The Plight of Mixed Race Kids

Roland G. Fryer, Jr.
Harvard University and NBER

Lisa Kahn
Yale School of Management

Steven D. Levitt
University of Chicago and American Bar Foundation

Jörg L. Spenkuch
University of Chicago

First Draft: August 2005
This Version: May 2008

* We are grateful to Gary Becker, Edward Glaeser, Claudia Goldin, Lawrence Katz, Kevin Murphy, Lawrence Summers, and a focus group of nine mixed race undergraduates at Harvard University for many useful conversations. Katherine Barghaus and Ethan Lieber provided outstanding research assistance. Financial Support was provided by the National Science Foundation, Harvard University’s Milton Fund (Fryer), the American Inequality Lab (Fryer), and the Sherman Shapiro Research Fund (Levitt). Correspondence can be addressed to Fryer at Department of Economics, Harvard University, 1805 Cambridge Street, Cambridge MA, 02138 (e-mail: rfryer@fas.harvard.edu); or Levitt at Department of Economics, University of Chicago, 1126 E. 59th Street, Chicago IL, 60637 (e-mail: slevitt@uchicago.edu).
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 kids during their adolescence and early adulthood. As one might expect, on a host of background and achievement characteristics, mixed race kids fall in between whites and blacks. When it comes to engaging in risky/anti-social adolescent behavior, however, mixed race kids 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 kids – 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 kids are outliers on risky behaviors nor why these behaviors are not strongly influenced by the racial composition at their school.
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). By 2000, interracial marriages increased dramatically; 5.9% of married black men chose a white bride and 2.7% of black women chose a white husband (Fryer 2007a).

\(^1\) 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.”
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.\(^3\) 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).\(^4\)

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 kids 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 regressions including indicator variables for a child’s race and controlling for age, age squared, gender, whether the student was born abroad, and region. Figure 3 shows the coefficients from

\(^3\) 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.

\(^4\) Romano (2003) reports that whites considering interracial marriage are nearly always asked the question “What about the Children?”
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 kids 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 kids 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.5

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 kids. We reinterpret and formalize the marginal man hypothesis using a 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 kids adhere to black norms and white kids adhere to white norms. Mixed race kids have a choice, they can choose to associate with black kids and adopt their norms, befriend white kids and adopt their norms, or both. It is this outside option that gives mixed race kids a higher cost of

5 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 kids are weakly above blacks in nearly all outcomes in both datasets.
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. 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. Ruebeck et al. (2008) is the paper most similar to ours. 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

---

6 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)).

7 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 exposted by Stonequist (1935, 1937) and Park (1928, 1931) had no predictions for Asian-White intermixing.

8 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.
“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 monoraciauls.

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 kids. Section III describes the plight of mixed race kids 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 Kids

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 too small to be useful, 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 Longitudinal Study of Adolescent Health (Add Health) – the primary data set used in this paper.

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

---

9 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.
The original data collection took place in 1995, with Wave II done in 1996 and Wave III carried out in 2001. 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. We use an array 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). We conceptualize the tradeoffs in 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

---

10 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%.

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

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


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 kids 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).
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. A child-based approach is more direct, but relies on the child’s self-identification.

Under our preferred “strict” definition at the kid 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. Nonetheless, the resulting frequency of mixed race children is 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.

---

14 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.

15 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.
on request. While results for particular variables do, of course, vary across these definitions, the same basic patterns are present 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. Achievement is proxied by score on the Add Health Picture Vocabulary Test (AHPVT), grade point average (GPA), whether or not a student has repeated a grade, and whether the student has a learning disability. 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.

16 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.

17 Harris and Thomas (2002) analyze three of these outcomes: grade point average, grade retention, and AHPVT.

18 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.
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.

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 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 Kids

In this section, we describe basic facts about the plight of mixed race kids 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. 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:

\[ \text{outcome}_i = a + R_i \beta + X_i \gamma + e_i, \]

where \( R_i \) represents a mutually exclusive and collectively exhaustive set of racial identifiers 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. In each of the figures that follow, we report the \( \beta \) coefficients for blacks, whites, normalizing mixed race black-white coefficients to be zero, so the other estimates are relative to mixed race children. 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 kids 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.

---

19 Students of all other races are dropped from the analysis.

20 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. Yet, 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.
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. 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 kids 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 kids 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 kids exhibit worse behavior than both of their single race counterparts. Behavior at school by mixed race kids 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 kids 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 race kids roughly occupy the lower envelope of good behaviors of blacks and whites.

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 35.21

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.

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

21 Under the parent-strict definition of mixed race kids, these differences become quite stark.
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. Discrimination-based models are unlikely to provide an answer. In all models we know of 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, discrimination models cannot explain our findings. 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 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. For Stonequist (1935) and Park (1928), being multiracial was

22 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.

23 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
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. 24

Formalizing the marginal man hypothesis via a two sector Roy model

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

---

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.

24 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)
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 $\arg\max_{x \in X} g(x - t)$. 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 kids 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 an 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 monoracial. 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 kids.\footnote{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.}

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 kids engage in more risky behavior than any of their single race peers can be rationalized through $p_j(x)$. Because mixed race kids 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).\footnote{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 kids 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 $q(q,x)$ 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:

\[
\max_{x_b, x_w} \left\{ g(x_w - t_w, x_b - t_b) + \lambda_w \int \phi_x(q) \phi(q, x_b, x_w) dq + \lambda_b \int \phi_x(q) \phi(q, x_b, x_w) dq \right\},
\]

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 kids are “flatter” than those for the monoracial groups, then the model explains why mixed race kids over invest in bad behaviors. As in the Roy model, for mixed race kids, 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 \( \lambda_w \) and \( \lambda_b \) as the fraction of blacks and whites in an individual’s school, then under the conformity model we would expect mixed race kids 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). A similar, but even more pronounced picture emerges if we split up the mixed race kids 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 kids 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 kids act (weakly) more black. When there are few blacks present, the costs of acting black for mixed race kids 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 less costly for a mixed race child to prove he can fight when the only opponents are whites. 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

27 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.

28 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 (Susskind 1999).
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 kids 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.
References


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. Moreover, participants were asked to sign a transcript release form and to provide a urine sample to be tested for sexually transmitted diseases (STDs).

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 range 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 variable 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: \( \text{BMI} = \frac{\text{Weight}}{\text{Height}^2} \).

**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 variable 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 vaginosisis, 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. 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 kid, 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>
<thead>
<tr>
<th>Demographics</th>
<th>Full Sample</th>
<th>White</th>
<th>Black</th>
<th>Mixed</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>.499</td>
<td>.495</td>
<td>.515</td>
<td>.484</td>
<td>School Survey</td>
</tr>
<tr>
<td></td>
<td>(.500)</td>
<td>(.500)</td>
<td>(.500)</td>
<td>(.501)</td>
<td></td>
</tr>
<tr>
<td>Age (in years), Wave I</td>
<td>14.983</td>
<td>15.015</td>
<td>14.869</td>
<td>14.821</td>
<td>School Survey</td>
</tr>
<tr>
<td></td>
<td>(1.742)</td>
<td>(1.738)</td>
<td>(1.757)</td>
<td>(1.651)</td>
<td></td>
</tr>
<tr>
<td>Born in US</td>
<td>.979</td>
<td>.982</td>
<td>.968</td>
<td>.959</td>
<td>School Survey</td>
</tr>
<tr>
<td></td>
<td>(.143)</td>
<td>(.132)</td>
<td>(.176)</td>
<td>(.199)</td>
<td></td>
</tr>
<tr>
<td>West</td>
<td>.118</td>
<td>.133</td>
<td>.061</td>
<td>.229</td>
<td>School Administrator Survey</td>
</tr>
<tr>
<td></td>
<td>(.323)</td>
<td>(.400)</td>
<td>(.239)</td>
<td>(.421)</td>
<td></td>
</tr>
<tr>
<td>Midwest</td>
<td>.300</td>
<td>.327</td>
<td>.203</td>
<td>.310</td>
<td>School Administrator Survey</td>
</tr>
<tr>
<td></td>
<td>(.458)</td>
<td>(.469)</td>
<td>(.402)</td>
<td>(.463)</td>
<td></td>
</tr>
<tr>
<td>Northeast</td>
<td>.146</td>
<td>.167</td>
<td>.072</td>
<td>.159</td>
<td>School Administrator Survey</td>
</tr>
<tr>
<td></td>
<td>(.353)</td>
<td>(.373)</td>
<td>(.258)</td>
<td>(.367)</td>
<td></td>
</tr>
<tr>
<td>South</td>
<td>.435</td>
<td>.372</td>
<td>.665</td>
<td>.302</td>
<td>School Administrator Survey</td>
</tr>
<tr>
<td></td>
<td>(.496)</td>
<td>(.483)</td>
<td>(.472)</td>
<td>(.460)</td>
<td></td>
</tr>
<tr>
<td>Home Environment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Household Income (in Dollar)</td>
<td>48,029</td>
<td>51,547</td>
<td>29,925</td>
<td>29,767</td>
<td>Parent Questionnaire</td>
</tr>
<tr>
<td></td>
<td>(48,046)</td>
<td>(47,838)</td>
<td>(45,472)</td>
<td>(18,655)</td>
<td></td>
</tr>
<tr>
<td>Not on Welfare</td>
<td>.916</td>
<td>.940</td>
<td>.794</td>
<td>.844</td>
<td>Parent Questionnaire</td>
</tr>
<tr>
<td></td>
<td>(.278)</td>
<td>(.237)</td>
<td>(.404)</td>
<td>(.366)</td>
<td></td>
</tr>
<tr>
<td>Father in Household</td>
<td>.769</td>
<td>.833</td>
<td>.534</td>
<td>.544</td>
<td>School Survey</td>
</tr>
<tr>
<td></td>
<td>(.422)</td>
<td>(.373)</td>
<td>(.499)</td>
<td>(.499)</td>
<td></td>
</tr>
<tr>
<td>Parents Married</td>
<td>.725</td>
<td>.787</td>
<td>.426</td>
<td>.245</td>
<td>Parent Questionnaire</td>
</tr>
<tr>
<td></td>
<td>(.446)</td>
<td>(.409)</td>
<td>(.495)</td>
<td>(.434)</td>
<td></td>
</tr>
<tr>
<td>Mother's Age (in years)</td>
<td>41.588</td>
<td>41.596</td>
<td>41.600</td>
<td>39.660</td>
<td>Parent Questionnaire</td>
</tr>
<tr>
<td></td>
<td>(6.396)</td>
<td>(5.964)</td>
<td>(8.273)</td>
<td>(3.737)</td>
<td></td>
</tr>
<tr>
<td>Mother is College Graduate</td>
<td>.319</td>
<td>.321</td>
<td>.308</td>
<td>.370</td>
<td>School Survey</td>
</tr>
<tr>
<td></td>
<td>(.466)</td>
<td>(.467)</td>
<td>(.462)</td>
<td>(.484)</td>
<td></td>
</tr>
<tr>
<td>Mother Ever Married</td>
<td>.955</td>
<td>.987</td>
<td>.800</td>
<td>.916</td>
<td>Parent Questionnaire</td>
</tr>
<tr>
<td></td>
<td>(.206)</td>
<td>(.115)</td>
<td>(.400)</td>
<td>(.281)</td>
<td></td>
</tr>
<tr>
<td>Years in Current Residence</td>
<td>7.817</td>
<td>8.065</td>
<td>6.719</td>
<td>4.654</td>
<td>Home Interview Wave I</td>
</tr>
<tr>
<td></td>
<td>(5.770)</td>
<td>(5.743)</td>
<td>(5.782)</td>
<td>(4.151)</td>
<td></td>
</tr>
<tr>
<td>Physical</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Birth Weight (in kilogram)</td>
<td>3.385</td>
<td>3.415</td>
<td>3.222</td>
<td>3.362</td>
<td>Parent Questionnaire</td>
</tr>
<tr>
<td></td>
<td>(.545)</td>
<td>(.543)</td>
<td>(.547)</td>
<td>(.443)</td>
<td></td>
</tr>
<tr>
<td>Weight (in kilogram)</td>
<td>64.390</td>
<td>63.890</td>
<td>66.980</td>
<td>66.006</td>
<td>Home Interview Wave I</td>
</tr>
<tr>
<td></td>
<td>(15.811)</td>
<td>(15.681)</td>
<td>(16.229)</td>
<td>(15.204)</td>
<td></td>
</tr>
<tr>
<td>Height (in meters)</td>
<td>1.692</td>
<td>1.692</td>
<td>1.692</td>
<td>1.707</td>
<td>Home Interview Wave I</td>
</tr>
<tr>
<td></td>
<td>(.106)</td>
<td>(.106)</td>
<td>(.108)</td>
<td>(.096)</td>
<td></td>
</tr>
<tr>
<td>BMI</td>
<td>22.328</td>
<td>22.136</td>
<td>23.239</td>
<td>22.601</td>
<td>(constructed)</td>
</tr>
<tr>
<td></td>
<td>(4.356)</td>
<td>(4.244)</td>
<td>(4.744)</td>
<td>(4.618)</td>
<td></td>
</tr>
<tr>
<td>Attractiveness</td>
<td>.000</td>
<td>.007</td>
<td>-.041</td>
<td>.414</td>
<td>Home Interview Wave I</td>
</tr>
<tr>
<td></td>
<td>(1.000)</td>
<td>(.998)</td>
<td>(1.008)</td>
<td>(.904)</td>
<td></td>
</tr>
<tr>
<td>Achievement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No Learning Disability</td>
<td>.861</td>
<td>.863</td>
<td>.850</td>
<td>.961</td>
<td>Parent Questionnaire</td>
</tr>
<tr>
<td></td>
<td>(.346)</td>
<td>(.344)</td>
<td>(.357)</td>
<td>(.197)</td>
<td></td>
</tr>
<tr>
<td>AHPVT</td>
<td>.000</td>
<td>.164</td>
<td>-.756</td>
<td>-.161</td>
<td>Home Interview Wave I</td>
</tr>
<tr>
<td></td>
<td>(.913)</td>
<td>(1.037)</td>
<td>(1.011)</td>
<td>(1.011)</td>
<td></td>
</tr>
<tr>
<td>GPA</td>
<td>2.865</td>
<td>2.934</td>
<td>2.592</td>
<td>2.720</td>
<td>School Survey</td>
</tr>
<tr>
<td></td>
<td>(.793)</td>
<td>(.788)</td>
<td>(.755)</td>
<td>(.783)</td>
<td></td>
</tr>
<tr>
<td>Never Repeated Grade</td>
<td>.797</td>
<td>.824</td>
<td>.673</td>
<td>.653</td>
<td>Home Interview Wave I</td>
</tr>
<tr>
<td></td>
<td>(.403)</td>
<td>(.381)</td>
<td>(.469)</td>
<td>(.480)</td>
<td></td>
</tr>
<tr>
<td>Behavior in School</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trouble with Teacher</td>
<td>.000</td>
<td>.051</td>
<td>-.198</td>
<td>.003</td>
<td>School Survey</td>
</tr>
<tr>
<td></td>
<td>(1.000)</td>
<td>(.967)</td>
<td>(1.098)</td>
<td>(.980)</td>
<td></td>
</tr>
<tr>
<td>Trouble Paying Attention</td>
<td>.000</td>
<td>.031</td>
<td>-.119</td>
<td>-.127</td>
<td>School Survey</td>
</tr>
<tr>
<td></td>
<td>(1.000)</td>
<td>(.972)</td>
<td>(1.095)</td>
<td>(.990)</td>
<td></td>
</tr>
<tr>
<td>Trouble with Homework</td>
<td>.000</td>
<td>.044</td>
<td>-.170</td>
<td>-.132</td>
<td>School Survey</td>
</tr>
<tr>
<td></td>
<td>(1.000)</td>
<td>(.976)</td>
<td>(1.071)</td>
<td>(.955)</td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Summary Statistics by Race

Strict Self-Declared Definition
<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Survey</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trouble with Students</td>
<td>0.00</td>
<td>0.065</td>
<td>-0.247</td>
<td>-0.214</td>
<td>School Survey</td>
</tr>
<tr>
<td></td>
<td>(1.000)</td>
<td>(.966)</td>
<td>(1.086)</td>
<td>(1.053)</td>
<td></td>
</tr>
<tr>
<td>Effort Schoolwork</td>
<td>0.00</td>
<td>-0.058</td>
<td>0.224</td>
<td>-0.095</td>
<td>School Survey</td>
</tr>
<tr>
<td></td>
<td>(1.000)</td>
<td>(1.002)</td>
<td>(.956)</td>
<td>(1.089)</td>
<td></td>
</tr>
<tr>
<td>Skipping School</td>
<td>0.00</td>
<td>-0.013</td>
<td>0.054</td>
<td>-0.207</td>
<td>School Survey</td>
</tr>
<tr>
<td></td>
<td>(1.000)</td>
<td>(1.002)</td>
<td>(.982)</td>
<td>(1.319)</td>
<td></td>
</tr>
<tr>
<td>Never Suspended or Expelled</td>
<td>0.00</td>
<td>0.104</td>
<td>-0.470</td>
<td>-0.547</td>
<td>Home Interview Wave I</td>
</tr>
<tr>
<td></td>
<td>(1.000)</td>
<td>(.939)</td>
<td>(1.123)</td>
<td>(1.132)</td>
<td></td>
</tr>
<tr>
<td>Behavior Outside School</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Watch TV</td>
<td>0.00</td>
<td>0.140</td>
<td>-0.527</td>
<td>-0.255</td>
<td>School Survey</td>
</tr>
<tr>
<td></td>
<td>(1.000)</td>
<td>(.947)</td>
<td>(1.021)</td>
<td>(0.987)</td>
<td></td>
</tr>
<tr>
<td>Drinking</td>
<td>0.00</td>
<td>-0.026</td>
<td>0.103</td>
<td>-0.145</td>
<td>School Survey</td>
</tr>
<tr>
<td></td>
<td>(1.000)</td>
<td>(1.004)</td>
<td>(.975)</td>
<td>(1.055)</td>
<td></td>
</tr>
<tr>
<td>Smoking</td>
<td>0.00</td>
<td>-0.078</td>
<td>0.303</td>
<td>-0.060</td>
<td>School Survey</td>
</tr>
<tr>
<td></td>
<td>(1.000)</td>
<td>(1.055)</td>
<td>(.667)</td>
<td>(1.073)</td>
<td></td>
</tr>
<tr>
<td>Daring</td>
<td>0.00</td>
<td>-0.057</td>
<td>0.227</td>
<td>-0.191</td>
<td>School Survey</td>
</tr>
<tr>
<td></td>
<td>(1.000)</td>
<td>(1.025)</td>
<td>(.856)</td>
<td>(1.154)</td>
<td></td>
</tr>
<tr>
<td>Lie to Parents</td>
<td>0.00</td>
<td>0.005</td>
<td>-0.015</td>
<td>-0.203</td>
<td>School Survey</td>
</tr>
<tr>
<td></td>
<td>(1.000)</td>
<td>(1.985)</td>
<td>(1.054)</td>
<td>(1.125)</td>
<td></td>
</tr>
<tr>
<td>Fight</td>
<td>0.00</td>
<td>0.019</td>
<td>-0.071</td>
<td>-0.241</td>
<td>School Survey</td>
</tr>
<tr>
<td></td>
<td>(1.000)</td>
<td>(.997)</td>
<td>(1.007)</td>
<td>(1.130)</td>
<td></td>
</tr>
<tr>
<td>Property Damage</td>
<td>0.00</td>
<td>-0.032</td>
<td>0.153</td>
<td>-0.196</td>
<td>Home Interview Wave I</td>
</tr>
<tr>
<td></td>
<td>(1.000)</td>
<td>(1.033)</td>
<td>(.816)</td>
<td>(1.020)</td>
<td></td>
</tr>
<tr>
<td>Steal</td>
<td>0.00</td>
<td>-0.011</td>
<td>0.063</td>
<td>-0.572</td>
<td>Home Interview Wave I</td>
</tr>
<tr>
<td></td>
<td>(1.000)</td>
<td>(1.013)</td>
<td>(.920)</td>
<td>(1.398)</td>
<td></td>
</tr>
<tr>
<td>Violent Acts</td>
<td>0.00</td>
<td>0.058</td>
<td>-0.252</td>
<td>-0.629</td>
<td>Home Interview Wave I</td>
</tr>
<tr>
<td></td>
<td>(1.000)</td>
<td>(.947)</td>
<td>(1.166)</td>
<td>(1.601)</td>
<td></td>
</tr>
<tr>
<td>Sell Drugs</td>
<td>0.00</td>
<td>0.004</td>
<td>-0.020</td>
<td>-0.019</td>
<td>Home Interview Wave I</td>
</tr>
<tr>
<td></td>
<td>(1.000)</td>
<td>(.989)</td>
<td>(1.048)</td>
<td>(1.041)</td>
<td></td>
</tr>
<tr>
<td>See Violence</td>
<td>0.00</td>
<td>0.065</td>
<td>-0.295</td>
<td>-0.303</td>
<td>Home Interview Wave I</td>
</tr>
<tr>
<td></td>
<td>(1.000)</td>
<td>(.933)</td>
<td>(1.213)</td>
<td>(1.290)</td>
<td></td>
</tr>
<tr>
<td>Ever Sex</td>
<td>0.00</td>
<td>-0.080</td>
<td>-0.364</td>
<td>-0.381</td>
<td>Home Interview Wave I</td>
</tr>
<tr>
<td></td>
<td>(1.000)</td>
<td>(.980)</td>
<td>(1.010)</td>
<td>(1.016)</td>
<td></td>
</tr>
<tr>
<td>Ever STD</td>
<td>0.00</td>
<td>-0.065</td>
<td>-0.295</td>
<td>-0.313</td>
<td>Home Interview Wave I</td>
</tr>
<tr>
<td></td>
<td>(1.000)</td>
<td>(.795)</td>
<td>(1.599)</td>
<td>(1.642)</td>
<td></td>
</tr>
<tr>
<td>Ever Illegal Drugs</td>
<td>0.00</td>
<td>-0.018</td>
<td>0.099</td>
<td>-0.526</td>
<td>Home Interview Wave I</td>
</tr>
<tr>
<td></td>
<td>(1.000)</td>
<td>(1.007)</td>
<td>(.954)</td>
<td>(1.081)</td>
<td></td>
</tr>
<tr>
<td>Psychological Variables</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mother Cares</td>
<td>0.00</td>
<td>-0.016</td>
<td>0.062</td>
<td>-0.017</td>
<td>School Survey</td>
</tr>
<tr>
<td></td>
<td>(1.000)</td>
<td>(1.012)</td>
<td>(.947)</td>
<td>(1.069)</td>
<td></td>
</tr>
<tr>
<td>Father Cares</td>
<td>0.00</td>
<td>0.009</td>
<td>-0.044</td>
<td>-0.298</td>
<td>School Survey</td>
</tr>
<tr>
<td></td>
<td>(1.000)</td>
<td>(.978)</td>
<td>(1.111)</td>
<td>(1.352)</td>
<td></td>
</tr>
<tr>
<td>Close to People</td>
<td>0.00</td>
<td>0.042</td>
<td>-0.170</td>
<td>-0.095</td>
<td>School Survey</td>
</tr>
<tr>
<td></td>
<td>(1.000)</td>
<td>(.985)</td>
<td>(1.042)</td>
<td>(0.961)</td>
<td></td>
</tr>
<tr>
<td>Feel Accepted</td>
<td>0.00</td>
<td>-0.024</td>
<td>0.098</td>
<td>-0.002</td>
<td>School Survey</td>
</tr>
<tr>
<td></td>
<td>(1.000)</td>
<td>(.997)</td>
<td>(1.007)</td>
<td>(0.926)</td>
<td></td>
</tr>
<tr>
<td>Feel Loved</td>
<td>0.00</td>
<td>-0.030</td>
<td>0.124</td>
<td>-0.068</td>
<td>School Survey</td>
</tr>
<tr>
<td></td>
<td>(1.000)</td>
<td>(.992)</td>
<td>(1.022)</td>
<td>(1.023)</td>
<td></td>
</tr>
<tr>
<td>Not Depressed</td>
<td>0.00</td>
<td>-0.035</td>
<td>0.145</td>
<td>-0.185</td>
<td>School Survey</td>
</tr>
<tr>
<td></td>
<td>(1.000)</td>
<td>(.995)</td>
<td>(1.006)</td>
<td>(1.103)</td>
<td></td>
</tr>
<tr>
<td>Like Oneself</td>
<td>0.00</td>
<td>-0.071</td>
<td>0.292</td>
<td>0.021</td>
<td>School Survey</td>
</tr>
<tr>
<td></td>
<td>(1.000)</td>
<td>(.999)</td>
<td>(.950)</td>
<td>(1.061)</td>
<td></td>
</tr>
<tr>
<td>Chances Live to 35</td>
<td>0.00</td>
<td>0.054</td>
<td>-0.216</td>
<td>-0.013</td>
<td>School Survey</td>
</tr>
<tr>
<td></td>
<td>(1.000)</td>
<td>(.947)</td>
<td>(1.163)</td>
<td>(1.065)</td>
<td></td>
</tr>
</tbody>
</table>

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)

Note: Both figures are from Fryer (2007a).
Figure 2: Black-White Births as Percentage of all Black and White Births

Source: Data are from 2000 Census 5% Public Use Micro Sample. See data appendix for details.
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.
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.
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.
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.
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.
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.
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.