THE PLIGHT OF MIXED RACE ADOLESCENTS

Roland G. Fryer, Jr
Lisa Kahn
Steven D. Levitt
Jörg L. Spenkuch

ABSTRACT

Over the past 40 years the fraction of mixed race black-white births has increased nearly nine-fold. There is little empirical evidence on how these children fare relative to their single-race counterparts. This paper describes basic facts about the plight of mixed race individuals during their adolescence and early adulthood. As one might expect, on a host of background and achievement characteristics, mixed race adolescents fall in between whites and blacks. When it comes to engaging in risky/anti-social adolescent behavior, however, mixed race adolescents are stark outliers compared to both blacks and whites. We argue that these behavioral patterns are most consistent with the "marginal man" hypothesis, which we formalize as a two-sector Roy model. Mixed race adolescents -- not having a natural peer group -- need to engage in more risky behaviors to be accepted. All other models we considered can explain neither why mixed race adolescents are outliers on risky behaviors nor why these behaviors are not strongly influenced by the racial composition at their school.

Roland G. Fryer, Jr
Department of Economics
Harvard University
Littauer Center 208
Cambridge, MA 02138
and NBER
rfryer@fas.harvard.edu

Lisa Kahn
Yale School of Management
135 Prospect St
PO Box 208200
New Haven, CT 06520
lisa.kahn@yale.edu

Steven D. Levitt
Department of Economics
University of Chicago
1126 East 59th Street
Chicago, IL 60637
and NBER
slevitt@midway.uchicago.edu

Jörg L. Spenkuch
Department of Economics
University of Chicago
1126 East 59th Street
Chicago, IL 60637
jspenkuch@uchicago.edu

I. Introduction

Throughout history racial mixing has been taboo.[1] Fear of interracial mixing was a driving force behind the Jim Crow system in the South and the Black Codes in the North (Romano 2003). Romano (2003) reports that children of mixed racial heritage were thought to be morally and physically inferior to "pure" blacks, and more prone to diseases such as tuberculosis. In *Perez v. Sharp* (1948) the State of California argued before the California Supreme Court that anti-miscegenation laws protected the larger social good because the children of racially mixed couples were biologically inferior.[2] Even staunch supporters of the civil rights movement were quick to draw a distinction between political and social equality (Moran 2003). Cohn (1944) argued that blacks should be given justice in the courts, protection of their property, a fair distribution of tax money, and equal wages. But, he insisted, white southerners would not forgo the segregation that kept blacks and whites separate for fear that any breach in the walls of social segregation would lead to racial mixture.

Despite its slow beginning, interracial intimacy has been increasing over the second half of the 20th century. Figure 1 plots the share of interracial marriages, by gender and race of spouse, from 1880 to 2000. In 1920, marriages to whites comprised roughly 0.3% of black marriages (Fryer 2007a)[3]. By 2000, interracial marriages increased dramatically; 5.9% of

---

[1] Laws governing the integration of schools, neighborhoods, and intimate relationships were among the last civil rights to be granted. Between 1913 and 1948, 30 out of 48 states banned interracial marriage. In 1948 California was the first state since 1887 to repeal its antimiscegenation law. This was done in response to *Perez v. Sharp*. On June 12, 1967, the landmark Supreme Court decision in *Loving v. Virginia* ruled, unanimously, that preventing marriages between individuals solely on the basis of racial classification violated the Equal Protection and Due Process Clauses of the Fourteenth Amendment. This ruling struck down anti-miscegenation laws in 16 remaining states. See Moran (2003) for a thoughtful review.

[2] Some whites even went as far as to claim that mixed-race children would be sterile like mules (Romano 2003). Hoffman (1896) argues that miscegenation is responsible for the increasing black mortality rate, as well as blacks' "consequent inferior social efficiency and diminishing power as a force in American national life."

[3] We are well aware that black-white individuals comprise only a fraction of all people of mixed race in the US today. However, given the historical opposition to the mixing of blacks and whites, and blacks' special position in American society, it seems warranted to focus our attention in this group of mixed race individuals.

married black men chose a white bride and 2.7% of black women chose a white husband (Fryer 2007a).

Accompanying the swift changes in interracial marriage were increases in the number of mixed-race births. Figure 2 presents a time series of black-white births from 1920-2000, computed using age-specific responses given on the 2000 Census, which allowed individuals to check multiple race categories for the first time.[4] Before 1960, interracial births were a negligible share of total births. Mixed race births steadily increase over time, accounting for one in 200 births by the 1980s and one in 70 births in 2000. Despite this recent increase, empirical evidence on the experience of mixed race children is scarce. This is particularly surprising given the nexus of opposition to interracial mixing has been the implications on the children of such marriages (Romano 2003).[5]

Using data from the National Survey of Adolescent Health (Add Health), which surveys over 90,000 individuals who were in grades 7-12 in the 1994-95 school year, our empirical analysis of the plight of mixed race adolescents unearths a rich set of new facts. Figure 3 presents a high-level summary of these findings. Using the wide range of variables available in Add Health we construct index measures of each child's home environment (e.g. household income, father in household, mother's education), physical characteristics (e.g. height, BMI, physical attractiveness), achievement (e.g. GPA, test scores), risky/anti-social behaviors (e.g. trouble with teacher, smoking, lying to parents, violent acts, etc.), and psychological wellbeing (e.g. feel loved, not depressed, chances of living to 35). On each of these index measures, we run

---

[4] It is possible that individuals in earlier cohorts are more reluctant to identify themselves as mixed race, exaggerating the growth in this category over time. Also, to the extent that survival rates differ between racial groups, our estimates will be biased.
[5] Romano (2003) reports that whites considering interracial marriage are nearly always asked the question "What about the Children?"

regressions including indicator variables for a child's race and controlling for age, age squared, gender, whether the student was born abroad, and region.[6] Figure 3 shows the coefficients from these regressions on the race dummies for blacks and whites, normalized such that black-white mixed race children have a coefficient of zero. For all composite measures higher values are better. Standard error bands are displayed only for mixed race children; the estimates on the other racial groups are much more precise because of larger sample sizes for those groups.

According to Figure 3, mixed race children are reared in home environments that overall are similar to those of black children. On the physical dimension, mixed race children parallel whites. School achievement results are intermediate between blacks and whites. Strikingly, however, mixed race children engage in substantially more risky/anti-social behavior than either blacks or whites, especially outside of school. Of the twenty-one different bad behavior variables that we analyze, mixed race adolescents are worse than both blacks and whites on 15 of them; they fall in between blacks and whites on the remaining 6 measures. Mixed race children also fare somewhat worse on measures of psychological well-being.[7]

We argue that these empirical patterns are largely consistent with the "marginal man" hypothesis (Park 1928, 1931, Stonequist 1935, 1937), a highly influential, yet rarely tested description of the experiences of mixed race adolescents. We reinterpret and formalize the marginal man hypothesis using a simplified two sector Roy model (Roy 1951), in which all adolescents face pressure to conform to peer norms. For mono-racial adolescents, this norm is determined by their race: black adolescents adhere to black norms and white adolescents adhere

---

[6]Results are robust to the inclusion of more extensive controls.
[7] We also have some very limited data on outcomes of young adults from the third wave of Add Health. Unfortunately, the sample sizes for mixed race children on these adult outcomes are extremely small, leading to estimates that are quite imprecise and often deviate greatly from estimates one could derive using the 2000 Census. Therefore, we do not focus on these outcomes, though it is important to note that mixed race adolescents are weakly above blacks in nearly all outcomes in both datasets.

to white norms. Mixed race children have a choice, they can choose to associate with black children and adopt their norms, befriend white children and adopt their norms, or both. It is this outside option that gives mixed race adolescents a higher cost of group acceptance, resulting in them choosing riskier behaviors to gain such acceptance. The key distinction between a Bernheim-type conformity model (Bernheim 1994) and the Roy model is that the former predicts that mixed race children who mostly interact with whites will adopt white behaviors, whereas mixed race children whose peer groups are mostly black will act black. In contrast, in the Roy model, when there are few blacks around, mixed race children can have a comparative advantage in black behaviors, inducing them to act particularly "black." Empirically, the evidence on this point favors the Roy model over conformity.

Our analysis builds upon a relatively small prior literature on the subject, especially Harris and Thomas (2002) and Ruebeck, Averett and Bodenhorn (2008), both of whom also use Add Health data.[8] Harris and Thomas (2002) focus on educational outcomes such as GPA, grade repetition, and test scores, generally finding that mixed race black-white children have outcomes between blacks and whites, but in some cases closer to and not statistically distinguishable from whites. They argue that the "marginal man" hypothesis predicts that mixed children should have educational outcomes closer to blacks than to whites, which leads them to reject it.[9] Ruebeck et

---

[8] Less closely related are studies of mulattoes in the antebellum South, showing that this group occupied a distinct position between blacks and whites, or at least strived to do so (Bodenhorn and Ruebeck 2003). Light skinned blacks were much more likely to operate farms and accumulated more wealth than their dark skinned counterparts (Bodenhorn 2002, Bodenhorn and Ruebeck 2007). There is a more recent literature that relates skin tone to economic outcomes (Keith and Herring (1991) and Hill (2000)).

[9] Furthermore, Harris and Thomas (2002) make the argument that the "marginal man" hypothesis also predicts that Asians should have worse educational outcomes than whites, and that mixed asian-white children should fall in between their monoracial peers, which is not the case in their analysis. It is important to note that the original hypothesis exposited by Stonequist (1935, 1937) and Park (1928, 1931) had no predictions for Asian-White intermixing.

al. (2008) is the paper most similar to ours.[10]   In independent research, they analyze many of the same outcomes that we consider through the lens of a Bernheim-type identity/conformity model. They find that mixed race children adopt behaviors that are characterized both as "white" and as "black." They find a greater variance in mixed race behavior than is observed for either whites or blacks, concluding that mixed-race identities are not as strongly codified as those for monoracials.

The remainder of the paper is structured as follows. Section II describes the data used in the analysis and the process of identifying mixed race adolescents. Section III describes the plight of mixed race adolescents from adolescence to early adulthood. Section IV interprets the facts described in section II and III through the lens of economic theory. Section V concludes. A data appendix provides further details of our sample construction.

II. Data Description and Identification of Mixed Race Adolescents

The absence of systematic empirical research on how mixed race children fare relative to their monoracial peers is largely due to a lack of adequate data. Few data sets record racial information in a way that mixed race children can be identified. Data sets that include a mixed race classification are either non-representative (Choi 2006) too small to be useful (Nasim et. al. 2007), or contain little information on childhood or adolescent experiences (e.g. Census 2000). The notable exception to this data shortcoming is the restricted-use version of the National

---

[10] We did not become aware of Ruebeck et al.'s (2008) parallel research agenda until shortly before the completion of the current draft of this paper.

Longitudinal Study of Adolescent Health (Add Health) – the primary data set used in this paper.[11]

Add Health began as a stratified random sample of all high schools in the US, resulting in a nationally representative sample of 90,118 students entering grades 7 through 12 in the 1994-95 school year. A sub sample (of the original in-school survey) of 20,745 students was given a series of in-home interviews.[12] The original data collection took place in 1995, with Wave II done in 1996 and Wave III carried out in 2001.[13] In addition, Wave I included a parent questionnaire conducted at home, in which 17,700 out of roughly 20,000 of the children's parents participated. As in all longitudinal data, some respondents could not be located or contacted after repeated attempts, refused to participate, or were unable to do so. Sample weights supplied with each wave attempt to correct for observed patterns of non-response.

A wide range of data is gathered on the adolescents in the study.[14] We use a wide range of data on demographics, family background, psychological well-being, behavior and academic achievement. In all our analysis we use missing value indicators and sample weights.

Classifying children as mixed race

Identifying individuals of mixed race in survey data can be tricky, and there is little consensus about the best way to do so (Robbin 2000).[15] We conceptualize the tradeoffs in

---

[11] We use the restricted-use data set that contains the full sample and more detailed information. The number of observations in restrictive-use version of Add Health is 90,118, while only roughly 6,500 observations are available in the public-use version. Furthermore, the restricted-use version contains more detailed information related to friend and sibling identification, respondents' romantic relationships, and spatial distances.

[12] This sample was selected in part to be representative of the full sample (a core of 12,105) as well as selection on several criteria for oversamples: disabled, blacks from well-educated families, Chinese, Cuban, Puerto Rican, and adolescents with siblings. The response rate on the first wave of the home interview was 78.9%.

[13] The response rates on the second and third waves of the survey were 88.2% and 77.4%, respectively.

[14] For a detailed description, see the Add Health website http://www.cpc.unc.edu/projects/addhealth

constructing such a measure using a 2×2 matrix. The first relevant dimension is whether one bases the classification on the responses of the parents or the child. The second relevant dimension for classifying mixed race children is how "strict" or "inclusive" one is in defining who is mixed race. Our preferred definition is one that is child-based and strict. We use a child-based definition both because data on fathers are often missing, and even if a male guardian is present, it is impossible in our data to determine whether he is the biological father.[16] A child-based approach is more direct, but relies on the child's self-identification.

Under our preferred "strict" definition at the child level, an individual is considered race A if and only if he consistently says he is race A whenever he is observed in the data. If there are any inconsistencies across waves, we code the race as missing. Using this approach we obtain 304 black-white mixed children; but we are likely to understate the number of mixed race adolescents in the data.[17] Nonetheless, the resulting frequency of mixed race children is

[15] There is a small literature in sociology and population studies on racial identification (Goldstein and Morning 2000, Lee 1993, Aspinall 1997, 2003, Harris 2002, Harris and Sim 2002, Anderson and Fienberg 1999, Davis 1991, for the UK Wilson 1984). Kao (1999), using the National Education Longitudinal Study, defines an adolescent as mixed race if their race differs from the race of their guardian who completes the parent survey. Harris and Thomas (2002) identify ki adolescents as mixed race if they self identify as mixed race or they provide inconsistent monoracial categories on different waves of the survey. Brunsma (2005) looks at children who select into the "More than one race" category on the Early Childhood Longitudinal Study. Xie and Goyette (1997) use the 1990 Public Use 5% Micro Sample of the US Census and classify children as multi-racial if they are living with two parents who check different single race categories. A few empirical studies allow an observer / interviewer to identify who is mixed race (Harris and Sim 2002, Hahn, Benedict and Barker 1996, Telles and Lim 1998 for the case of Brazil).

[16] Ruebeck, Averett, and Bodenhorn (2008) make use of the racial classification of an adolescent's parents in Add Health and code an individual as mixed race if one of his parents is mixed or black and the other one is white or mixed. To increase their sample size and lessen the degree of selection associated with children living in intact families, they also employ a definition based on children's self-selection into black and white on the home survey.

[17] As a robustness check, we looked at the consistency of racial identification for whites and blacks in the Add Health and the Early Childhood Longitudinal Survey (ECLS). In the Add Health 16.3% of blacks would be dropped from the data for inconsistent racial identification and 10.7% of whites. In ECLS, these numbers are 1.1% for blacks and 0.2% for whites. We have also explored a more "inclusive" child-level definition, under which there are three additional ways in which a student can be classified as mixed race. First, if he is consistently non-Hispanic and marked both black and white, and no other race, in all waves in which he participated. This corresponds to our strict definition. Second, a student is coded as mixed if he qualified as mixed under the above definition on any single survey, even if he failed to do so in other instances. Third, a student is coded as mixed if he marks a combination of black and white and no other race across surveys. For example if the student marks only black at school and only white at home then he is coded mixed. This is the same procedure used in Harris and Thomas (2002) and is done so to obtain as many potentially mixed race students as possible.

consistent with that observed in the 2000 Census among children of the appropriate age. In the 2000 Census children who check black and white as race constitute approximately 0.38% of the population between the ages of 12 and 18—the age range of the overwhelming majority of children in Add Health. The percentage of mixed race children in Add Health, employing our strict, child-based definition, is 0.34%. All of the results we report in the paper correspond to this definition of mixed race; full results for the other definitions are available from the authors on request. While results for particular variables do, of course, vary across these definitions, the same basic patterns are present (with slightly smaller standard errors) under each of these definitions.

Variables used in the analysis

　　We broadly classify variables into five categories: home environment, physical characteristics, school achievement, risky/anti-social behavior, and psychological wellbeing. For further details on these variables and their construction, see the data appendix.

　　Demographic variables include age, gender, whether born in the United States, and region of residence. Our set of home environment variables consists of 10 variables. These include household income, receipt of public assistance such as welfare, father in the household, the marital status of the parent filling out the questionnaire; mother's age, whether or not their mother is a college graduate or has ever been married, and years in current residence,.

　　Our measures of physical attributes include birth weight, height, BMI, and interviewer rated attractiveness.[18]   Achievement is proxied by score on the Add Health Picture Vocabulary

---

[18] At the end of the in-home interview in Wave I the interviewer was asked to rate the physical attractiveness of the respondent on a scale from 1 to 5, where 0 indicates "very unattractive", and 5 indicates "very attractive." Whenever we use this variable in our analysis, we include interviewer fixed effects to account for interviewer specific tastes.

Test (AHPVT), grade point average (GPA), whether or not a student has repeated a grade, and whether the student has a learning disability.[19] The AHPVT is an abridged and computerized version of the well-known Peabody Picture Vocabulary Test-Revised, conducted as part of the Wave I home interview for 19,713 children. GPA is based on student self reports; grade repetition and learning disabilities are drawn from the parent questionnaire. [20]

We analyze twenty-one variables designed to capture a student's risky/anti-social behavior in and out of school. The in-school behavior variables include: trouble with teachers, trouble paying attention, trouble with homework, trouble with students, effort on schoolwork, skipping school, and never suspended or expelled. The variables designed to measure behavior out of school include: watch TV, drink, smoke, dare, lie to parents, fight, property damage, steal, violent acts, sell drugs, encounter violence, ever sex, ever STD, and ever illegal drugs. In the survey many of these questions take the form "Since school started this year, how often have you had trouble ….?" Answers to these questions range, for example, from 0 to 4, where 0 indicates "never," and 4 indicates "everyday." For all such questions with answers that are ordinal, but do not have clear cardinality, we normalize responses to be mean zero with standard deviation equal to one in our weighted sample for ease of comparison. [21]

Our final category of variables corresponds to a child's (broadly defined) psychological wellbeing. These variables include the child's responses to questions such as the degree to

---

[19] Harris and Thomas (2002) analyze three of these outcomes: grade point average, grade retention, and AHPVT.
[20] To get a sense of how much exaggerating there might be in the data among different racial groups, we compared the survey data with data collected from official transcripts, for the first year of high school of 12,115 Wave III respondents who signed a transcript release form. Blacks reported that their GPAs were 10.9% higher than the actual GPA compared to 4.5% for whites and 8.7% for mixed race students. Due to a lack of objective measures we are forced to rely on self-reports for other outcomes we report. Thus, differential self-reporting is a possible source of bias in our analysis, but the available data on GPA suggests it cannot explain our patterns.
[21]We have also estimated ordered logistic regressions which yield the same qualitative results. We focus on the normalized regressions for ease of interpretation of the main coefficients.

which his mother or father cares about him, how close he feels to other people, whether he feels accepted or loved, is depressed, likes himself, and expects to live to 35.

III. Empirical Evidence on Mixed Race Adolescents

In this section, we describe basic facts about the plight of mixed race adolescents on the myriad dimensions we consider: demographics, home environment, physical characteristics, academic achievement, behaviors in and outside of school, and psychological wellbeing. Summary statistics for the variables we use in our analysis are displayed in Table 1. The table presents means with standard deviations in parenthesis for whites, blacks, and mixed black and white students under our strict child centered definition.[22] As noted earlier, except where there are natural units for a variable (e.g. household income or weight), we have normalized the responses to be mean zero with a standard deviation of one in our sample.

Rather than discuss the summary statistics in tabular form, we find it more informative to display the results, adjusted for a small set of controls for background characteristics, graphically. For each outcome measure, we estimate an equation taking the form:

$$outcome_i = a + R_i\beta + X_i\gamma + \varepsilon_i,$$

where R_i represents a mutually exclusive set of racial identifiers with whites as the omitted category and X_i is a vector of controls for age and its square, gender, whether or not a student was born in the United States, and the region in which she lives. It is imperative to note that our results are remarkably robust to the inclusion of all our home environment variables as further controls. In most cases, controlling for a child's home slightly decreases the distance between the racial groups, but it widens the gap for three outcome variables. In only three cases (Sell Drugs,

---

[22] Students of all other races are dropped from the analysis.

Mother Cares, and Feel Loved) do we observe changes in the relative ordering of the racial groups; the absolute changes in the coefficients are very small.

In what follows, we report values for black, white, and mixed race individuals from estimating the simple regression model above with limited controls; the value for whites is given by the constant. We report all estimates relative to mixed race children by normalizing the mixed race coefficient to 0. Estimates for blacks and whites are then reduced by the mixed race coefficient to maintain their relative distance. We also display standard error bands for the mixed race coefficient. In all cases, the standard errors for blacks and whites are much smaller because of larger sample sizes for these groups.

Figure 4 shows home environment variables. Mixed race adolescents have family incomes similar to black children and have the same likelihood of having a father in the home. The parents of mixed race children are less likely to have been married than those of other races (although rates of ever married are higher than for these mothers than for blacks). The mothers of mixed race children are younger on average, but slightly more likely to be a college graduate. Mixed race children are much more mobile than their single race counterparts.[23]

In Figure 4, and elsewhere throughout the paper, we rely on the relatively small sample of mixed race children available in Add Health. For the home environment variables, unlike the other variables we use in the paper (e.g. anti-social behaviors, academic achievement, psychological wellbeing), it is possible to verify the patterns in Add Health using the 2000 Census. Census results for a set of home environment variables similar to those in Add Health are reported in Figure 5. The results are generally quite similar, but with some differences.

---

[23] One possible explanation for the greater mobility of mixed race children might be that their parents are more likely to be members of the military. We are able to dismiss this explanation as the same pattern emerges in the 2000 Census after we exclude all children whose parents currently or formerly served in the armed forces.

Mixed race children fare slightly better with respect to household income, having a father in the home, and having married parents in the census, but still lag far behind whites. The mothers of mixed race children are not as young in the Census as they are in Add Health.

Figure 6 plots our set of physical variables. Mixed race birth weights look more like whites than blacks. Differences in adolescent height and BMI are small. Mixed race adolescents are rated as .41 standard deviations more attractive than white children, and .47 standard deviations more attractive than blacks.

Figure 7 presents academic outcomes. Mixed race adolescents are less likely than blacks or whites to have a learning disability. Their AHPVT scores are roughly in the middle of blacks and whites. While blacks fare .89 standard deviations worse than whites, mixed race children lag .33 standard deviations behind. On our other two achievement variables (grade point average and whether or not a student repeated a grade), mixed race adolescents are between blacks and whites but more similar to blacks.

The next set of variables we analyze are behavioral measures, displayed in Figure 8. In all cases, higher values are indicative of better behavior. Strikingly, on fifteen out of the twenty-one variables relating to bad behavior mixed race adolescents exhibit worse behavior than both of their single race counterparts. Behavior at school by mixed race adolescents generally mirrors that of blacks, except with regard to exerting effort and skipping school – two dimensions on which mixed race children are significantly worse than blacks. The bad behavior of mixed race children stands out even more clearly outside of school. With the exception of watching television (which blacks do more of), mixed race adolescents are the worst or essentially tied for worst on every other behavior considered. This is true whether the risky behaviors are those more common to whites (e.g. drinking and smoking) or to blacks (e.g. sex and violence). Mixed

14

race adolescents roughly occupy the lower envelope of good behaviors of blacks and whites.[24] In

a similar manner Choi et al (2006), using data from four middle schools in Seattle, WA, find that

multi-racial adolescents are at greater risk for substance abuse.

One potential explanation for the worse self-reporting of behavior for mixed race

adolescents is that they have different reference points or norms in their peer group relative to

blacks. For instance, it is possible that among whites one incident of fighting in a school year is

"very often" whereas among black adolescents that frequency is considered "not at all." One test

of this hypothesis is to compare the responses of adolescents across racial groups within the

same school or even the same peer network. Controlling for school or connected component

fixed effects does little to alter the results.

Figure 9 reports on psychological wellbeing. Although generally not statistically

significantly different, mixed race children do worse on most of the psychological dimensions

explored. The greatest observed difference is with respect to whether the child perceives his

father as caring, which mixed race children do significantly less. Interestingly, blacks tend to be

more content on most dimensions relative to whites and mixed race children. The exception to

this finding is when asked about their chances of surviving to age 35.[25]

Broadly summarizing, the data on mixed race black-white children suggests that they grow

up in home environments that are similar to blacks, have academic achievement in between that

of whites and blacks, but engage in much more risky behaviors and are slightly worse off

psychologically.

---

[24]One possible explanation for why mixed race adolescents fight more often is that they might get picked on more
than blacks or whites. However, if mixed race children got picked on a lot, we would expect them to say that they
don't feel accepted. Yet, on that measure they are practically identical to whites (see figure 9).
[25] Under the parent-strict definition of mixed race adolescents, these differences become quite stark.

IV. Interpreting the Data through the Lens of Economic Theory

In this section we attempt to make sense of the wealth of information presented above using social and economic theory. In what follows, we see to what extent existing theories can account for this disparate set of facts. The key issue is that it is not obvious what type of economic model can reconcile the patterns in the data, particularly their especially bad behavior.[26] Discrimination-based models are unlikely to provide an answer. In almost all models with human capital investment and discrimination, lower levels of discrimination lead to more investment in human capital (e.g. Becker 1957, Arrow 1973, Fryer and Jackson 2008). If mixed race individuals face weakly lower discrimination than blacks (for instance because they have lighter skin), one would expect weakly more human capital investment on their part. So if one interprets bad behavior as interfering with human capital investment, traditional discrimination models cannot explain our findings.[27] We begin our analysis by considering the most prominent sociological theory, re-interpreting it through an economic lens. We conclude by comparing and contrasting this model with a conformity model.

The 'Marginal Man' Hypothesis

The most prominent theory of mixed race individuals is the marginal man hypothesis, developed by Park (1928, 1931) and Stonequist (1935, 1937). The "marginal man" is depicted as

---

[26] Additionally, we have examined young adult outcomes for black-white mixed individuals in Add Health and the 2000 Census, but the results are inconsistent across the two data sets, so we have not made this the focus of our analysis. Broadly summarizing, mixed race children fall in between their monoracial counterparts in both data sets. Detailed young adult outcomes are available from the authors upon request.

[27] Arcidiacono, Bayer and Hizmo (2008) and Lang and Manove (2006) present models in which discrimination can actually increase educational investment. In these models discrimination decreases with educational attainment. Thus, it can be beneficial to "overinvest" in education to face lower levels of discrimination. These models predict that, at least for certain parts of the ability distribution, blacks invest more in education than mixed race individuals, and mixed race adolescents invest more than whites. However, neither model explains why mixed race children invest less than both blacks and whites.

someone who lives in a bi-cultural environment and who is caught between the two conflicting

cultures. This produces dual identification and loyalty to both groups, neither of which he is

willing to let go, thus causing inner conflict. Inner conflict is seen as the defining characteristic of

the "marginal man's" personality.[28] For Stonequist (1935) and Park (1928), being multiracial was

a defining quality of an individual, and the marginal man was a distinctive personality type.

Stonequist and Park argued that anxiety associated with racial marginality "initiates a process of

disorganization which finds expression in statistics of delinquency, crime, suicide and mental

instability" (Stonequist 1935, p. 12).

Although broadly consistent with our data on behavior and psychology, testing the

marginal man hypothesis is difficult because it makes few readily falsifiable predictions. For

instance, the relationship between inner conflict and achievement can take on many forms and

the theory, as it stands, does not provide many restrictions.[29]

Formalizing the marginal man hypothesis via a two sector Roy model

---

[28] The "marginal man" concept has been criticized by various authors, largely based on counterexamples and subtle theoretical grounds (Goldberg 1941, Golovensky 1952, Green 1947, Antonovsky 1956). Most importantly these authors have argued that living in a bicultural environment does not automatically result in inner conflict. In his qualitative analysis Antonovsky (1956), for instance, finds multiple coping strategies among American Jews.

[29] Even one of the most concrete comparative statics from the marginal man hypothesis – marginal individuals prefer to associate with the social group of higher status than to bond with the one of lower status – is difficult to pin down. Neither Stonequist nor Park were clear as to whether status was measured at a local or global level and whether that could differ for individuals. If it's a global phenomenon, between blacks and whites in contemporary America, this implies that mixed race black-white children will more closely identify with whites. Only 34.5% of mixed race black and white adolescents identify with whites. (During the home interview in Wave I children we reported to be multiracial were asked "Which one [racial] category best describes your racial background?" 19 out of the 55 black-white children who answer this question say "white," while 36 report "black." If we further limit the sample to students in schools with more than 75% white children, then 5 out of 13 children answer "white," and 8 say "black.") If it is a local measure, then it is difficult to test because within a local community it may not be possible to determine which group has higher status. In predominately black communities, whites are often low status (Conley 2001)

One way to generate more precise predictions from the marginal man hypothesis is to import the key insights into a formal economic model. We do this using a classic Roy model (Roy 1951) with two sectors (peer groups).

Consider a population of many agents, each of whom selects some publicly observable behavior x from the set X. These behaviors may include patterns of speech, style of clothing, time engaged in certain activities, music on their iPod, and so on. Each agent has intrinsic preferences over the set X, which we summarize by a utility function $g(x-t)$. The parameter t represents an agent's intrinsic bliss point: the $\underset{x \in X}{\arg\max}\, g(x-t)$. Following Bernheim (1994), we assume that $g(\cdot)$ is twice continuously differentiable, strictly concave, symmetric, and achieves its maximum at $g(0)$.

In addition to these fairly standard preferences, each individual also cares about popularity/social esteem in school. We assume that popularity depends solely on whether or not an individual is an accepted member of the peer group. Blacks and whites have a predetermined peer group (equivalent to assuming that the cost of switching sectors is prohibitively high) whereas mixed race adolescents are able to choose whether to identify with whites or blacks, though they pay a cost.

Let $p_b(x)$ (resp. $p_w(x)$) denote the probability that the black (resp. white) peer group accepts and individual with observable behavior x. Examples abound. If x is a measure of academic achievement, wearing clothes from GAP, or whistling Vivaldi $p_b(x)$ is likely to be decreasing in x. Individuals who are accepted by the group garner a payoff $\beta > 0$. We denote the cost of choosing sector j by c_j.

Each agent chooses x to solve:

$$\max\left\{\max_{x\in X} g(x-t) + p_b(x)\beta - c_b, \max_{x\in X} g(x-t) + p_w(x)\beta - c_w\right\}.$$

To keep the model very simple, we will assume that $c_j = 0$ and $p_j(x) = 1$ for all monoracials.

Thus, their decision problem can be represented as: $\max_{x\in X} g(x-t) + \beta$. The solution of which is

$x = t$. This simplification abstracts away from peer dynamics that are surely going on within

racial groups. All of the results of this simple model are consistent with a more general model

were all racial groups can choose their peer group if the cost of switching sectors is significantly

lower for multiracial adolescents. [30]

Our formalization of the marginal man hypothesis as a two-sector Roy model is consistent

with the facts described in the previous section. The fact that mixed race adolescents engage in

more risky behavior than any of their single race peers can be rationalized through $p_j(x)$.

Because mixed race adolescents have an outside option, they must go to greater lengths to

demonstrate their affiliation with the group; one way of demonstrating solidarity is to go to

extremes in carrying out group-sanctioned misbehavior (see Austen-Smith and Fryer (2005) or

Fryer (2007b) for a micro model of group dynamics which justifies this assumption). [31] Berman

(2000) argues that groups provide public goods and charge their member for group membership

by requiring them to make costly, but unproductive, investments. When it is easy to leave the

group after receiving the public good, groups do better by holding their members to tougher

standards.

---

[30] In many Roy models, the individuals with the most choice have higher utility. In our model, there is tradeoff between having the benefit of more choice and incurring the cost of not having a natural peer group, so we are unable to sign the change in utility.
[31] Patterns in extracurricular sports roughly follow the distribution one might expect in the Roy model. Whites tend to participate more in baseball, field hockey, ice hockey, soccer, swimming, tennis, volleyball, and wrestling; whereas blacks are more likely to do track or play football and basketball. Mixed race adolescents do more of all sports than blacks, and only slightly less of the typically white sports than whites.

Comparing and contrasting a Roy model and a conformity model

The Roy model described above is similar in spirit to a conformity model along the lines of Bernheim (1994), generalized to allow different groups to have different social norms. In the classic conformity model, an individual's type is unobservable to her peers. Thus, popularity is determined by other students' perceptions of a student's type. We assume that all agents within a racial group will, in equilibrium, form the same inferences about an individual's type; thus we can summarize a student's perceived type by a real number, q. Let $s(q)$ denote the social esteem of an individual thought to be type q. Types are unobservable, but others can infer an individual's type from her observable choices. Let $\varphi(q,x)$ be the inference function that links observable behaviors to perceived types and $\lambda_j, j \in \{b,w\}$ the weight that an individual puts on acceptance by group j.

Using the same notation as above, the utility maximization of an individual in the classic conformity model can be expressed as follows:

$$\max_{x_b, x_w \in X} \left\{ g(x_w - t_w, x_b - t_b) + \lambda_w \int_T s_w(q)\varphi(q, x_b, x_w)dq + \lambda_b \int_T s_b(q)\varphi(q, x_b, x_w)dq \right\},$$

where x_w and x_b represent "white" and "black" behaviors, respectively, and other racially dependent parameters are defined similarly.

If the weight on group acceptance is large enough and the inference functions for mixed race adolescents are "flatter" than those for the monoracial groups, then the model explains why mixed race adolescents over invest in bad behaviors.[32] As in the Roy model, for mixed race

---

[32] If extreme levels of bad behavior are detrimental to group acceptance, then, compared to monoracial children, mixed race adolescents will still overinvest in bad behaviors; unless, of course, *any* bad behavior decreases their chances of acceptance.

adolescents, it takes more black behaviors to be accepted by the black peer group and more white behaviors to be accepted by the white peer group.

If one interprets λ_b and λ_w as the fraction of blacks and whites in an individual's school, then under the conformity model we would expect mixed race adolescents who attend more black schools to act more black and those that attend more white schools to act more white. The predicted pattern is not, however, observed in the data. In Add Health the median mixed race child attends a school at which whites comprise 53.7% of the student body. Surprisingly, after controlling for a student's race and school-fixed effects, mixed race children attending schools with a share of whites below the median actually exhibit slightly *less* of the "black" bad behaviors (.06 standard deviations, standard error = .12).[33] A similar, but even more pronounced picture emerges if we split up the mixed race adolescents by the fraction of blacks at school rather than the fraction of whites. In this case, the mixed children at schools with a below average share of blacks (less than 12 percent) score .18 standard deviations *higher* (standard error = .12) on our measure of black behavior than those mixed race adolescents with above average shares of black peers.

Unlike the conformity model, the two-sector Roy model can be made consistent with the observation that when there are very few blacks in a school, mixed race adolescents act (weakly) more black. It is important to note that when there are few blacks present, the costs of acting black for mixed race adolescents is lower. For example, fighting is one aspect of behavior more associated with blacks than whites. If blacks are more experienced fighters than whites, than it is

---

[33] Our composite measure of black behavior is the sum of typically black behaviors (i.e. getting in trouble with ones teacher, having trouble paying attention, trouble with homework and other students, watching TV, and fighting) standardized to have mean 0 and standard deviation 1; and higher values indicate that a child engages more in these behaviors Before the standardization missing values for single behaviors have been replaced with a racial group's mean.

less costly for a mixed race child to prove he can fight when the only opponents are whites.[34] This force works in the opposite direction of conformity.

V. Conclusion

The number of mixed-race children has increased dramatically. Sociologists have theorized about the challenges facing these individuals since early in the 20th century, but little systematic empirical research has explored their outcomes. Using the Add Health data set, we find high rates of risky/anti-social behavior on the part of mixed race adolescents on virtually every dimension we are able to measure. Formalizing the existing sociological theories using a Roy model, we conclude that the observed pattern of behaviors appears consistent with such a model. While the predictions of the Roy model and a conformity model are generally similar, the Roy model alone can reconcile the observation that when mixed race adolescents are in environments where their peers are predominately black, they are no more likely to adopt black behaviors than when they have peers who are predominately white.

---

[34] Anecdotally, this phenomenon has been observed among the programs for gifted minority youth held at M.I.T. each summer. These programs attract a subset of black and mixed race children who are among the "whitest" acting in their schools. At M.I.T., however, they have a comparative advantage in acting "black," and engage in a wide range of behaviors to signal how "black" they are (Suskind 1999).

References

Anderson, Margaret L., and Fienberg, Stephen E., 1999, *Who Counts? The Politics of Census-Taking in Contemporary America?* Russel Sage, New York.

Antonovsky, Aaron, 1956, "Toward a Refinement of the 'Marginal Man' Concept," *Social Forces,* 35, 57-62.

Arcidiacono, Peter, Bayer, Patrick J. and Hizmo, Aurel, 2008, "Beyond Signaling and Human Capital: Education and the Revelation of Ability". NBER Working Paper #13951.

Arrow, Kenneth., 1973., "The Theory of Discrimination" in *Discrimination in Labor Markets*, Orley Ashenfelter and Albert Rees eds. Princeton University Press, p. 3-33.

Aspinall, Peter, J. 1997, "The Conceptual Bias of Ethnic Group Terminology and Classifications," *Social Science & Medicine*, 45, 689-698.

Aspinall, Peter J., 2003, "The Conceptualisation of and Categorisation of Mixed Race/Ethnicity in Britain and North America: Identity Options and the Role of the State," *International Journal of Intercultural Relations,* 27, 269-296.

Austen-Smith, David, and Fryer, Roland G., 2005, "An Economic Analysis of 'Acting White'," *Quarterly Journal of Economics*, 120, 551-583.

Becker, Gary S., 1957, *The Economics of Discrimination*, University of Chicago Press, Chicago.

Berman, Eli, 2000, "Sect, Subsidy, and Sacrifice: An Economist's View of Ultra-Orthodox Jews," *Quarterly Journal of Economics*, 115, 905-953.

Bernheim, Douglas B., 1994, "A Theory of Conformity," *Journal of Political Economy*, 102, 841-877.

Bodenhorn, Howard, 2002, "The Complexion Gap: The Economic Consequences of Color Among Free African Americans in the Rural Antebellum South," NBER Working Paper 8957.

Bodenhorn, Howard, and Ruebeck, Christopher S., 2007, "Colourism and African–American Wealth: Evidence from the Nineteenth-Century South," *Journal of Population Economics*, 20, 599-620.

Brunsma, David L., 2005, "Interracial Families and the Racial Identification of Mixed-Race Children: Evidence from the Early Childhood Longitudinal Survey," *Social Forces*, 84, 1131-1157.

Choi, Y., T.W. Harachi, M.R. Gillmore, and R. Catalano, 2006, "Are Multiracial Adolescents at Greater Risk? Comparison of Rates, Patterns, and Correlates of Substance Use and Violence

between Monoracial and Multiracial Adolescents," *American Journal of Orthopsychiatry* 76(1) 86-97.

Conley, Dalton, 2001, *Honky*. Vintage: New York.

Cohn, David, 1944, "How the South Feels," *Atlantic Monthly*, January, 47-51.

Davis, F. James, 1991, *Who is Black? Ones Nation's Definition*, Pennsylvania State University Press, University Park, PA.

Fryer, Roland G., 2007a, "Guess Who's Been Coming to Dinner? Trends in Interracial Marriage over the 20th Century," *Journal of Economic Perspectives*, 21, 71-90.

Fryer, Roland G., 2007b, "A Model of Social Interactions and Endogenous Poverty Traps," *Rationality and Society*, 19, 335-366.

Fryer, Roland G., and Jackson, Matthew O., 2008, "A Categorical Model of Cognition and Biased Decision Making," *B.E. Journal of Theoretical Economics*, 8, Article 6.

Goldberg, Milton M., 1941, "A Qualification of the Marginal Man Theory," American Sociological Review, 6, 52-58.

Golovensky, David I., 1952, "The Marginal Man Concept: An Analysis and Critique," *Social Forces*, 30, 333-339.

Green, Arnold W., 1947, "A Re-Examination of the Marginal Man Concept," *Social Forces*, 26, 167-171.

Hahn, Robert A., Truman, Benedict I., and Barker, Nancy D., 1996, "Identifying Ancestry: The Reliability of Ancestral Identification in the United States by Self, Proxy, Interviewer, and Funeral Director," *Epidemiology,* 7, 75-80.

Harris, David R., 2002 "Does It Matter How We Measure? Racial Classification and the Characteristics of Multiracial Youth," In *The New Race Question*, edited by J. Perlmann and M. C. Waters, Russell Sage, New York.

Harris, David R., and Sim, Joseph J., 2002, "Who Is Multiracial? Assessing the Complexity of Lived Race," *American Sociological Review*, 67, 614-627.

Harris, David R., and Thomas, Justin L., 2002, "The Educational Costs of Being Multiracial: Evidence from a National Survey of Adolescents," PSC Research Report No. 02-0521.

Hill, Mark E., 2000, "Color Differences in the Socioeconomic Status of African American Men: Results of a Longitudinal Study," *Social Forces,*" 78, 1437-1460.

Hoffman, Frederick, 1896, *Race, Traits and Tendencies of the American Negro*, Macmilan, New York.

Kao, Grace, 1999, "Racial Identity and Academic Performance: An Examination of Biracial Asian and African Youth," *Journal of Asian American Studies*, 2, 223-249.

Keith, Verna M., and Herring, Cedric, 1991, "Skin Tone and Stratification in the Black Community," *American Journal of Sociology*, 97, 760-778.

Lang, Kevin and Manove, Michael, 2006, "Education and Labor-Market Discrimination" NBER Working Paper #12257.

Lee, Sharon M, 1993, "Racial Classifications in the U.S. Census: 1890-1990," *Ethnic and Racial Studies*, 16, 75-94.

Moran, Rachel, 2003, *The Regulation of Race and Romance,* University of Chicago Press, Chicago.

Nasim, A., F.Z Belgrave, R.J. Jagers, K.D. Wilson, K. Owens (2007) "The Moderating Effects of Culture on Peer Deviance and Alcohol Use among High-Risk African-American Adolescents" *Journal of Drug Education* 37(3) 335-363.

Romano, Reene, 2003, *Race Mixing: Black-White Marriage in Postwar America*, Harvard University Press, Cambridge, MA.

Park, Robert E., 1928, "Human Migration and the Marginal Man," *American Journal of Sociology,* 33, 881-893.

Park, Robert E., 1931, "Mentality of Racial Hybrids," *American Journal of Sociology*, 36, 534-551.

Roy, Andrew D., 1951 "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers*, 3, 135-46.

Reubeck, Christopher, Averett, Susan, and Bodenhorn, Howard, 2008, "Acting White or Acting Black: Mixed Race Adolescents' Identity and Behavior," NBER Working Paper 13793.

Stonequist, Everett V., 1935, "The Problem of the Marginal Man," *American Journal of Sociology*, 41, 1-12.

Stonequist, Everett V., 1937, *The Marginal Man: A Study in Personality and Culture Conflict,* Charles Scribener's Sons, New York.

Suskind, Ron, 1999, *A Hope in the Unseen: An American Odyssey from the Inner City to the Ivy League*, Broadway: New York.

Telles, Edward E., Lim, Nelson, 1998, "Does it Matter Who Answers the Race Question? Racial Classification and Income Inequality in Brazil," *Demography*, 35, 465-474.

Wilson, Anne, 1984, "'Mixed Race' Children in British Society: Some Theoretical Considerations," *The British Journal of Sociology*, 35, 42-61.

Xie, Yu, Goyette, Kimberly, 1997, "The Racial Identification Biracial Children with One Asian Parent: Evidence from the 1990 Census," *Social Forces*, 76, 547-570.

Appendix

The National Longitudinal Study of Adolescent Health (Add Health) is a nationally representative sample of 90,118 children, who attended grades 7 through 12 during the 1994-95 academic year. Thus far, information on these children has been collected at three separate points in time. During stage 1 of Wave I (September 1995 – April 1995) an in-school questionnaire was administered to the full sample of children, and 164 questionnaires were completed by school administrators. 20,745 computer assisted in-home interviews were conducted for a core sample of students plus oversamples during stage 2 of Wave I (April 1995 – December 1995). Additionally, a subset of 19,713 children completed the Add Health Picture Vocabulary Test (AHPVT) — an abridged and computerized version of the Peabody Picture Vocabulary Test-Revised; and 17,700 parents or legal guardians did so for the parent questionnaire. The data for Wave II was collected from April 1996 – August 1996 and contains information on 14,738 children and 128 schools. These children were, again, interviewed at home and are a subset of the ones interviewed during Wave I. So are the schools. In Wave III of Add Health (July 2001 – April 2001) 14,979 of the original Wave I in-home interview participants were re-interviewed and given the AHPVT again.[1] Moreover, participants were asked to sign a transcript release form and to provide a urine sample to be tested for sexually transmitted diseases (STDs).[2]

All of the variables described in this section have been standardized to have mean 0 and standard deviation 1 in the weighted population of white, black, and black-white children according to our strict child-based definition. Children of all other races have been dropped from the analysis. Unless otherwise noted, all of our regressions use missing value indicators, and are weighted by the cross-sectional weights supplied with Add Health for the survey from which the outcome variable was taken.

Below we describe how we combined and recoded some of the Add Health variables used in our analysis.

Home Environment Composite Measure
This variable was constructed by taking the mean of all standardized home environment variables, i.e. *Household Income, Not on Welfare, Father in Household, Parents Married, Mother's Age, Fewer Siblings, Mother is College Graduate, Mother Ever Married,* and *Years in Current Residence*, weighted by the number of observations in an individual's racial group. Before averaging missing values on a particular variable have been replaced with the racial group's mean on this variable

Physical Composite Measure
This variable was constructed by taking the mean of all standardized physical variables, i.e. *Birth Weight, Height, BMI, and Attractiveness*, weighted by the number of observations in an individual's racial group. Before averaging missing values on a particular variable have been replaced with the racial group's mean on this variable.

Achievement Composite Measure

---

[1] The Wave III data sets also contains 218 young adults who were interviewed during the pretest.
[2] Upon special requests researchers can also obtain data on 1,507 romantic partners of Wave III Add Health respondents collected in separate interviews.

This variable was constructed by taking the mean of all standardized achievement variables, i.e. *AHPVT Score, GPA, Never Repeated a Grade,* and *No Learning Disability*, weighted by the number of observations in an individual's racial group. Before averaging missing values on a particular variable have been replaced with the racial group's mean on this variable.

Behavior in School Composite Measure
This variable was constructed by taking the mean of all standardized school behavior variables, i.e. *Trouble with Teachers, Trouble Paying Attention, Trouble with Homework, Trouble with Students, Less Effort Schoolwork, Skipping School, and Ever Suspended or Expelled,* weighted by the number of observations in an individual's racial group. Before averaging missing values on a particular variable have been replaced with the racial group's mean on this variable.

Behavior Outside School Composite Measure
This variable was constructed by taking the mean of all standardized home behavior variables, i.e. *Watch TV, Drink, Smoke, Dare, Lie to Parents, Fight, Property Damage, Steal, Violent Acts, Sell Drugs, Encounter Violence, Ever Sex, Ever STD, and Ever Illegal Drugs*, weighted by the number of observations in an individual's racial group. Before averaging missing values on a particular variable have been replaced with the racial group's mean on this variable.

Household Income
On the parent questionnaire in Wave I respondents were asked, "About how much total income, before taxes did your family receive in 1994? Include your own income, the income of everyone else in your household, and income from welfare benefits, dividends, and all other sources." The values of this variable ranges from $0 to $999,000.

Not on Welfare
On the parent questionnaire in Wave I the following question was asked, "A you receiving public assistance, such as welfare?" A value of 0 indicates "yes," whereas 1 indicates "no" as an answer.

Father in Household
On the in-school questionnaire in Wave I respondents were asked, "Do you live with your biological father, stepfather, foster father, or adoptive father?" A value of 1 indicates "yes," whereas 0 indicates "no" as an answer.

Parents Married
On the parent questionnaire in Wave I the following question was asked, "What is your current marital status?" This variable takes on a value of 1 if the respondent reports to be married, and 0 otherwise.

Mother's Age
On the parent questionnaire in Wave I the following question was asked, "How old are you?" We use the answer to this question only if the respondent was female.

Mother is College Graduate

On the in-school questionnaire in Wave I respondents were asked with respect to their mother, "How far in school did she go?" The answer choices range from "she never went to school" to "professional training beyond a four-year college." We recoded this variable to take on a value of 1 if the child reported his mother to have at least "graduated from college or a university," and 0 otherwise.

Mother Ever Married
On the parent questionnaire in Wave I the following question was asked. "Have you ever been married?" For female respondents we coded this variable as 1 if the answer was "yes;" and as 0 if the respondent chose "no."

Years in Current Residence
During the in-home interview in Wave I respondents were asked about the month and year of their birth and, "How old were you when you moved here to your current residence?" This variables takes the difference between their imputed age and the answer to the latter question.

Birth Weight
On the parent questionnaire in Wave I respondents were asked for the birth weight (in pounds and ounces) of their child. We convert the answers into kilograms.

Height
During the in-home interview in Wave I respondents were asked, "What is your height in feet and inches?" The answer choices ranged from "4 feet, 0 inches" to "6 feet, 9 inches and over." We converted the respondents' weight into meters.

BMI
During the in-home interview in Wave I respondents were asked, "What is your weight?" The answers ranged from 50 to 430 pounds. We converted the answer in kilogram and constructed this variable using this answer and the one to the previously mentioned question according to the formula: $BMI = Weight / Height^2$.

Attractiveness
At the end of the in-home interview in Wave I the interviewer was asked to rate the physical attractiveness of the respondent on a scale from 1 to 5, where 0 indicates "very unattractive", and 5 indicates "very attractive."

AHPVT Score
Linked to a respondent's answers to the Wave I in-home interview is his score on the Add Health Picture Vocabulary Test. The test scores range from 0 to 87.

GPA
On the in-school questionnaire in Wave I respondents were asked for their grades in English, Math, History, and Science. Each of these variables range from 0 to 4.0. We take the average of the given answers.

Never Repeated a Grade

During the in-home interview in Wave I respondents were asked, "Have you ever repeated a grade or been held back a grade?" We recode this variable to take on a value of 1 if the given answer was "no," and 0 if it was "yes."

No Learning Disability
On the parent questionnaire in Wave I respondents were asked with respect to their child, "Does (he/she) have a specific learning disability, such as difficulties with attention, dyslexia, or some other reading, spelling, writing, or math disability?" We recoded this variables so that 0 indicates "yes" and 1 indicates "no."

Trouble with Teachers
On the in-school questionnaire in Wave I respondents were asked, "Since school started this year, how often have you had trouble getting along with your teachers?" The variable ranges from 0 to 4, where 0 indicates "never," and 4 indicates "everyday."

Trouble Paying Attention
On the in-school questionnaire in Wave I respondents were asked, "Since school started this year, how often have you had trouble paying attention in school?" The variable ranges from 0 to 4, where 0 indicates "never," and 4 indicates "everyday."

Trouble with Homework
On the in-school questionnaire in Wave I respondents were asked, "Since school started this year, how often have you had trouble getting your homework done?" The variable ranges from 0 to 4, where 0 indicates "never," and 4 indicates "everyday."

Trouble with Students
On the in-school questionnaire in Wave I respondents were asked, "Since school started this year, how often have you had trouble getting along with other students?" The variable ranges from 0 to 4, where 0 indicates "never," and 4 indicates "everyday."

Less Effort Schoolwork
On the in-school questionnaire in Wave I respondents were asked, "In general, how hard do you try to do your school work well?" The variable ranges from 0 to 4, where 0 indicates "I try very hard to do my best," and 4 indicates "I never try at all."

Skipping School
On the in-school questionnaire in Wave I respondents were asked, "In the past twelve months, how often did you skip school without an excuse?" The variable ranges from 0 to 6, where 0 indicates "never" and 6 indicates "nearly every day."

Never Suspended or Expelled
During the in-home interview in Wave I respondents were asked, "Have you ever received an out-of-school suspension from school?" and "Have you ever been expelled from school?" We coded the variable as 1 if the answer to both questions was "no," and as 0 if at least one question had been answered with "yes."

Watch TV
On the in-school questionnaire in Wave I respondents were asked, "Outside of school hours, about how much time do you spend watching television or video cassettes on an average school day?" The variable ranges from 0 to 4, where 0 indicates none, and 4 indicates "more than 4 hours."

Drinking
On the in-school questionnaire in Wave I respondents were asked, "In the past twelve months, how often did you drink beer, wine, or liquor?" The variable ranges from 0 to 6, where 0 indicates "never" and 6 indicates "nearly every day."

Smoking
On the in-school questionnaire in Wave I respondents were asked, "In the past twelve months, how often did you smoke cigarettes?" The variable ranges from 0 to 6, where 0 indicates "never" and 6 indicates "nearly every day."

Daring
On the in-school questionnaire in Wave I respondents were asked, "In the past twelve months, how often did you do something dangerous because you were dared to?" The variable ranges from 0 to 6, where 0 indicates "never" and 6 indicates "nearly every day."

Lie to Parents
On the in-school questionnaire in Wave I respondents were asked, "In the past twelve months, how often did you lie to your parents or guardians?" The variable ranges from 0 to 6, where 0 indicates "never" and 6 indicates "nearly every day."

Fight
On the in-school questionnaire in Wave I respondents were asked, "In the past year, how often have you gotten into a physical fight?" The variable ranges from 0 to 4, where 0 indicates "never" and 4 indicates "more than 7 times."

Property Damage
During the in-home interview in Wave I respondents were asked, "In the past 12 months, how often did you deliberately damage property that didn't belong to you?" The variable ranges from 0, implying "never," to 3, implying "5 or more times."

Steal
During the in-home interview in Wave I respondents were asked, "How often did you take something from a store without paying for it?" The variable ranges from 0, implying "never," to 3, implying "5 or more times."

Violent Acts
In three separate questions during the in-home interview in Wave I respondents were asked how many times each of the following things happened to them during the last 12 months: they got into a physical fight, they pulled a knife or gun on someone, they shot or stabbed someone. The

answer choices to each question range from 0 to 2, where 0 indicates "never", and 2 indicates "more than once." We coded this variable as the mean of the given answers.

Sell Drugs
During the in-home interview in Wave I respondents were asked, "How often did you sell marijuana or other drugs?" The variable ranges from 0, implying "never," to 3, implying "5 or more times."

Encounter Violence
In five separate questions during the in-home interview in Wave I respondents were asked how many times each of the following things happened to them during the last 12 months: they saw someone shoot or stab another person, someone pulled a knife or gun on them, someone shot them, someone cut or stabbed them, they were jumped. The answer choices to each question range from 0 to 2, where 0 indicates "never", and 2 indicates "more than once." We coded this variable as the mean of the given answers.

Ever Sex
During the in-home interview in Wave I respondents were asked whether they ever had vaginal intercourse. A value of 0 implies "no" as an answer, and 1 implies "yes."

Ever STD
In ten separate questions during the in-home interview in Wave I respondents were asked whether they had ever been told by a doctor or nurse that they had any of the following sexually transmitted diseases: chlamydia, syphilis, gonorrhea, HIV or AIDS, genital herpes, genital warts, trichomoniasis, hepatitis B, bacterial vaginosis, or non-gonococcal vaginitis. We coded this variable so that it takes a value of 0 if the respondent answered "no" to each question, and 1 if he answered with "yes" to at least one question.

Ever Illegal Drugs
In four separate questions during the in-home interview in Wave I respondents were asked how old they were when they tried the following drugs: marijuana, any kind of cocaine (including powder, freebase and crack cocaine), inhalants (such as glue, or solvents), and other type of illegal drug (such as LSD, PCP, ecstasy, mushrooms, speed, ice, heroin, or pills without a prescription). One answer choice to each question was "You never tried any [...]." This variable was coded as 0 if the respondent reported to have never tried any of the drugs to which these four questions refer, and as 1 if he reported to have tried at least one of them.

Mother Cares
On the in-school questionnaire children were asked with respect to their mother, "How much do you think she cares about you?" The variable ranges from 1 to 5, where 0 indicates "not at all" and 5 indicates "very much."

Father Cares
On the in-school questionnaire children were asked with respect to their father, "How much do you think he cares about you?" The variable ranges from 1 to 5, where 0 indicates "not at all" and 5 indicates "very much."

Close to People
On the in-school questionnaire children were asked how strongly they agree or the disagree with the statement, "I feel close to people at the school." The variable ranges from 1 to 5. We coded it such that 1 implies "strongly disagree" and 5 implies "strongly agree."

Feel Accepted
On the in-school questionnaire children were asked how strongly they agree or the disagree with the statement "I feel socially accepted." The variable ranges from 1 to 5. We coded it such that 1 implies "strongly disagree" and 5 implies "strongly agree."

Feel Loved
On the in-school questionnaire children were asked how strongly they agree or the disagree with the statement "I feel loved and wanted." The variable ranges from 1 to 5. We coded it such that 1 implies "strongly disagree" and 5 implies "strongly agree."

Not Depressed
On the in-school questionnaire children were asked, "In the last month, how often did you feel depressed or blue?" The variable ranges from 0 to 4. We recoded it such that 0 indicates "everyday" and 4 indicates "never."

Like Oneself
On the in-school questionnaire children were asked how strongly they agree or the disagree with the statement "I like myself just the way I am." The variable ranges from 1 to 5. We coded it such that 1 implies "strongly disagree" and 5 implies "strongly agree."

Chances Live to 35
On the in-school questionnaire children were asked, "What do you think are the chances you will live to age 35?" The variable ranges from 0 to 8, where 0 indicates "no chance" and 8 indicates "it will happen."

In addition to Add Health we also utilize data from the 2000 US Census. We use the integrated public use 5% sample.[3] The 2000 Census allowed respondents to check more than one race. Based on their answers the Census Bureau then assigns them one of more than 150 racial categories in its detailed version of the race variable. The Census Bureau codes an individual as "White and Black" if he reports to be only of white and black origin. Since this coincides with our strict definition, we code an individual as mixed race if the Census Bureau does so, i.e. if it codes him as "Black/White."

The Integrated Public Use Microdata Series (IPUMS) allows researchers to link characteristics of different household members to individuals with same household identifier. This lets us link an individual to the race of his parents, and explore a different definition of mixed race. Under this definition a child is coded as a member of any single race group if and only if both parents are of that race as well. If a child has two mixed parents, or one mixed parent and one who is either white or black, or one white and one black parent, he is coded as mixed

---

[3] Our Census data has been obtained from http://usa.ipums.org.

under this alternative definition.

All of our adult outcome variables are restricted to individuals who were between 17 and 30 years old at the time the Census was taken, and are standardized to have mean 0 and a standard deviation of 1 in this (unweighted) population. Our home environment variables are restricted to children between the age of 10 and 19, except for cases in which we explicitly state a different age range, and are standardized in the same manner as the adult outcomes.

Unless otherwise noted, all our Census data regressions are unweighted and include missing value identifiers.

Household Income
This variable reports the sum of the total money income of all household members living in the household at the time of the Census, who are 15 years old and above.

Not on Welfare
The Census reports the pre-tax income from public assistance programs, referred to as "welfare," during the previous year for each household member. We coded this variable as 0 if the individual received any income from welfare, and as 1 if no welfare payments were received.

Father in Household
The Census Bureau provides a constructed variable indicating whether the father of a person lives in the same household, and if so the father's person number. Our variable takes on a value of 0 if the person's father does not live in the same household, and 1 if he does.

Parents Married
We retrieve the marital status of a child's parents, and code this variable as 1 if the parents who are present in the household report to be married, and 0 if they don't.

Mother's Age
We retrieve the age of a child's mother and use it for this variable.

Mother is College Graduate
We retrieve the educational attainment of a child's mother; and code this variable as 1 if she has at least obtained a bachelor's degree, and as 0 otherwise.

Mother Ever Married
After linking a mother's current marital status to her child, we let this variable take on a value of 1 if she reports "Married, spouse present," "Married, spouse absent," "Separated," "Divorced," or "Widowed;" and as 0 if her current status is "Never married/single."

Not Migrated within Last 5 Years
The Census asked people whether they had lived in the same, or a different house 5 years earlier. If an individual answered to have lived in the same house, he is coded as 1. He is coded as 0 if he reported to have lived in a different house.
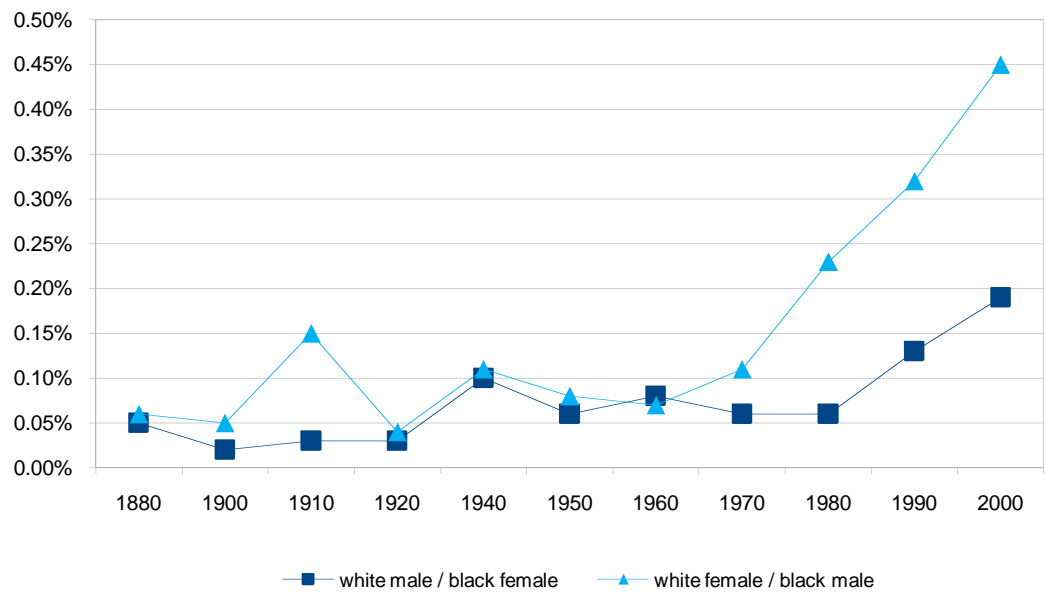
Table 1: Summary Statistics by Race

| Variable | Strict Self-Declared Definition | | | | Source |
	Full Sample	White	Black	Mixed	
Demographics					
Female	.499	.495	.515	.484	School Survey
	(.500)	(.500)	(.500)	(.501)	
Age (in years), Wave I	14.983	15.015	14.869	14.821	School Survey
	(1.742)	(1.738)	(1.757)	(1.651)	
Born in US	.979	.982	.968	.959	School Survey
	(.143)	(.132)	(.176)	(.199)	
West	.118	.133	.061	.229	School Administrator Survey
	(.323)	(.340)	(.239)	(.421)	
Midwest	.300	.327	.203	.310	School Administrator Survey
	(.458)	(.469)	(.402)	(.463)	
Northeast	.146	.167	.072	.159	School Administrator Survey
	(.353)	(.373)	(.258)	(.367)	
South	.435	.372	.665	.302	School Administrator Survey
	(.496)	(.483)	(.472)	(.460)	
Home Environment					
Household Income (in Dollar)	48,029	51,547	29,925	29,767	Parent Questionnaire
	(48,046)	(47,838)	(45,472)	(18,655)	
Not on Welfare	.916	.940	.794	.844	Parent Questionnaire
	(.278)	(.237)	(.404)	(.366)	
Father in Household	.769	.833	.534	.544	School Survey
	(.422)	(.373)	(.499)	(.499)	
Parents Married	.725	.787	.426	.245	Parent Questionnaire
	(.446)	(.409)	(.495)	(.434)	
Mother's Age (in years)	41.588	41.596	41.600	39.660	Parent Questionnaire
	(6.396)	(5.964)	(8.273)	(3.737)	
Mother is College Graduate	.319	.321	.308	.370	School Survey
	(.466)	(.467)	(.462)	(.484)	
Mother Ever Married	.955	.987	.800	.916	Parent Questionnaire
	(.206)	(.115)	(.400)	(.281)	
Years in Current Residence	7.817	8.065	6.719	4.654	Home Interview Wave I
	(5.770)	(5.743)	(5.782)	(4.151)	
Physical					
Birth Weight (in kilogram)	3.385	3.415	3.222	3.362	Parent Questionnaire
	(.545)	(.539)	(.547)	(.443)	
Weight (in kilogram)	64.390	63.890	66.698	66.006	Home Interview Wave I
	(15.811)	(15.681)	(16.229)	(15.204)	
Height (in meters)	1.692	1.692	1.692	1.707	Home Interview Wave I
	(.106)	(.106)	(.108)	(.096)	
BMI	22.328	22.136	23.239	22.601	(constructed)
	(4.356)	(4.244)	(4.744)	(4.618)	
Attractiveness	.000	.007	-.041	.414	Home Interview Wave I
	(1.000)	(.998)	(1.008)	(.904)	
Achievement					
No Learning Disability	.861	.863	.850	.961	Parent Questionnaire
	(.346)	(.344)	(.357)	(.197)	
AHPVT	.000	.164	-.756	-.161	Home Interview Wave I
	(1.000)	(.913)	(1.037)	(1.011)	
GPA	2.865	2.934	2.592	2.720	School Survey
	(.793)	(.788)	(.755)	(.783)	
Never Repeated Grade	.797	.824	.673	.653	Home Interview Wave I
	(.403)	(.381)	(.469)	(.480)	
Behavior in School					
Trouble with Teacher	.000	.051	-.198	.003	School Survey
	(1.000)	(.967)	(1.098)	(.980)	
Trouble Paying Attention	.000	.031	-.119	-.127	School Survey
	(1.000)	(.972)	(1.095)	(.990)	
Trouble with Homework	.000	.044	-.170	-.132	School Survey
	(1.000)	(.976)	(1.071)	(.955)	

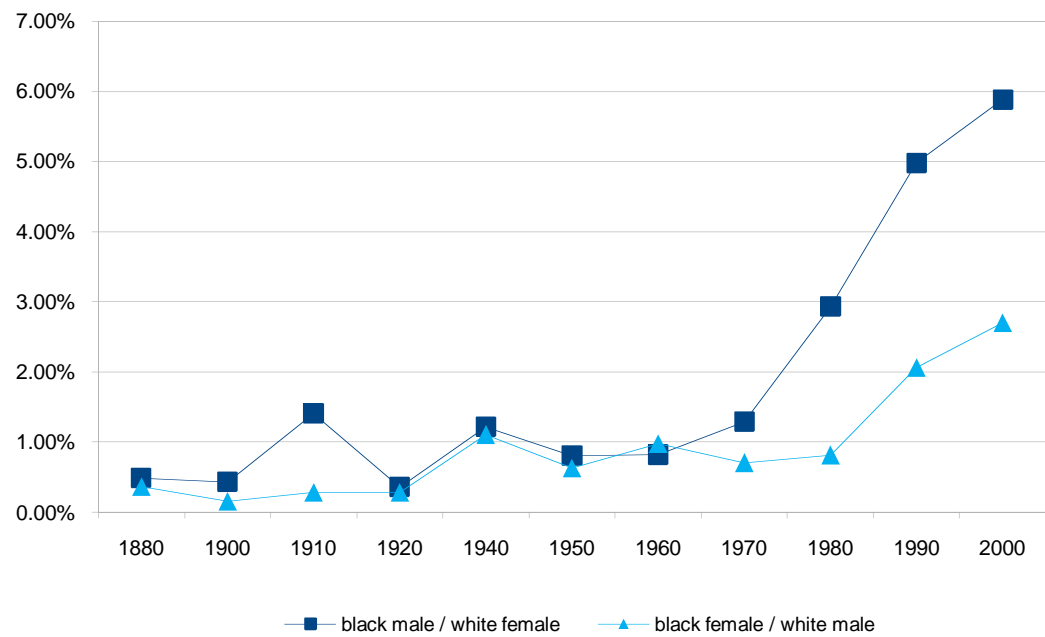|					
Trouble with Students	.000	.065	-.247	-.214	School Survey
	(1.000)	(.966)	(1.086)	(1.053)	
Effort Schoolwork	.000	-.058	.224	-.095	School Survey
	(1.000)	(1.002)	(.956)	(1.089)	
Skipping School	.000	-.013	.054	-.207	School Survey
	(1.000)	(1.002)	(.982)	(1.319)	
Never Suspended or Expelled	.000	.104	-.470	-.547	Home Interview Wave I
	(1.000)	(.939)	(1.123)	(1.132)	
Behavior Outside School					
Watch TV	.000	.140	-.527	-.255	School Survey
	(1.000)	(.947)	(1.021)	(.987)	
Drinking	.000	-.026	.103	-.145	School Survey
	(1.000)	(1.004)	(.975)	(1.055)	
Smoking	.000	-.078	.303	-.060	School Survey
	(1.000)	(1.055)	(.667)	(1.073)	
Daring	.000	-.057	.227	-.191	School Survey
	(1.000)	(1.025)	(.856)	(1.154)	
Lie to Parents	.000	.005	-.015	-.203	School Survey
	(1.000)	(.985)	(1.054)	(1.125)	
Fight	.000	.019	-.071	-.241	School Survey
	(1.000)	(.997)	(1.007)	(1.130)	
Property Damage	.000	-.032	.153	-.196	Home Interview Wave I
	(1.000)	(1.033)	(.816)	(1.020)	
Steal	.000	-.011	.063	-.572	Home Interview Wave I
	(1.000)	(1.013)	(.920)	(1.398)	
Violent Acts	.000	.058	-.252	-.629	Home Interview Wave I
	(1.000)	(.947)	(1.166)	(1.601)	
Sell Drugs	.000	.004	-.020	-.019	Home Interview Wave I
	(1.000)	(.989)	(1.048)	(1.041)	
See Violence	.000	.065	-.295	-.303	Home Interview Wave I
	(1.000)	(.933)	(1.213)	(1.290)	
Ever Sex	.000	.080	-.364	-.381	Home Interview Wave I
	(1.000)	(.980)	(1.010)	(1.016)	
Ever STD	.000	.065	-.295	-.313	Home Interview Wave I
	(1.000)	(.795)	(1.599)	(1.642)	
Ever Illegal Drugs	.000	-.018	.099	-.526	Home Interview Wave I
	(1.000)	(1.007)	(.954)	(1.081)	
Psychological Variables					
Mother Cares	.000	-.016	.062	-.017	School Survey
	(1.000)	(1.012)	(.947)	(1.069)	
Father Cares	.000	.009	-.044	-.298	School Survey
	(1.000)	(.978)	(1.111)	(1.352)	
Close to People	.000	.042	-.170	-.095	School Survey
	(1.000)	(.985)	(1.042)	(.961)	
Feel Accepted	.000	-.024	.098	-.002	School Survey
	(1.000)	(.997)	(1.007)	(.926)	
Feel Loved	.000	-.030	.124	-.068	School Survey
	(1.000)	(.992)	(1.022)	(1.023)	
Not Depressed	.000	-.035	.145	-.185	School Survey
	(1.000)	(.995)	(1.006)	(1.103)	
Like Oneself	.000	-.071	.292	.021	School Survey
	(1.000)	(.999)	(.950)	(1.061)	
Chances Live to 35	.000	.054	-.216	-.013	School Survey
	(1.000)	(.947)	(1.163)	(1.065)	

Notes: The entries are means and standard deviations of student-level data for those students in Add Health who are either white, black, or mixed race according to our strict child-level definition. All behavior and psychological variables, AHPVT, and Attractiveness have been normalized to have mean of 0 and standard deviation of 1 in the weighted sample of white, black, and mixed children. The final column shows the survey from which the variable was extracted. In all cases, sample weights provided with Add Health are used in the calculations. See the appendix for further details on the construction of our sample.

Figure 1: Percent of Whites (A) and Blacks (B) Marrying out of Race, by Gender and
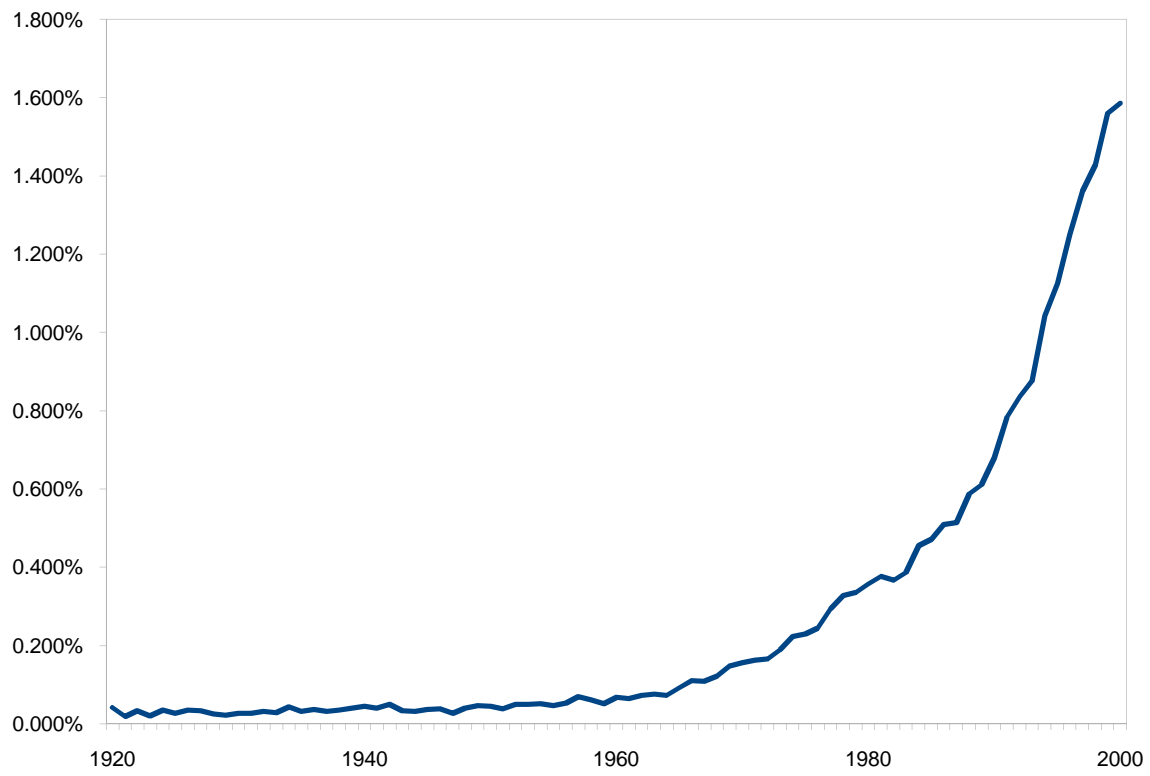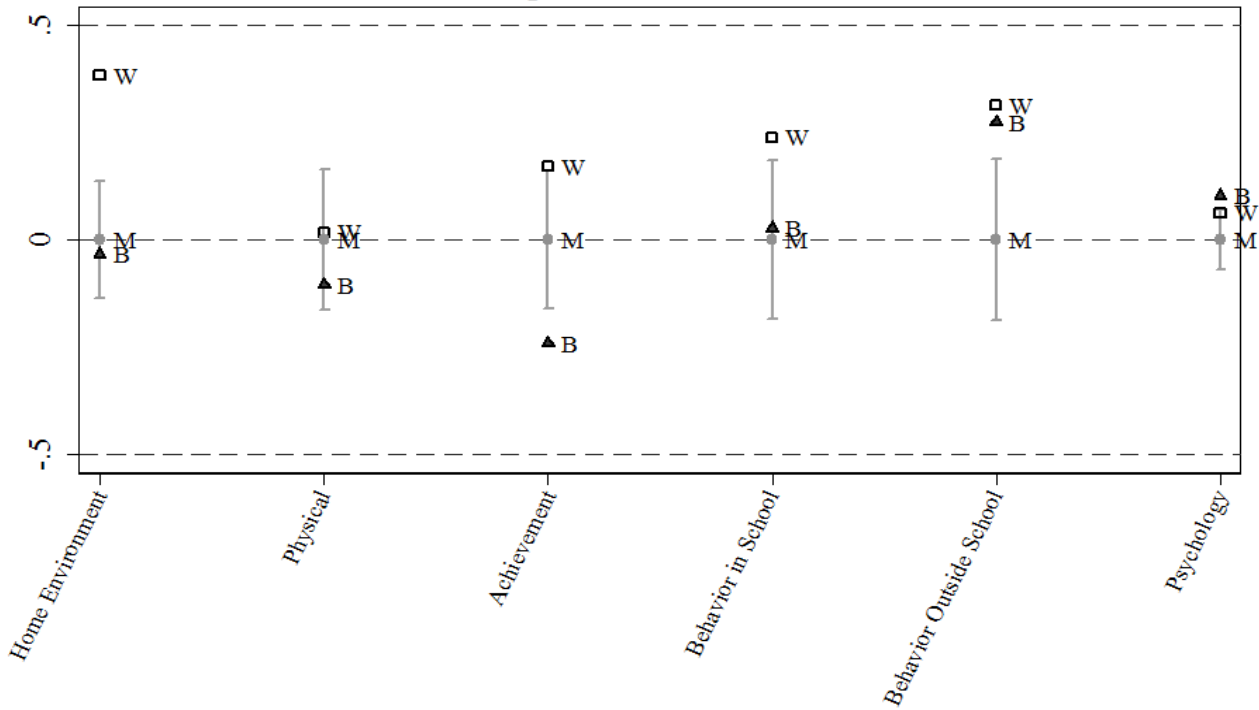Race of Spouse (as percent of all marriages)

(A)

— white male / black female   — white female / black male

(B)

— black male / white female   — black female / white male

Note: Both figures are from Fryer (2007a).

Figure 2: Black-White Births as Percentage of all Black and White Births

Figure 3: Overview

Notes: This figure shows the coefficients from our composite outcome measures for blacks and whites. Black-white mixed race children are the omitted category. Every regression in each composite measure includes indicator variables for a child's race and controls for age, age squared, gender, whether the student was born abroad, and region. Each composite measure was constructed by taking the mean of all standardized variables within one category weighted by the number of observations in an individual's racial group. Before averaging, missing values on a particular variable have been replaced with the racial group's mean on this variable. See the appendix for further details.
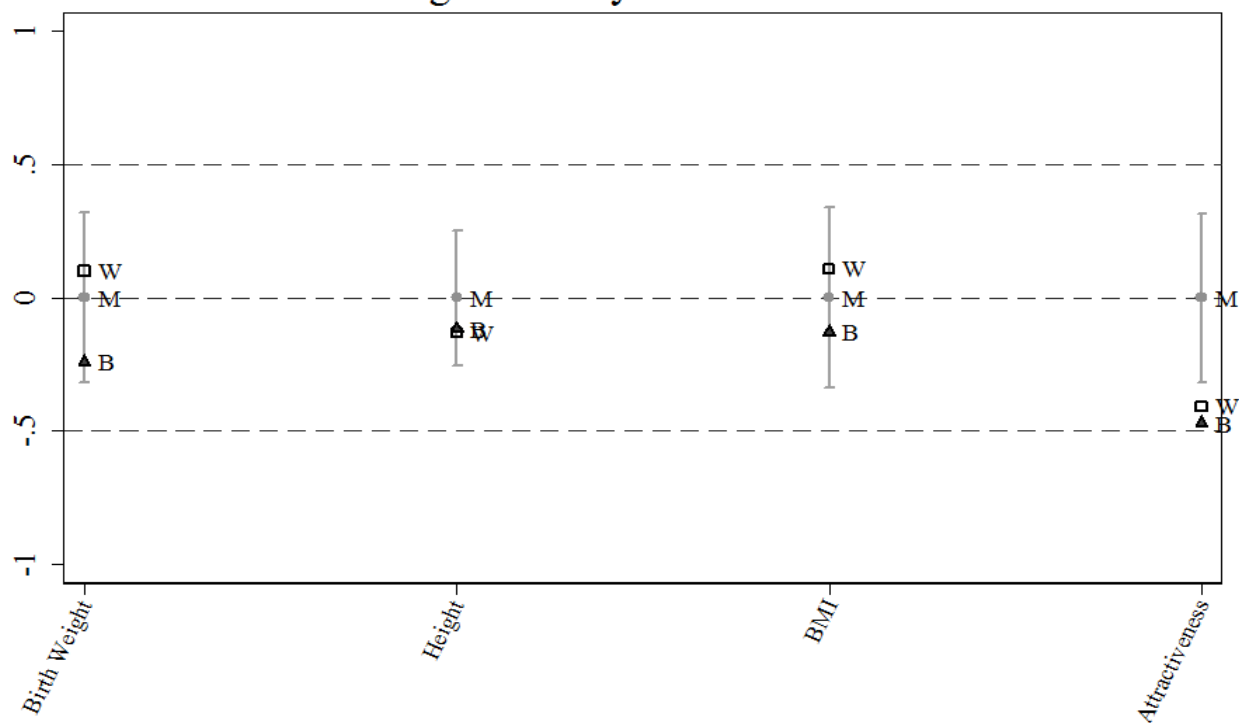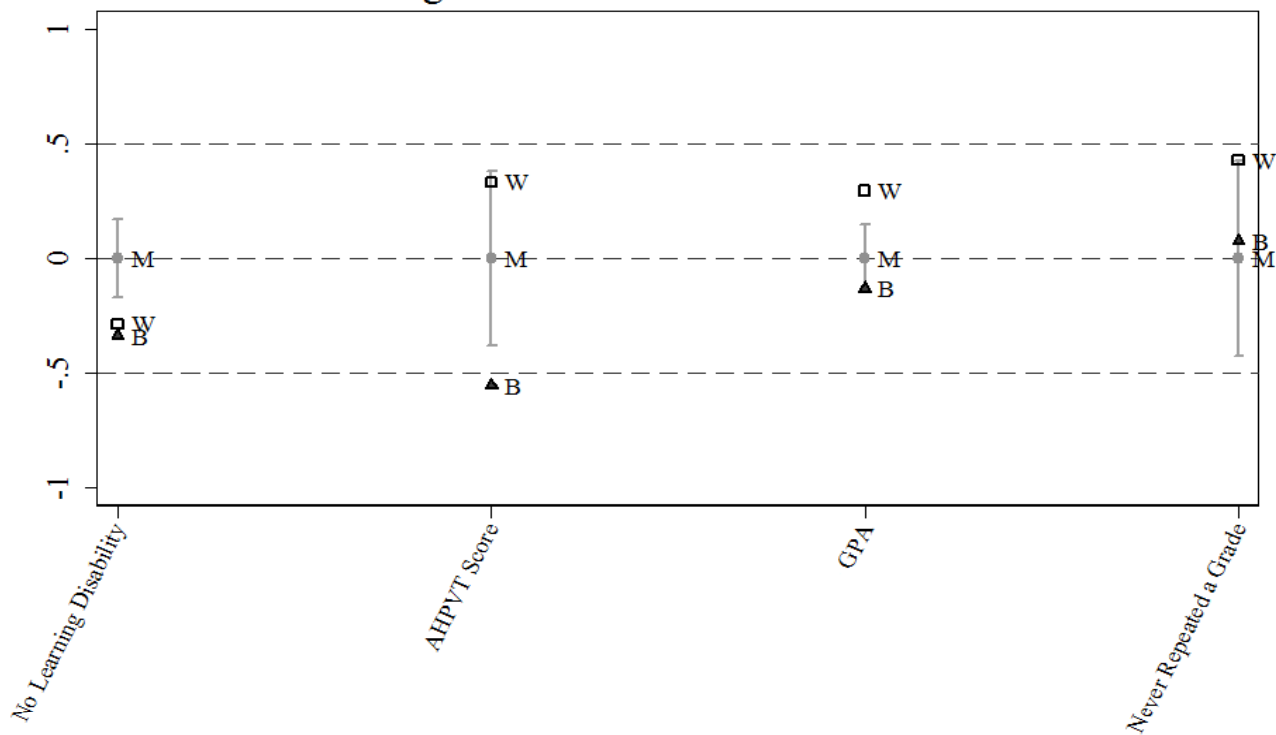
Figure 4: Home Environment Variables
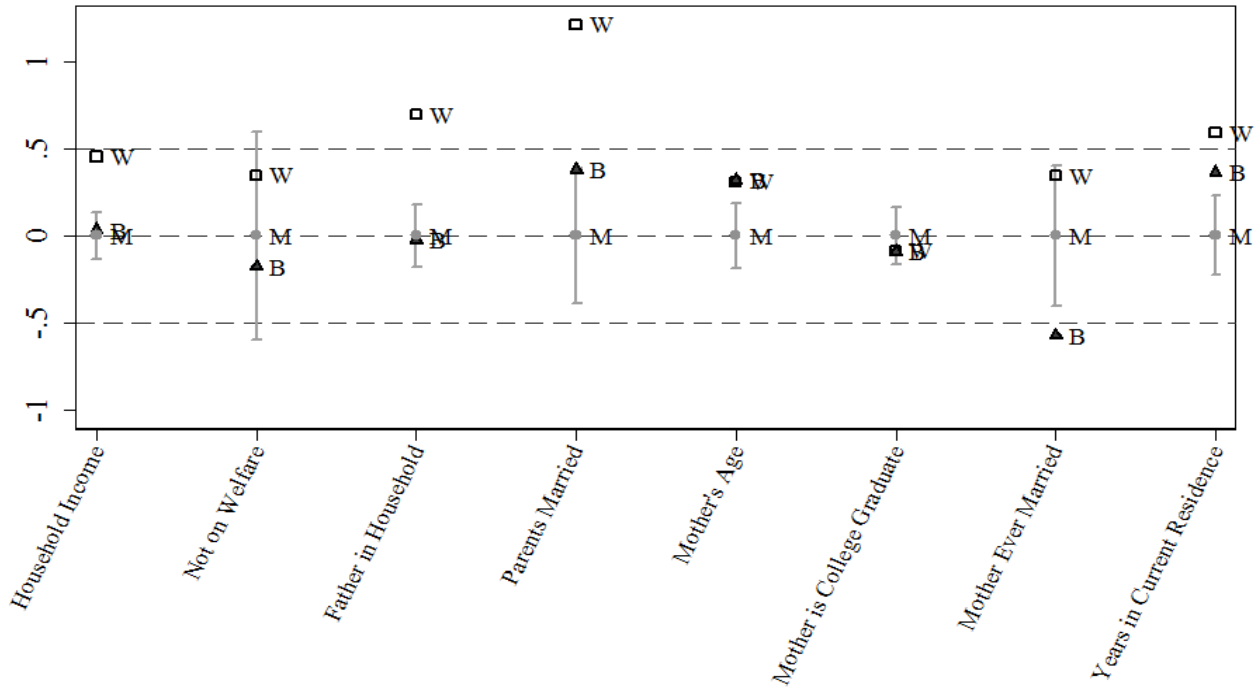
Notes: This figure shows the coefficients from regressions of our home environment variables on the race dummies for blacks and whites. Black-white mixed race children are the omitted category. All variables are taken from Add Health and have been standardized to have a mean of 0 and a standard deviation of 1 in our sample of white, black, and mixed race children. Every regression is weighted and includes indicator variables for a child's race and controls for age, age squared, gender, whether the student was born abroad, and region. See the appendix for further details.

Figure 5: Home Environment

Census
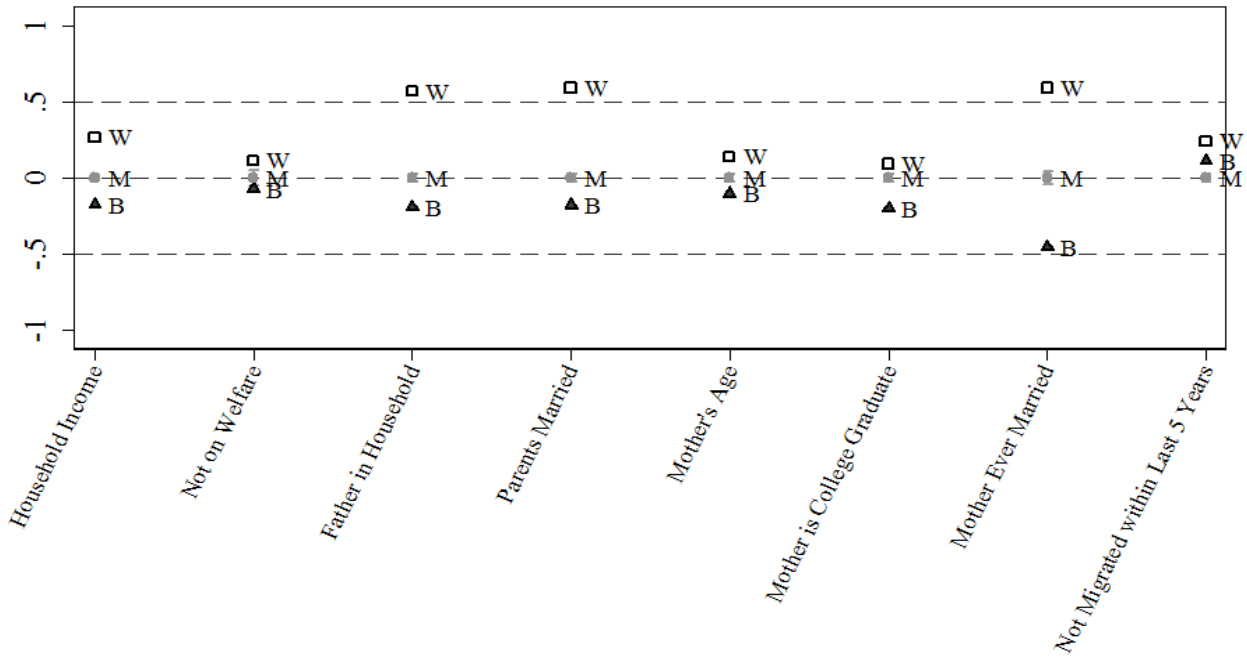
Notes: This figure shows the coefficients from regressions of our home environment variables on the race dummies for blacks and whites. Black-white mixed race children are the omitted category. All variables are taken from 2000 US Census and have been standardized to have a mean of 0 and a standard deviation of 1 in our sample of white, black, and mixed race children. Every regression includes indicator variables for a child's race and controls for age, age squared, gender, whether the student was born abroad, and region. See the appendix for further details.
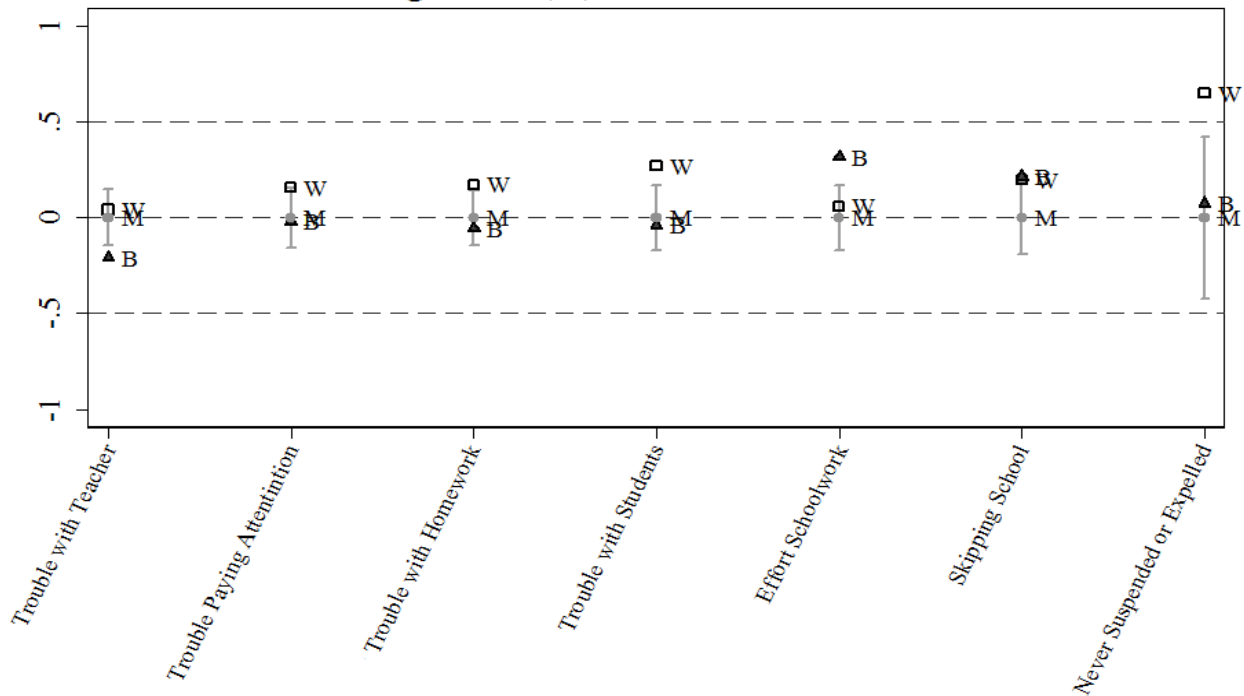
Figure 6: Physical Variables

Notes: This figure shows the coefficients from regressions of our physical variables on the race dummies for blacks and whites. Black-white mixed race children are the omitted category. All variables are taken from Add Health and have been standardized to have a mean of 0 and a standard deviation of 1 in our sample of white, black, and mixed race children. Every regression is weighted and includes indicator variables for a child's race and controls for age, age squared, gender, whether the student was born abroad, and region. See the appendix for further details.
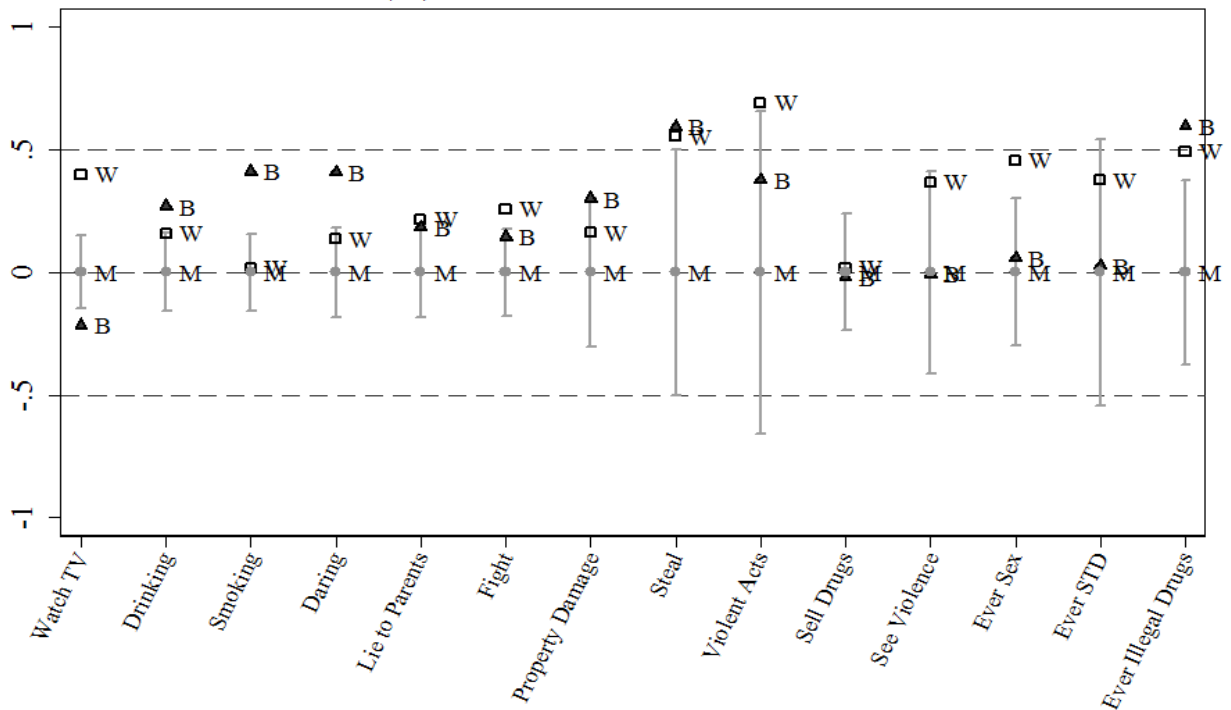
Figure 7: Achievement Variables

Notes: This figure shows the coefficients from regressions of our achievement variables on the race dummies for blacks and whites. Black-white mixed race children are the omitted category. All variables are taken from Add Health and have been standardized to have a mean of 0 and a standard deviation of 1 in our sample of white, black, and mixed race children. Every regression is weighted and includes indicator variables for a child's race and controls for age, age squared, gender, whether the student was born abroad, and region. See the appendix for further details.

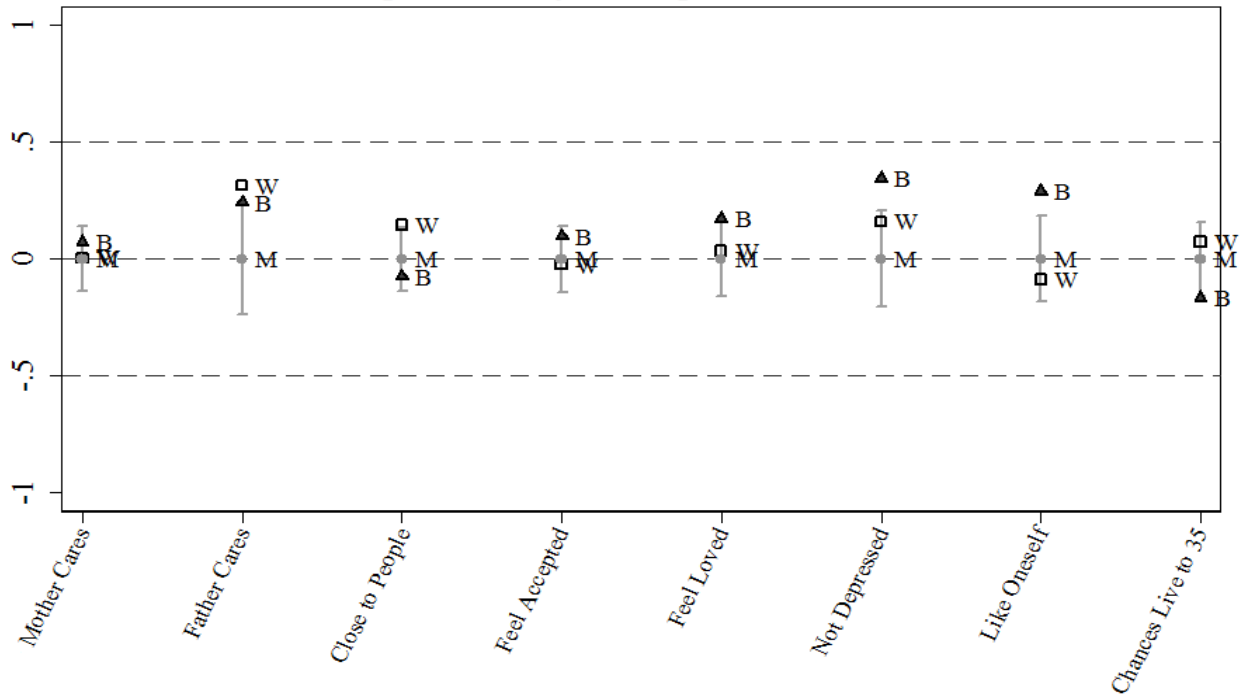## Figure 8: (A) Behavior in School

(B) Behavior Outside School

Notes: This figure shows the coefficients from regressions of our behavior variables on the race dummies for blacks and whites. Black-white mixed race children are the omitted category. All variables are taken from Add Health and have been standardized to have a mean of 0 and a standard deviation of 1 in our sample of white, black, and mixed race children. Every regression is weighted and includes indicator variables for a child's race and controls for age, age squared, gender, whether the student was born abroad, and region. See the appendix for further details.

Figure 9: Psychological Variables

Notes: This figure shows the coefficients from regressions of our psychological variables on the race dummies for blacks and whites. Black-white mixed race children are the omitted category. All variables are taken from Add Health and have been standardized to have a mean of 0 and a standard deviation of 1 in our sample of white, black, and mixed race children. Every regression is weighted and includes indicator variables for a child's race and controls for age, age squared, gender, whether the student was born abroad, and region. See the appendix for further details.