroll and Higher	0.229	0.010	0.325	0.008	0.239	0.011
College Grad. and Higher	0.042	0.002	0.110	0.004	0.048	0.003
Females						
HS Grad. and Higher	0.587	0.016	0.665	0.016	0.618	0.019
College Enroll and Higher	0.321	0.009	0.342	0.006	0.269	0.009
College Grad. and Higher	0.071	0.003	0.137	0.003	0.073	0.002

Note: Columns (1), (3) and (5) show the distribution of education outcomes by race and gender for those experiencing an arrest in the simulation. Columns (2), (4) and (6) show the change in the education level shares when we re-simulate the data for these individuals and set the coefficient on arrests in the schooling equation to zero.

such that increasing the family factor for black youth to match the family factor for white youth significantly reduces the gender gap in college enrollment among black youth and eliminates the black-white racial gap among males. Finally, once accounting for the family, math/verbal and behavior factors, arrests have little influence on educational attainment and thus have a minimal effect on educational gender gaps. This is because there is a high correlation between arrests and the behavior factor, such that arrests are due mainly to behavioral differences. We discuss this further in the next section.

6 Understanding the Race/Gender Gaps in Arrests

Although the primary reason we incorporate arrests into our analysis is to understand how differential arrest rates between males and females shape gender gaps in educational attainment, we can also use the estimates from our model to gain insight into the sources of the gender and racial gaps in arrests observed in Table 2. Table 15 studies the role of differences in factor means on the observed gender and racial gaps in the fraction arrested prior to age 25. Columns (1) to (3) focus on the gender gaps in arrests within race. The bottom row of each table shows the baseline gaps in the fraction ever arrested for each race. For example, black males are 32.8 percentage points more likely to be arrested prior to age 25 than black females. Each of the rows of the table shows the counterfactual arrest gaps when we increase each of the factor means of the target group to match the factor means of the comparison group, in the gender analysis, the factor means of black males are increased to match the factor means of black females. Columns (1) to (3) show that differences in the mean of the behavior factor between males and females within race account for more than half of the gender gap in arrest for each race, although even after accounting for differences in all of the factors males are still statistically more likely to be arrested.

Columns (4) and (5) analyze the racial gaps in arrest rates among males, with columns (6) and (7) focusing on females. Looking at the baseline probabilities, columns (4) through (7) show that the arrest gaps between Hispanic males and white males, black females and white females, and Hispanic females and white females, are not statistically significant even without controlling for the factors. However, the racial gap between black males and white males is large and statistically

Table 15: Counterfactual Gender/Race Gap in Fraction Ever Arrested Prior to Age 24

	Equalizing Factor Mean Male to Factor Mean Female			Equalizing Factor Mean Non-White to Factor Mean White			
	Male v. Female			Male		Female	
	Black	White	Hispanic	Black v. White	Hisp. v. White	Black v. White	Hisp. v. White
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Family Only	0.327***	0.182***	0.231***	0.095***	-0.015	-0.029	-0.046*
	(0.020)	(0.013)	(0.022)	(0.035)	(0.035)	(0.026)	(0.024)
Math/Verbal Only	0.349***	0.194***	0.243***	0.199***	0.069**	0.025	-0.001
	(0.026)	(0.013)	(0.023)	(0.040)	(0.035)	(0.028)	(0.028)
Behavior Only	0.113***	0.068***	0.080**	-0.023	0.007	-0.049**	-0.015
	(0.030)	(0.021)	(0.032)	(0.031)	(0.020)	(0.020)	(0.024)
All Factors	0.130***	0.080***	0.089***	-0.009	0.010	-0.044*	-0.011
	(0.036)	(0.024)	(0.034)	(0.036)	(0.029)	(0.023)	(0.030)
Baseline Gap	0.328***	0.181***	0.232***	0.140***	0.024	-0.006	-0.027
	(0.021)	(0.013)	(0.022)	(0.017)	(0.019)	(0.018)	(0.018)

Note: Columns (1) through (3) report the counterfactual gender gap in arrests if males of a given racial group were given the factor means of their female counterparts. Columns (4) and (5) report the counterfactual racial gap in arrests if males in each racial group were given the factor means of white males. Columns (6) and (7) report the counterfactual racial gap in arrests if females in each racial group were given the factor means of white females. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

significant, with black males being 14 percentage points more likely to be arrested prior to age 25 than white males. Column (4) shows the counterfactual black-white racial gap in arrests among males when we equalize the factor means for black males to the factor means of white males. This analysis shows that differences in the behavior factor account for the entirety of the black-white racial gap in arrests among males, where the counterfactual racial gap falls from 14 to -2.3 percentage points when the behavior factor is equalized.

In summary, given the predominant influence of the behavior factor on arrest outcomes, differences in the behavior factor among males and females account for more than half of the observed gender differences in arrest rates. Furthermore, differences in the behavior factor among black males and white males fully explain racial disparities in arrests.

7 Conclusion

Understanding the causes of gender disparities in educational attainment has been and continuous to be of central interest to researchers. Given the substantial variation in the gender gap across race, and recognizing that different mechanisms documented in the literature may operate differently across races, this paper extends the literature by examining which mechanisms are most influential to the different gender gaps within race. We develop a novel empirical framework to study this question that combines a sequential schooling and arrest model with a multi-factor model of student characteristics. This approach allows us to separate the importance of three distinct mechanisms relevant for the gender gap in educational attainment: First, the role of different student-level characteristics between males and females. Second, the role of differential responses to student-level characteristics between males and females; and third, the role of differential arrest rates between males and females. Furthermore, by modeling decisions at each grade-level transition, we examine how the roles of each of these three mechanisms evolve over the schooling career.

The factor aspect of our model uses 59 measures of early student outcomes to recover three latent student-level characteristics: a family background factor, a math/verbal factor, and a normative/externalizing behavior factor. Our findings indicate that equalizing the mean of the behavior factor between males and females within racial group would completely explain the gender gap in high school graduation, and it would substantially reduce the gap in postsecondary participation. For example, the large gender gap in college enrollment between black males and females would be reduced by 51%. A similar counterfactual analysis for math/verbal skills leads to a reduction in the gap on the order of 15% to 35%, depending on the racial group. Moreover, we show that the large arrests rates of males, and in particular black males, do not play a relevant role in explaining the gender gap in educational attainment. Finally, we present evidence showing that family background characteristics have a disproportionately larger effect on black males than black females on high school completion and college enrollment. In summary, once all these elements are taken together, our model can fully explain the gender gap in educational attainment.

The findings of this paper suggest that initiatives aiming to foster normative/externalizing behavior of males, and in particular on black males, could effectively reduce current inequalities in educational attainment. In addition, the fact that black males are more responsive to family background characteristics than their female counterparts when considering educational decisions indicates that more disadvantaged household environments exacerbate the gender gap in educational attainment. Finally, our findings show that racial disparities in educational attainment are mainly driven by the outcomes of black males, therefore policies that focus on keeping them in the education system will largely contribute to reduce the black-white gap.

References

D. Acemoglu and D. Autor. Skills, tasks and technologies: Implications for employment and earnings. *Handbook of Labor Economics*, 4:1043–1171, 2011.

M. Almlund, A. L. Duckworth, J. J. Heckman, and T. D. Kautz. Personality psychology and economics. Technical report, National Bureau of Economic Research, 2011.

E. M. Aucejo and J. James. The path to college education: Are verbal skills more important than math skills? 2015.

D. Autor and M. Wasserman. Wayward sons: The emerging gender gap in labor markets and education. *Third Way Report*, 2013.

D. Autor, D. Figlio, K. Karbownik, J. Roth, M. Wasserman, et al. Family disadvantage and the gender gap in behavioral and educational outcomes. *Working Paper*, (21936), 2016.

J. G. Bernburg and M. D. Krohn. Labeling, life chances, and adult crime: The direct and indirect effects of official intervention in adolescence on crime in early adulthood. *Criminology*, 41(4): 1287–1318, 2003.

M. Bertrand and J. Pan. The trouble with boys: Social influences and the gender gap in disruptive behavior. *American Economic Journal: Applied Economics*, 5:32–64, 2013.

L. Borghans, B. H. Golsteyn, J. J. Heckman, and J. E. Humphries. What do grades and achievement tests measure? 2015.

R. Brame, S. D. Bushway, R. Paternoster, and M. G. Turner. Demographic patterns of cumulative arrest prevalence by ages 18 and 23. *Crime & Delinquency*, 60(3):471–486, 2014.

L. Broidy, D. Nagin, R. Tremblay, J. Bates, B. Brame, K. Dodge, and D. Fergusson. Developmental trajectories of childhood disruptive behaviors and adolescent delinquency: A six-site, cross-national study. *Developmental Psychology*, 39(2):222–245, 2003.

C. Buchmann and T. A. DiPrete. The growing female advantage in college completion: The role of family background and academic achievement. *American Sociological Review*, 71(4):515–541, 2006.

S. V. Cameron and J. J. Heckman. The dynamics of educational attainment for black, hispanic, and white males. *Journal of Political Economy*, 109(3):455–499, 2001.

P. Carneiro, J. Heckman, and D. Masterov. Labor market discrimination and racial differences in premarket factors. *Journal of Law and Economics*, 48(1):1–39, 2005.

P. A. Cavazos-Rehg, M. J. Krauss, E. L. Spitznagel, M. Schootman, K. K. Bucholz, J. F. Peipert, V. Sanders-Thompson, L. B. Cottler, and L. J. Bierut. Age of sexual debut among u.s. adolescents. *Contraception*, 80(2):158–162, 2009.

J. Cawley, J. Heckman, and E. Vytlacil. Three observations on wages and measured cognitive ability. *Journal of Political Economy*, 8(4):419–442, 2001.

R. Chetty, N. Hendren, F. Lin, J. Majerovitz, and B. Scuderi. Childhood environment and gender gaps in adulthood. *Working Paper*, (21936), 2016.

J. Cooley, S. Navarro, and Y. Takahashi. How the timing of grade retention affects outcomes: Identification and estimation of time-varying treatment effects. 2011.

K. Deater-Deckard, K. Dodge, J. Bates, and G. Pettit. Multiple risk factors in the development of externalizing behavior problems: Group and individual differences. *Development and Psychopathology*, 10(03):469–493, 1998.

T. A. DiPrete and J. L. Jennings. Social and behavioral skills and the gender gap in early educational achievement. *Social Science Research*, 41(1):1–15, 2012.

N. M. Fortin, P. Oreopoulos, and S. Phipps. Leaving boys behind: Gender disparities in high academic achievement. *Journal of Human Resources*, 50(3):549–579, 2015.

R. G. Fryer Jr. An empirical analysis of racial differences in police use of force. Technical report, National Bureau of Economic Research, 2016.

C. Goldin, L. F. Katz, and I. Kuziemko. The homecoming of american college women: The reversal of the college gender gap. *The Journal of Economic Perspectives*, 20(4):133–156, 2006.

R. Goodman. The strengths and difficulties questionnaire: a research note. *Journal of Child Psychology and Psychiatry*, 38(5):581–586, 1997.

J. Heckman, J. Stixrud, and S. Urzua. The effects of cognitive and non-cognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics*, 24(3):411–482, 2006.

J. Heckman, R. Pinto, and P. Savelyev. Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. *American Economic Review*, 103(6):1–35, 2013.

J. J. Heckman. The american family in black & white: A post-racial strategy for improving skills to promote equality. *Daedalus*, 140(2):70–89, 2011.

J. J. Heckman, J. E. Humphries, and G. Veramendi. Web appendix for estimating the causal effects of schooling. 2015.

L. R. Huesmann, L. D. Eron, and E. F. Dubow. Childhood predictors of adult criminality: are all risk factors reflected in childhood aggressiveness? *Criminal Behaviour and Mental Health*, 12(3): 185–208, 2002.

B. A. Jacob. Where the boys aren't: Non-cognitive skills, returns to school and the gender gap in higher education. *Economics of Education review*, 21(6):589–598, 2002.

J. James. MM algorithm for general mixed multinomial logit models. *Journal of Applied Econometrics*, 2016.

J. Lavigne, R. Gibbons, K. Christofel, R. Arend, H. Binns, N. Dawson, H. Sobel, and C. Isaacs. Prevalence rates and correlates of psychiatric disorders among preschool children. *Journal of the American Academy of Child and Adolescent Psychiatry*, 35:204–214, 1996.

R. Lerner and L. Steinberg. *Handbook of Adolescent Psychology, Volume 1: Individual Bases of Adolescent Development*, volume 1. John Wiley & Sons, 2009.

S. Lundberg and R. A. Pollak. Cohabitation and the uneven retreat from marriage in the united states, 1950–2010. In *Human capital in history: The American record*, pages 241–272. University of Chicago Press, 2014.

A. McDaniel, T. A. DiPrete, C. Buchmann, and U. Shwed. The black gender gap in educational attainment: Historical trends and racial comparisons. *Demography*, 48(3):889–914, 2011.

D. Neal and W. Johnson. The role of premarket factors in black-white wage differences. *Journal of Political Economy*, 104(5):869–895, 1996.

M. Ree and T. Carretta. Group differences in aptitude factor structure on the asvab. *Educational and Psychological Measurement*, 55(2):268–277, 1995.

H.-L. Schofield, K. Bierman, B. Heinrichs, R. Nix, C. P. P. R. Group, et al. Predicting early sexual activity with behavior problems exhibited at school entry and in early adolescence. *Journal of Abnormal Child Psychology*, 36(8):1175–1188, 2008.

C. Segal. Misbehavior, education, and labor market outcomes. *Journal of the European Economic Association*, 11(4):743–779, 2013.

G. Sweeten. Who will graduate? Disruption of high school education by arrest and court involvement. *Justice Quarterly*, 23(4):462–480, 2006.

R. Thompson, J. K. Tabone, A. Litrownik, E. Briggs, J. Hussey, D. English, and H. Dubowitz. Early adolescent risk behavior outcomes of childhood externalizing behavioral trajectories. *The Journal of Early Adolescence*, 2010.

F. Vitaro, M. Brendgen, S. Larose, and R. E. Trembaly. Kindergarten disruptive behaviors, protective factors, and educational achievement by early adulthood. *Journal of educational psychology*, 97(4):617, 2005.

Appendix

A Reduced Form Gap

Table 16: Unexplained Gender Gap in College Enrollment (Before age 25)

Model Controls	Black	White	Hispanic
Time only	-0.166	-0.105	-0.079
Time and disruption only	-0.167	-0.099	-0.077
Time, disruption, and arrest only	-0.171	-0.099	-0.079

Note: This table shows the simulated gender gap in college enrollment by race from our sequential schooling model without controls for the factors. These are the results from three reduced form models. The first includes only controls for time since completing the ninth grade as a control in the sequential schooling decision. The second adds controls for an impact of schooling disruption on the schooling decision. The third additionally accounts for the impact of being arrested on the schooling decision. This table shows that these controls by themselves do little to explain the gaps.

B Measures

We incorporate a wide variety of data from the NLSY97 to identify the family factor, the math/verbal factor, and the behavioral factor. We only include measurement data that are determined prior to the beginning of high school as these variables are further removed from later in life education decisions. Table 17 provides a complete list of the measures used in our analysis. For data that are grade dependent, rather than make unnecessary assumptions to normalize or aggregate the measures, we estimate the measurement equation parameters at the grade level. For example, we use a question on mother's parenting style to identify our family factor. This question is asked in each round of the survey. Rather than aggregating these data, we incorporate all of the data by using parenting style in grade 6, parenting style in grade 7, and parenting style in grade 8 as three separate measures. This approach allows the family factor to influence parenting style depending on the grade of the youth, accounting for the fact that parenting styles may change as youth get older.

Other measures, like number of household members under the age of 18 do not vary significantly across grade level, so for these variables we compute averages.

Because we only use information prior to the end of middle school, due to the sampling design of the NLSY97, some of the variables have missing values for the oldest individuals in the sample. For example, the PIAT exam was only given to respondents in Round 1 who were in the 9th grade or lower. Under our research design, we only use exam scores for the sixth, seventh, and eighth graders, treating the exam as missing for the other respondents. If the respondent's information is missing due to age at first interview, we treat this data as missing at random. For other variables where data is missing because the respondent either did not know the answer to the question or refused to answer the question, in order to avoid issues of selection bias, we endogenize the missing information into our model. For example, mother's degree takes values: missing, high school drop out, high school graduate, some college, or college graduate. In this way the fact that the person has missing information in this variable is informative about their family background just as it would be if the respondent had reported college graduate or any other level of education for their mother.

Finally, because our estimator allows us to easily incorporate discrete and continuous measures, for some of the variables, we constructed measures at both the extensive and intensive margin to recover more information from the data. For example, the survey asks the respondent about family income. We converted this information into two measures. The first is a discrete, extensive measure for if family income is missing, below poverty ($15K/yr), or above poverty. For the those reporting above poverty, the second variable is the log of family income, measuring the intensive margin. This approach allows the estimator to recover the family factor from observed family income in a more flexible way.

Table 17: Description of Measures

Name	Description (Values)	% Missing
Measures Loading on Family Only		
DadDegree	Biological fathers's highest degree completed (missing, high school drop out, high school graduate, some college, college graduate)	0% (ENDOG)
MomDegree	Biological mother's highest degree completed (missing, high school drop out, high school graduate, some college, college graduate)	0% (ENDOG)
brokenHH	Any household arrangement except both biological parents during grades 6 to 8 (yes, no)	7.5% (MAR)
family incomeLOG	Log of family income for those conditional on greater than 15k family income reported (continuous)	40% (ENDOG)
family income group	Grouped average family income level reported between age 10 and 17 (missing, $<$15k,$\geq$15k)	0% (ENDOG)
mom parent style 6	Mother's parenting style when youth is in 6th grade (uninvolved, permissive, authoritarian, authoritative)	91.8% (MAR)
mom parent style 7	"··············" in 7th grade	78.9% (MAR)
mom parent style 8	"··············" in 8th grade	70.4% (MAR)
num HH under18	Average number of household members under the age of 18 while youth is age 10 to 17 (continuous)	0.5% (MAR)
Measures Loading on Math/Verbal and Behavior		
GPAin8 8	Self-reported grade point average in 8th grade (continuous)	4.3% (MAR)
HW hrs wkANY 6	Any positive weekly homework hours in 6th grade (yes, no)	91.7 % (MAR)
HW hrs wkANY 7	"··············" in 7th grade	80.2% (MAR)
HW hrs wkANY 8	"··············" in 8th grade	81.2% (MAR)
HW hrs wkLOG 6	Log of hours spent per week on homework if reported positive hours in 6th grade (continuous)	92.9%
HW hrs wkLOG 7	"··············" in 7th grade	82.3%

(*Continued on next page*)

Note: This table provides a description of the measurements used to identify each of the factors. MAR denotes missing at random. ENDOG denotes modeled endogenously in the likelihood. Table 20 shows how the different factors load on each of these measures.

Table 17: Description of Measures

Name	Description (Values)	% Missing
HW hrs wkLOG 8	"..............." in 8th grade	83.4%
READ hrs wkANY 6	Any positive weekly hours spent reading for pleasure in 6th grade (yes, no)	91.6% (MAR)
READ hrs wkANY 7	"..............." in 7th grade	80.1% (MAR)
READ hrs wkANY 8	"..............." in 8th grade	81.2% (MAR)
READ hrs wkLOG 6	Log weekly hours spent reading for pleasure if reported positive hours in 6th grade (continuous)	94.7%
READ hrs wkLOG 7	"..............." in 7th grade	87.6%
READ hrs wkLOG 8	"..............." in 8th grade	88.9%
days suspendedANY 6	Any suspensions in 6th grade (yes, no)	0.4% (MAR)
days suspendedANY 7	"..............." in 7th grade	0.8% (MAR)
days suspendedANY 8	"..............." in 8th grade	1.9% (MAR)
days suspendedLOG 6	Log of total days suspended if reported any suspensions in 6th grade (continuous)	90.4%
days suspendedLOG 7	"..............." in 7th grade	86.5%
days suspendedLOG 8	"..............." in 8th grade	85.4%
retention 6	Retained in the 6th grade (yes, no)	0.4% (MAR)
retention 7	Retained in the 7th grade (yes, no)	0.8% (MAR)
retention 8	Retained in the 8th grade (yes, no)	1.9% (MAR)
Measures Loading on Math/Verbal Only		

(*Continued on next page*)

Note: This table provides a description of the measurements used to identify each of the factors. MAR denotes missing at random. ENDOG denotes modeled endogenously in the likelihood. Table 20 shows how the different factors load on each of these measures.

Table 17: Description of Measures

Name	Description (Values)	% Missing
ASVAB AR 7	ASVAB Arithmetic Reasoning score of youth in 7th grade (continuous)	93.1% (MAR)
ASVAB AR 8	"·············· " in 8th grade	83% (MAR)
ASVAB MK 7	ASVAB Mathematics Knowledge score of youth in 7th grade (continuous)	93.1% (MAR)
ASVAB MK 8	"·············· " in 8th grade	83.1% (MAR)
ASVAB PC 7	ASVAB Paragraph Comprehension score of youth in 7th grade (continuous)	93.1% (MAR)
ASVAB PC 8	"·············· " in 8th grade	83% (MAR)
ASVAB WK 7	ASVAB Word Knowledge score of youth in 7th grade (continuous)	93.1% (MAR)
ASVAB WK 8	"·············· " in 8th grade	83% (MAR)
piat 6	Peabody Individual Achievement Test Mathematics score of youth in 6th grade (continuous)	91.8% (MAR)
piat 7	"·············· " in 7th grade	79.2% (MAR)
piat 8	"·············· " in 8th grade	72.5% (MAR)
Measures Loading on Behavior Only		
alcohol 6	Youth self-report of drinking alcohol in 6th grade (yes, no)	10.1% (MAR)
alcohol 7	"·············· " in 7th grade	10.3% (MAR)
alcohol 8	"·············· " in 8th grade	10.1% (MAR)
attack 6	Youth self-report of attacking or assaulting another person in 6th grade (yes, no)	6.1% (MAR)
attack 7	"·············· " in 7th grade	5.9% (MAR)
attack 8	"·············· " in 8th grade	6% (MAR)
marij 6	Youth self-report of smoking marijuana in 6th grade (yes, no)	2.3% (MAR)

(*Continued on next page*)

Note: This table provides a description of the measurements used to identify each of the factors. MAR denotes missing at random. ENDOG denotes modeled endogenously in the likelihood. Table 20 shows how the different factors load on each of these measures.

Table 17: Description of Measures

Name	Description (Values)	% Missing
marij 7	"·············· " in 7th grade	2.7% (MAR)
marij 8	"·············· " in 8th grade	3.4% (MAR)
sex 6	Youth self-report of having sexual intercourse in 6th grade (yes, no)	3.6% (MAR)
sex 7	"·············· " in 7th grade	3.9% (MAR)
sex 8	"·············· " in 8th grade	4.7% (MAR)
smoke 6	Youth self-report of smoking cigarette in 6th grade (yes, no)	11.4% (MAR)
smoke 7	"·············· " in 7th grade	11.6% (MAR)
smoke 8	"·············· " in 8th grade	10.9% (MAR)
Measures Loading on Family, Math/Verbal, and Behavior		
InEdSamp	Individual has an observed transition from 8th grade highest grade completed to 9th grade highest grade completed. The sequential schooling and arrest model is on the 'yes' population (yes, no)	0% (ENDOG)
age6th	Age which the individual first attempted the 6th grade (continuous)	0.4% (MAR)

This table provides a description of the measurements used to identify each of the factors. MAR denotes missing at random. ENDOG denotes modeled endogenously in the likelihood. Table 20 shows how the different factors load in these measures.

C Estimation Algorithm

To illustrate our estimation algorithm, assume there are T possible outcomes and measures for each individual. Let y_{it} indicate the observed outcome t for $t \in 1, 2, \ldots, T$ of individual i for $i \in 1, 2, \ldots, n$. The data $y_i = \{y_{i1}, y_{i2}, \ldots, y_{iT}\}$ contains a mix of continuous and discrete variables. If outcome t is a continuous variable, then y_{it} is a scalar. If t is a discrete outcome, it takes one of J_t values, and y_{it} is a $J_t \times 1$ vector, i.e. $y_{it} = [y_{it1} \quad y_{it2} \quad \cdots \quad y_{itJ_m}]'$, with $y_{itj} = 1$ if i reports $j \in J_t$ and zero everywhere else. In our context, T is about 77: 59 measurement outcomes, 9 schooling outcomes and 9 arrest outcomes.

Let $\Pr(y_{it}|x_i, \beta_t, \sigma_t)$ denote the probability of observing y_{it} if t is a continuous measure and $\Pr(y_{it}|x_i, \beta_t)$ the probability if t is a discrete measures. $\mathbf{x_i}$ **contains observed data and unobserved factors.** When y is continuous, β_t is a vector of parameters. For the discrete measures, β_t is a matrix, with $(J_t - 1)$ columns, where β_{tj} indicate the jth column. To simplify the indexing, in this example we normalize the coefficients for the J_t outcome to zero, while in the body of the paper we normalize the first outcome to zero. In this framework, some of the values of β_t must be normalized, e.g. arrests in future periods do not effect schooling decisions today. Let $\mathtt{I}_t$ index all of the elements in β_t to be estimated and $!\mathtt{I}_t$ (the complement of $\mathtt{I}_t$) the elements that are not estimated and fixed in the model, e.g. to 0 or 1. The joint likelihood is $L(y_i|x_i, \beta, \sigma) = \prod_{t=1}^{T} \Pr(y_{it}|x_i, \beta_t, \sigma_t)$. Letting $f(\theta|\gamma)$ denote the probability density function of the factors, with parameters γ, the full likelihood can be written compactly as,

$$LL(\beta, \sigma, \gamma) = \sum_{i=1}^{n} \ln \left(\int L(y_i|x_i, \beta, \sigma) f(\theta|\gamma) d\theta \right)$$

Our estimation strategy uses the minorization-maximization algorithm developed in James (2016), which maximizes the log-likelihood using surrogate functions, which are linear in parameters and can be easily maximized equation by equation. Given a current guess of parameters, β^o, σ^o, and γ^o, the log-likelihood can be bounded below by a function

$$Q(\beta, \sigma, \gamma|\beta^o, \sigma^o, \gamma^o) = \sum_{i=1}^{n} \int \ln \left(L(y_i|x_i, \beta, \sigma) f(\theta|\gamma) \right) h(\theta|y_i, x_i, \beta^o, \sigma^o, \gamma^o) d\theta$$

where

$$h(\theta|y_i, x_i, \beta^o, \sigma^o, \gamma^o) = \frac{L(y_i|x_i, \beta^o, \sigma^o)f(\theta|\gamma^o)}{\int L(y_i|x_i, \beta^o, \sigma^o)f(\theta'|\gamma^o)d\theta'}$$

Drawing R values of θ from $f(\theta|\gamma^o)$ for each individual labeled θ^o_{ir}, the density $h(\theta|y_i, x_i, \beta^o, \sigma^o, \gamma^o)$ can be approximated with the weights

$$w^o_{ir} = \frac{L(y_i|x^o_{ir}, \beta^o, \sigma^o)}{\sum_{r'=1}^{R} L(y_i|x^o_{ir'}, \beta^o, \sigma^o)} \tag{8}$$

Notice x is now indexed by r because the factors enter the x vector and x is superscripted by o because it depends on the initial guess of the parameters.

Using this approximation to the density, the surrogate function becomes.

$$Q(\beta, \sigma, \gamma|\beta^o, \sigma^o, \gamma^o) = \sum_{i=1}^{n}\sum_{r=1}^{R} w^o_{ir} \ln\left(L(y_i|x^o_{ir}, \beta, \sigma)f(\theta|\gamma)\right)$$

James (2016) proposes a new surrogate function $\widetilde{Q}(\beta, \sigma, \gamma|\beta^o, \sigma^o, \gamma^o)$ that has closed form solutions for all of the parameters. The algorithm works as follows.

Initialize with starting values $\beta^o, \sigma^o, \gamma^o$. Iterate until converged
Minorization Step:

- For each $i \in 1, 2, \ldots, n$
  1) Draw R values of θ from $f(\theta|\gamma^o)$, labeled θ^o_{ir}
  2) Construct x^o_{ir} using θ_{ir}
  3) For each $t \in 1, 2, \ldots, T$

     If t is continuous
     - $SR_{irt} = \left(y_{it} - (x^o_i)'\beta_t\right)^2$
     - $P_{irt} = \dfrac{1}{\sqrt{2(\sigma^o_t)^2\pi}} \exp\left(-\dfrac{SR_{irt}}{2(\sigma^o_t)^2}\right)$

     If t is discrete
     - $P_{irtj} = \left[\prod_{j'=1}^{J_t-1} \exp\left((x^o_i)'\beta_{tj'}\right)^{(j'=j)}\right]\left[\dfrac{1}{1+\sum_{j'=1}^{J_t-1}\exp\left((x^o_i)'\beta_{tj'}\right)}\right]$
     - $P_{irt} = \prod_{j=1}^{J_t} P^{y_{itj}}_{irtj}$

4) $L_{ir} = \prod_{t=1}^{T} P_{irt}$

5) $w_{ir}^o = L_{ir} / \left(\sum_{r'=1}^{R} L_{ir'} \right)$

6) Compute $\mathbf{XX}_i = \sum_{r=1}^{R} w_{ir}^o (x_{ir}^o)(x_{ir}^o)'$

7) For each $t \in 1, 2, \ldots, T$

If t is continuous

- $SR_{it} = \sum_{r=1}^{R} w_{ir}^o SR_{irt}$
- $\mathbf{XY}_{it} = \sum_{r=1}^{R} w_{ir}^o x_{ir}^o y_{it}$

If t is discrete

- $\mathbf{XY}_{it} = \sum_{r=1}^{R} w_{ir}^o x_{ir}^o \left[y'_{it[1:J_t-1]} - P'_{irt[1:J_t-1]} \right]$

Maximization Step:

- For each $t \in 1, 2, \ldots, T$
  - Let n_t denote the set of individuals with observation on t
  - $\mathbf{XX}_t = \sum_{i \in n_t} \mathbf{XX}_i$
  - $\mathbf{XY}_t = \sum_{i \in n_t} \mathbf{XY}_{it}$

  If t is continuous

  - $\sigma_t^o = \sqrt{\sum_{i \in n_t} SR_{it}/n_t}$
  - $\beta^o_{t[\mathtt{I}_t]} = \mathbf{XX}^{-1}_{t[\mathtt{I}_t,\mathtt{I}_t]} (\mathbf{XY}_{t[\mathtt{I}_t]} - \mathbf{XX}_{t[\mathtt{I}_t,!\mathtt{I}_t]} \beta^o_{t[!\mathtt{I}_t]})$

  If t is discrete

  - $\mathbf{B_t} = -.5(eye_{J_t-1} - ones_{J_t-1}/J_t) \otimes \mathbf{XX}_t$
  - $\beta^o_{t[\mathtt{I}_t]} = \beta^o_{t[\mathtt{I}_t]} - \mathbf{B}^{-1}_{t[\mathtt{I}_t,\mathtt{I}_t]} \mathbf{XY}_{t[\mathtt{I}_t]}$

- The update for γ^o depends on the chosen distribution. The update treats θ^o_{ir} as data using the weights w^o_{ir}. For example if γ is the mean of the θ, then the update is $\gamma^o = \sum_{i=1}^{n} \sum_{r=1}^{R} w^o_{ir} \theta^o_{ir} / n$.

D Factor Estimates

Table 18: Factor Type Estimates

	Type Factor Mean			Type Shares					
				Black		White		Hispanic	
	Family	Math/ Verbal	Behavior	Male	Female	Male	Female	Male	Female
Type 1	4.376	3.475	-1.232	0.019	0.007	0.111	0.037	0.025	0.034
Type 2	3.348	2.827	-1.179	0.191	0.382	0.008	0.031	0.455	0.345
Type 3	3.768	3.271	-1.112	0.045	0.143	0.193	0.300	0.039	0.090
Type 4	3.974	3.374	-0.020	0.047	0.111	0.222	0.357	0.091	0.139
Type 5	3.480	2.782	-1.619	0.435	0.045	0.105	0.044	0.062	0.003
Type 6	3.519	2.967	0.306	0.056	0.210	0.014	0.064	0.131	0.356
Type 7	4.552	3.591	-0.005	0.001	0.023	0.089	0.125	0.011	0.007
Type 8	3.858	2.976	-1.262	0.206	0.078	0.257	0.043	0.184	0.026

Note: This table reports for each factor the 8 type means and shares that were obtained from the mixture of normals. Section 3 describes how these types are recovered.

Table 19: Factor Moments by Race and Gender

	Male			Female		
	Family	Math/ Verbal	Behavior	Family	Math/ Verbal	Behavior
Black						
Mean	3.591	2.905	-1.147	3.597	3.012	-0.661
	(0.018)	(0.050)	(0.145)	(0.016)	(0.047)	(0.093)
Variance	0.085	0.095	0.441	0.108	0.113	0.594
	(0.011)	(0.011)	(0.114)	(0.012)	(0.013)	(0.167)
Correlation						
Math/Verbal	0.446	–	–	0.560	–	–
	(0.042)			(0.033)		
Behavior	0.317	0.198	–	0.311	0.155	–
	(0.058)	(0.060)		(0.054)	(0.077)	
White						
Mean	3.935	3.209	-0.787	3.924	3.288	-0.482
	(0.011)	(0.049)	(0.108)	(0.012)	(0.044)	(0.076)
Variance	0.121	0.126	0.517	0.127	0.107	0.523
	(0.008)	(0.015)	(0.126)	(0.009)	(0.014)	(0.131)
Correlation						
Math/Verbal	0.547	–	–	0.536	–	–
	(0.028)			(0.034)		
Behavior	0.426	0.323	–	0.393	0.181	–
	(0.035)	(0.053)		(0.044)	(0.061)	
Hispanic						
Mean	3.588	2.961	-0.828	3.592	3.022	-0.430
	(0.019)	(0.047)	(0.119)	(0.020)	(0.053)	(0.086)
Variance	0.113	0.104	0.520	0.106	0.110	0.633
	(0.012)	(0.013)	(0.136)	(0.013)	(0.019)	(0.179)
Correlation						
Math/Verbal	0.521	–	–	0.575	–	–
	(0.037)			(0.032)		
Behavior	0.272	0.183	–	0.221	0.069	–
	(0.051)	(0.066)		(0.079)	(0.111)	

Note: This table reports mean, variance, and correlation of the factors by race and gender. Standard errors are reported in parentheses.

Table 20: Summary of Factor Loadings: Average Marginal Effects of Factors

Measure	Categorical Values	Average Value	Family	Math/ Verbal	Behavior
1. ASVAB AR 7	Continuous	-0.789	0	0.801*** (0.023)	0
2. ASVAB AR 8	Continuous	-0.579	0	0.821*** (0.018)	0
3. ASVAB MK 7	Continuous	-0.893	0	0.575*** (0.024)	0
4. ASVAB MK 8	Continuous	-0.485	0	0.719*** (0.016)	0
5. ASVAB PC 7	Continuous	-0.673	0	0.726*** (0.023)	0
6. ASVAB PC 8	Continuous	-0.517	0	0.736*** (0.018)	0
7. ASVAB WK 7	Continuous	-1.026	0	0.705*** (0.029)	0
8. ASVAB WK 8	Continuous	-0.837	0	0.684*** (0.023)	0
9. DadDegree	Missing	0.165	-0.116*** (0.004)	0	0
	HS Dropout	0.157	-0.108*** (0.004)	0	0
	HS Deg.	0.319	-0.091*** (0.008)	0	0
	Some Col.	0.163	0.046*** (0.009)	0	0
	Col. Deg.	0.196	0.268*** (0.008)	0	0
10. GPAin8 8	Continuous	2.866	0	0.362*** (0.019)	0.331*** (0.017)
11. HW hrs wkANY 6	Yes	0.898	0	0.030*** (0.011)	0.025** (0.012)
12. HW hrs wkANY 7	Yes	0.912	0	0.034*** (0.007)	0.027*** (0.008)
13. HW hrs wkANY 8	Yes	0.898	0	0.030*** (0.010)	0.056*** (0.010)

(*Continued on next page*)

Note: Average marginal effects correspond to the coefficients from the measurement equations described in Section 3 and Table 17. Standard errors are reported in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10% level, respectively.

Table 20: Summary of Factor Loadings: Average Marginal Effects of Factors

Measure	Categorical Values	Average Value	Family	Math/ Verbal	Behavior
14. HW hrs wkLOG 6	Continuous	1.338	0	0.004 (0.042)	0.097** (0.048)
15. HW hrs wkLOG 7	Continuous	1.400	0	0.031 (0.027)	0.214*** (0.027)
16. HW hrs wkLOG 8	Continuous	1.467	0	0.069* (0.040)	0.183*** (0.036)
17. InEdSamp	Yes	0.910	0.023*** (0.007)	0.029*** (0.008)	0.048*** (0.004)
18. MomDegree	Missing	0.064	-0.043*** (0.002)	0	0
	HS Dropout	0.168	-0.128*** (0.004)	0	0
	HS Deg.	0.351	-0.093*** (0.008)	0	0
	Some Col.	0.232	0.030*** (0.009)	0	0
	Col. Deg.	0.185	0.234*** (0.009)	0	0
19. READ hrs wkANY 6	Yes	0.672	0	0.069*** (0.017)	0.043* (0.026)
20. READ hrs wkANY 7	Yes	0.637	0	0.090*** (0.012)	0.009 (0.014)
21. READ hrs wkANY 8	Yes	0.583	0	0.051** (0.020)	0.093*** (0.018)
22. READ hrs wkLOG 6	Continuous	1.064	0	0.138* (0.073)	0.019 (0.062)
23. READ hrs wkLOG 7	Continuous	1.052	0	0.077* (0.045)	0.057 (0.046)
24. READ hrs wkLOG 8	Continuous	1.058	0	0.069 (0.057)	0.035 (0.044)
25. age6th	Continuous	11.648	-0.036 (0.025)	-0.144*** (0.028)	-0.126*** (0.013)
26. alcohol 6	Yes	0.083	0	0	-0.042*** (0.003)

(*Continued on next page*)

Note: Average marginal effects correspond to the coefficients from the measurement equations described in Section 3 and Table 17. Standard errors are reported in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10% level, respectively.

Table 20: Summary of Factor Loadings: Average Marginal Effects of Factors

Measure	Categorical Values	Average Value	Family	Math/ Verbal	Behavior
27. alcohol 7	Yes	0.101	0	0	-0.040*** (0.005)
28. alcohol 8	Yes	0.135	0	0	-0.055*** (0.005)
29. attack 6	Yes	0.034	0	0	-0.023*** (0.002)
30. attack 7	Yes	0.039	0	0	-0.026*** (0.002)
31. attack 8	Yes	0.042	0	0	-0.031*** (0.002)
32. brokenHH	Yes	0.504	-0.227*** (0.006)	0	0
33. days suspendedANY 6	Yes	0.065	0	-0.024*** (0.005)	-0.049*** (0.004)
34. days suspendedANY 7	Yes	0.097	0	-0.034*** (0.006)	-0.076*** (0.005)
35. days suspendedANY 8	Yes	0.113	0	-0.033*** (0.006)	-0.085*** (0.005)
36. days suspendedLOG 6	Continuous	0.307	0	-0.290*** (0.052)	-0.547*** (0.069)
37. days suspendedLOG 7	Continuous	0.283	0	-0.224*** (0.048)	-0.737*** (0.091)
38. days suspendedLOG 8	Continuous	0.474	0	-0.024 (0.022)	-0.753*** (0.099)
39. family incomeLOG	Continuous	3.829	0.376*** (0.010)	0	0
40. family income group	Missing	0.237	-0.031*** (0.005)	0	0
	<$15K	0.105	-0.078*** (0.004)	0	0
	≥$15k	0.658	0.109*** (0.005)	0	0
41. marij 6	Yes	0.031	0	0	-0.023*** (0.002)

(*Continued on next page*)

Note: Average marginal effects correspond to the coefficients from the measurement equations described in Section 3 and Table 17. Standard errors are reported in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10% level, respectively.

Table 20: Summary of Factor Loadings: Average Marginal Effects of Factors

Measure	Categorical Values	Average Value	Family	Math/ Verbal	Behavior
42. marij 7	Yes	0.044	0	0	-0.030*** (0.003)
43. marij 8	Yes	0.064	0	0	-0.043*** (0.003)
44. mom parent style 6	Uninvolved	0.080	-0.032*** (0.008)	0	0
	Permissive	0.329	-0.004 (0.016)	0	0
	Authoritarian	0.121	-0.042*** (0.007)	0	0
	Authoritative	0.470	0.077*** (0.016)	0	0
45. mom parent style 7	Uninvolved	0.119	-0.039*** (0.007)	0	0
	Permissive	0.321	0.017 (0.015)	0	0
	Authoritarian	0.123	-0.020** (0.009)	0	0
	Authoritative	0.436	0.042*** (0.015)	0	0
46. mom parent style 8	Uninvolved	0.123	-0.040*** (0.006)	0	0
	Permissive	0.322	0.018* (0.010)	0	0
	Authoritarian	0.147	-0.027*** (0.007)	0	0
	Authoritative	0.408	0.048*** (0.009)	0	0
47. num HH under18	Continuous	2.202	-0.200*** (0.015)	0	0
48. piat 6	Continuous	95.997	0	9.133*** (0.431)	0
49. piat 7	Continuous	96.408	0	10.330*** (0.323)	0

(*Continued on next page*)

Note: Average marginal effects correspond to the coefficients from the measurement equations described in Section 3 and Table 17. Standard errors are reported in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10% level, respectively.

Table 20: Summary of Factor Loadings: Average Marginal Effects of Factors

Measure	Categorical Values	Average Value	Family	Math/ Verbal	Behavior
50. piat 8	Continuous	95.632	0	9.706^{***} (0.417)	0
51. retention 6	Yes	0.009	0	-0.005^{***} (0.001)	-0.006^{***} (0.001)
52. retention 7	Yes	0.014	0	-0.007^{***} (0.001)	-0.011^{***} (0.001)
53. retention 8	Yes	0.019	0	-0.008^{***} (0.002)	-0.017^{***} (0.002)
54. sex 6	Yes	0.029	0	0	-0.023^{***} (0.002)
55. sex 7	Yes	0.065	0	0	-0.043^{***} (0.003)
56. sex 8	Yes	0.110	0	0	-0.068^{***} (0.003)
57. smoke 6	Yes	0.088	0	0	-0.050^{***} (0.003)
58. smoke 7	Yes	0.097	0	0	-0.042^{***} (0.004)
59. smoke 8	Yes	0.112	0	0	-0.054^{***} (0.004)

Note: Average marginal effects correspond to the coefficients from the measurement equations described in Section 3 and Table 17. Zero denotes that a given factor does not load in the measure. Standard errors are reported in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10% level, respectively.