NBER WORKING PAPER SERIES

THE DYNAMICS OF EDUCATIONAL ATTAINMENT FOR BLACKS, HISPANICS, AND WHITES

Stephen V. Cameron
James J. Heckman

Working Paper 7249
http://www.nber.org/papers/w7249

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 1999

The Dynamics of Educational Attainment for Blacks, Hispanics, and Whites

ABSTRACT

This paper estimates a dynamic model of schooling attainment to investigate the sources of discrepancy by race and ethnicity in college attendance. When the returns to college education rose, college enrollment of whites responded much more quickly than that of minorities. Parental income is a strong predictor of this response. However, using NLSY data, we find that it is the long-run factors associated with parental background and income and not short-term credit constraints facing college students that account for the differential response by race and ethnicity to the new labor market for skilled labor. Policies aimed at improving these long-term factors are far more likely to be successful in eliminating college attendance differentials than are short-term tuition reduction policies.

Stephen V. Cameron   James J. Heckman
Department of Economics   Department of Economics
Columbia University   University of Chicago
420 West 118th Street   1126 East 59th Street
New York, NY 10027   Chicago, IL 60637
                                       and NBER
                                       jimh@cicero.spc.uchicago.edu

The recent rise in the economic return to higher education demonstrates the importance of college education as a major determinant of economic success. Disparity in levels of educational attainment between majority and minority groups has translated into growing disparity in earnings between the two groups. This disparity has been accentuated by differentials in the growth of college attendance of majority and minority youth in response to the rise in the economic return to college education.

The time series of the educational attainment of males classified by racial and ethnic status are presented in Figures 1A and 1B. For both high school completion, inclusive of GED certification, and college attendance, there are clear gaps. The Hispanic - White differential in high school completion has been stable over time. The Black-White differential has narrowed but most of this is due to the growing use of GED certification by Blacks.[1] (See Cameron and Heckman, 1993; Cavallo et.al, 1998). More disturbing is the disparity in college entry. Whites responded quickly to the rise in the return to college education that began in the early 80^s. However, Blacks and Hispanics lagged in their response. It is only in the early 90^s that minority college attendance begins to respond to the increased returns to college education. The conventional explanation for the level of disparity in educational attainment and the differential response to the increased return to college education is disparity in the family resources required to finance a college education. (See, e.g. Kane, 1994). Figure 2 apparently supports this claim. The educational response to the rising return to education was the most rapid and the largest for young adults whose families are in the top half of the family income distribution. The response of children from families at the bottom of the family income distribution was substantially delayed. Minority children are concentrated in families near the bottom of the overall family income distribution where real earnings have declined in the past 15 years. The lower real earnings of minority parents coupled with the rise in real tuition costs would seem to suggest that short-term liquidity constraints on educational finance explain much of what is happening in Figure

---

[1] Cameron and Heckman (1993) and Cavallo et.al (1998) document that this disparity contributes to about 10% of the growth in measured black-white inequality within the complete high school category in the 80^s since GED recipients earn less than High School graduates.

1

1B. Recent educational policy is founded on this interpretation of the evidence. Policies based on this premise advocate tuition offsets and income supplements to stimulate the college attendance of students from low income families. (See the essays in Kosters, 1998).

This paper uses better data and models than have been used by previous analysts to analyze the determinants of inequality in educational attainment among majority and minority youth. Using longitudinal data from the National Longitudinal Survey of Youth (NLSY), we estimate educational choices as sequential decisions made at each age and each grade from feasible person-specific choice sets. Previous cross-sectional studies have focused only on explaining years of schooling completed or have analyzed the transition from high school to college without considering the determinants of who graduates from high school.[2] Conditioning on a choice variable such as completion of high school masks the total effect of socioeconomic variables on college entry decisions. The decision to go to college is near the end of a long chain of educational decisions made earlier in the life cycle. To understand who goes to college requires that we also understand who is eligible to do so. In this paper, we build on our previous work (Cameron and Heckman, 1998) and · develop a sequential model of schooling attainment that accounts for the selective nature of high school graduates. More able people progress to higher grades and it is necessary to account for this selection in order to estimate the *ceteris paribus* "causal effect" of the socioeconomic variables we study on educational attainment.

We estimate how family background, family income, college tuition costs, labor market opportunities and cognitive ability affect age - and grade-specific schooling transitions. We reach the following major conclusions about the sources of disparity in college attendance between majority and minority group members.

(1) In agreement with the conclusions reported in a substantial literature in economics and sociology, we find that when we do not control for family background, family income

---

[2] Kane (1994) is an exception to this rule. He jointly analyzes the determinants of high school graduation and college attendance. We consider a more general model in which age - and grade-specific schooling choices are analyzed, and schooling transitions prior to high school graduation are analyzed. Many Blacks and Hispanics never get to 12th grade, much less graduate.

is a major determinant of college attendance. Given minority deficits in family resources, it is not surprising that minority college enrollment is lower than that of majority White children.

(2) When we control for the long-run effect of family and the environment on college entry as measured by a widely- used measure of cognitive ability, or by controlling for family background (parental education), the estimated effect of family income on college attendance is greatly weakened. It is the long-term influence of families and not short-term "liquidity constraints" that accounts for much of the ethnic and racial disparity in college entry. Equalizing these long-term factors, minority youth are *more likely* to complete high school and enter college than are majority youth, even after controlling for selectivity in educational attainment. Conditioning on background variables, family income has a stronger effect on who stays in high school and who graduates than it has on who attends college.

(3) College tuition costs do not explain the disparity between majority and minority schooling attainment.

Taken together, these conclusions suggest that long-term factors like family environments matter greatly in fostering the abilities and attitudes required for entering college. Short-term tuition and income supplements are unlikely to have much effect in alleviating the gap in college entry between minority and majority group members. The fact that take-up rates are so low for Pell Grant and other programs targeted toward students from low-income families reinforces this point. (Orfield, 1992).

The plan of this paper is as follows. (1) In part 1, we present some basic facts about educational attainment using the NLSY data and briefly summarize the conventional interpretation of the evidence. (2) In part 2, we present the econometric model used to make a more refined interpretation of the basic facts of educational attainment presented in Part 1. Formal econometric issues are discussed in Appendix B. (3) In part 3, we discuss the estimates of the model. A concluding section summarizes the paper.

3

1. Our Data and Some Facts About Schooling Attendance

This paper analyzes the dynamics of schooling attainment from early high school through college entry for Black, Hispanic, and White males. We analyze males because their schooling decisions are less complicated by fertility considerations. Our data are drawn from the 1979-1989 waves of the National Longitudinal Survey of Youth (the military subsample and the non-Black non-Hispanic disadvantaged samples are excluded). Because the NLSY collected detailed information about school attendance and completion starting from January, 1978, it is a valuable data set for studying schooling dynamics. Information on schooling is available starting at age 15 and extends through age 24 for all individuals included in our sample. Because information before 1978 is retrospective and limited, we confine our analysis to males between the ages of 13 and 16 in January, 1978 in order to have reliable schooling information on individuals beginning at age 15. In fact, few males quit school before age 16. An extensive discussion of the NLSY survey, and the construction of individual schooling histories, is presented in Appendix A.

Previous influential work by Hauser (1991), Kane (1994) and others is based on Current Population Survey (CPS) Data. These data suffer from major limitations of special importance to analyses of the role of family background and family income on college enrollment decisions. The CPS data report parental characteristics only for persons who are living in the parental home or for those attending college who live in a dormitory or other group quarters. Data on parental characteristics are not available for youth who live on their own, outside of group quarters.

Virtually all of the existing evidence on the importance of family background and family income on schooling choices is derived from samples of "dependents," *i.e.* persons living in the parental home or students living in group quarters. Implicitly, these studies analyze the effect of parental variables on college entry conditional on the choice of a living arrangement. These studies run the risk of obscuring the effects of parental variables on educational attainment because they condition on another choice variable (residential living decisions). In our previous work we demonstrate that as consequence of this conditioning, CPS-based

4

studies tend to underestimate the contribution of family variables to college attendance decisions. (Cameron and Heckman, 1992). The last section of Appendix A presents a more complete discussion of the problems in using CPS data to analyze college enrollment decisions.

The NLSY contains much more information relevant to schooling enrollment decisions than does the CPS. We demonstrate the value of this information in understanding schooling decisions. In addition, we marshal several auxiliary data sources to augment our analysis. To measure college tuition costs, we combine the NLSY with data on two- and four-year college tuition measured at both the state and the county levels. We also construct measures of proximity to both two- and four-year colleges. College proximity is an important factor in college entry decisions and has been used to instrument endogenous schooling attainment in several recent studies of the returns to schooling (see Kane and Rouse, 1995, and Card, 1996). In addition, we draw on a number of measures of state and county labor market conditions better gauge how labor market opportunities in low skill markets affect college enrollment decisions. We now present the main contours of the schooling dynamics of Black, White and Hispanic males.

Basic Facts about Schooling Attainment

We analyze educational attainment decisions from age 15 through age 24. Throughout this paper we measure age as of October. We adopt this convention for two reasons. First, October approximately corresponds to the age-cutoff rules for first grade entry used by many school districts in the United States. Second, the CPS surveys on schooling attainment that are widely used to study schooling are conducted in October. The October CPS data cannot be used to study individual schooling transitions, but they do provide a cross-check on basic schooling attainment facts from the NLSY. Since most of the literature is based on **this data**, we follow this convention to make our analyses comparable to those of previous **studies**.

Figure 3A shows the highest grade completed at age 15 for Blacks, Hispanics, and Whites. At age 15, grade nine is the modal highest grade completed, and the grade a

5

student would have completed had he entered the first grade at age 6 and attended school continuously through age 15. By age 15, large differences in schooling attainment already exist across racial-ethnic groups. While all racial-ethnic groups are equally likely to be above the modal grade, minorities are 16 to 21 percent more likely to be below it. These White-minority differences are apparently due to uneven chances of grade advancement during elementary school as well as to postponed entry into first grade or episodes of dropping out of school and returning before age 15. For the cohort of youth we study, Blacks start behind and stay behind. Disparity in the age of first-grade entry among racial-ethnic groups is confirmed by tabulations of age and grade from October CPS data. For Hispanics and Blacks of this cohort, twice as many are start first grade at age 7 or later than Whites. (U.S. Department of Education, 1996).

Differences in final schooling attainment emerge after age 16.[3] Panel B of the figure shows the large, well-documented disparities among racial-ethnic groups in final attainment. The two most important features of panel B are the large differences in college attendance (category ">12" on the figure) and high school dropping out ("<12").[4] Whites are about 12 to 14 percent more likely to enroll in college by age 24, and Hispanics are about 15 percent more likely than Whites to have not finished high school. Analyses for later ages confirm that there is very little college entry or high school completion after age 24 so the story told by Figure 3B applies as a characterization of life cycle schooling.

Additional facts about schooling achievement at age 24 are highlighted in Tables 1 and 2. Table 1 reveals that high school completion rates are higher for Whites than minorities but the opposite is true for GED-certification. Eleven percent of Whites earn high school credentials via the GED compared to 17% of Blacks and 23% of Hispanics.

Table 2 presents differentials in college entry rates for two - and four-year college entry by age 24. Panel A summarizes the proportions for high school graduates (excluding GED

---

[3] This finding is due in part to compulsory schooling attendance laws and to the lack of labor market opportunities for people younger than 16.

[4] To accord with Census and CPS convention, GED attainment and traditional high school graduation have been combined into the category "12". We break this out in Table 1.

6

recipients). The third column for each group displays the fraction who attend either type of college. White graduates are more likely to enter college than are Blacks (by seven percentage points) or Hispanics (by three percentage points). Whites have the highest four-year college entry rate and Hispanics the lowest. Hispanics, however, show the highest two-year entry rate, which is partly attributable to the regional concentration of Hispanics in states such as California and Texas with extensive community college networks and low-tuition costs during the period of our study. Panel B of the table shows that GED recipients are about half as likely as traditional high school graduates to enter college and are more likely to attend two-year than four-year college. GED recipients represent only six or seven percent of all college entrants.

The link between early schooling attainment and later educational outcomes is illustrated in Table 3, which shows the distribution of schooling attainment at age 24 by highest grade attained at age 15. The grade distribution at age 15 is given in percentages in column 5. The table shows that individuals at grade levels below the modal grade are much more likely to drop out of high school and are more likely not to enter college if they complete high school. Subsequent high school graduation and college entry chances are highly correlated with early schooling attainment.

Dynamics of Schooling Transitions

Most educational attainment histories in secondary school follow the standard pattern of no interruption or delay. At age 15, less than 1% of Blacks and Whites and about 2% of Hispanics do not attend school. In fact, most dropping out occurs at age 17 or 18, and a significant number of dropouts attend 12 years of secondary school. It is not a surprise, then, that eleven years is the modal highest grade completed among dropouts for all racial-ethnic groups. Table 4 shows the fraction of dropouts, high school graduates, and GED recipients attending school full-time at age 15 who leave and return to secondary or elementary school anytime between ages 15 and 24. We classify dropouts liberally as students who leave school for at least eight consecutive weeks during the regular school year. For both high school graduates and those who never complete high school, returning to school is a rare event: only 2 to 6 percent of high school graduates and 6 to 12 percent

7

of eventual dropouts report at least one episode of leaving and subsequent return to school (panels A and C). For GED recipients the corresponding numbers are high and range between 36 and 42 percent.[5]

Panel A of Figure 4 shows the cumulative probability of traditional high school graduation between age 15 and 24. High school graduation occurs almost exclusively at age 18 or 19. A small percent graduate before age 18 but almost none after. GED acquisition is the main vehicle of high school completion after age 19 (panel B), though at least half of all GED completion also occurs by age 19.

Panel C of the figure illustrates the age distribution of college entry. This figure reveals that most college entry happens at age 19 or 20, immediately after high school completion. Table 5 presents the delay between high school completion and initial college entry. Panel A of the table shows that, among high school graduates, 82% percent of Whites and Hispanics and 73% of Blacks who ever enter college do so within a year of high school graduation. Between 5 and 8 percent wait more than three years. Panel B shows that about half of GED recipients who ever enter college enter it within a year of attaining their GED.

The Family Income-Schooling Connection

A central question addressed in this paper is how family income influences schooling attendance. Table 6 addresses this question in a simple way. Panel A of the table displays two- and four- year college entry proportions by family income levels for a combined group of GED recipients and high school graduates as a summary of Table 2. Panels B through E show the same proportions broken down by the position of family income in terms of quartiles of the White family income distribution. For instance, Panel B shows college attendance rates for the children of families in the top quartile of the White family income distribution. The same income cutoffs in the White distribution are used to categorize the family incomes of Blacks and Hispanics. Column (4) shows that Blacks and Hispanics

---

[5] A more thorough description of the schooling transition data can be found in an earlier version of this paper (Cameron and Heckman, 1996, available upon request.) That paper examines age-and grade-specific transitions into and out-of-school as well as the numerical importance of students who attend only part of the school year. Most dropping out happens during the summer as students fail to reenter school in the Fall.

are underrepresented in the top category because of their overall lower position in the overall family income distribution.[6] Looking at column (3) from top to bottom shows that as family income increases, youth of all racial-ethnic backgrounds are more likely to enter college.[7] For all race groups, the chance of attending a four-year college rises with family income. Nevertheless, there are substantial differences in sensitivity to family income among the three groups. Hispanics are the least sensitive to family income differences and exhibit the smallest difference in the chances of college attendance between the top and bottom quartiles. Blacks are the most sensitive and have the highest rate of four year college attendance overall and the lowest rate in both the top, third, and bottom quartiles. Two-year attendance shows no definite pattern. Moving up the family income scale, more children enter college, but there appears to be substitution from two-year to four-year colleges as family income increases.

An interesting counterfactual question we address is "how would the White-minority gap in panel A differ if Blacks and Hispanics had the same family income distribution as Whites?" The answer to this question is given in panel F.[8] Column (3) shows the fraction of Blacks and Hispanics going to college when family income distributions are equated to the White distribution. Comparing panel F to the actual enrollment rates reported in panel A shows that, for Blacks, nine percentage points of the eleven percentage point gap in overall college entry is eliminated by equalizing family income. Hispanics recover five percentage points of an eight percentage point gap. The remaining adjusted gaps for both Blacks and Hispanics are statistically insignificant. In addition, looking at column (1), Blacks are slightly more likely than Whites, by one percentage point, to attend a four-year

---

[6]Parental family income is measured between age 13 and 17. It should be noted that the income distribution used to compute the categories in Table 6 does not represent the population distribution of White family income in the U.S. but the population distribution for White families with children who are recent high school completers.

[7]The only exception is a curious drop in college attendance for Hispanics as we go from the second to top quartile (panels B and C). This fact seems to be due to the large rise in the number of Hispanic youth attending two-year college between the top and second quartiles.

[8]This exercise amounts to a averaging of the numbers in columns (1)-(3) of panels B through E using the appropriate weights given in column (4).

school once family income is equalized.

Because the proportions of Blacks and Hispanics in each of the income cells shown in panels B though E are approximately equal (that is, their family income distributions are about the same), these numbers reveal that family income is an important factor for Hispanics and Blacks in determining both college entry and the type of college attended.

Explanatory Variables

Table 7 defines the basic variables used below in our multivariate analysis of schooling attainment. In addition to these variables we also include time dummies. Table 8 summarizes mean differences in family background, scholastic ability, labor market characteristics, location and tuition costs among Blacks, Whites and Hispanics for the variables defined in Table 7. Some of these differences arise from the different geographical distributions of these groups. There are two notable omissions from the list. The first is expected future returns to and costs of a college education. In general, these variables are difficult to estimate (for us or for the people we study) and in this paper, we use time dummy variables to proxy the changing structure of returns to education. Second, standard school quality measures were also investigated, but their contribution was found to be negligible when family background measures are included in the list of explanatory variables in the analysis presented in Section 3.

Family background differences favor Whites. Minorities live in geographical areas with lower tuition costs and with lower commuting costs. Local labor market effects on the opportunity costs of schooling tend to be neutral across race groups. Measures of scholastic ability are largest for whites, intermediate for Hispanics and lowest for Blacks. These univariate relationships indicate that tuition costs play a small role in accounting for racial and ethnic differences and that scholastic ability plays a much more substantial role. These findings are sustained in the more refined econometric analysis presented in Section 3.

Sizeable differences are apparent in family background characteristics among Blacks, Hispanics, and Whites. Family size is much larger for Blacks and Hispanics than for

Whites.[9] Parental education is lowest for Hispanics and highest for Whites. Broken homes are more prevalent among Blacks and least common among Whites. Family incomes are most favorable for Whites and least favorable for Blacks.

Tuition rates at both two-year and four-year public colleges are highest for Whites and lowest for Hispanics (tuition is measured at both the state and the county level, when available). The imputed size of Pell grant awards for college tuition and expenses are about $1200 for Blacks and Hispanics and about $500 for Whites. Hence, net tuition at four-year college ranges from $180 for Hispanics to $1200 for Whites on average. The last row of the table shows that college access is highest for Hispanics. Ninety-two percent of Hispanics live in a county in which either a two- or four-year college is located compared to only 82% of Whites. Local labor market conditions also vary across racial-ethnic lines. Though the county unemployment rate shows little average variation across groups, average wages for unskilled workers are about 10% higher in counties where Hispanics reside than they are for Whites. Black opportunity costs are in between.

Finally, large differences between White and minority scholastic ability as measured by the Armed Forces Qualification Test (AFQT) are evident. AFQT figures prominently in our analysis below. Our measure of the AFQT is age-adjusted. Moreover, because our sample is 13-16 years of age when the AFQT tests are taken, there is no effect of high school graduation or of college attendance on the test score so that the test score is relatively free of endogeneity from schooling.[10] Further information about the AFQT is given in Appendix B.

The major findings of this section are:

(1) By age 15 before any significant dropping out occurs, there are large differences in racial-ethnic schooling attainment. Whites start ahead and stay ahead.

(2) Early schooling attainment is highly correlated with later schooling attainment.

---

[9]These reported family sizes are inflated due to the family-size biased nature of the NLSY sampling frame.

[10]Throughout this paper we use an age-adjusted measure of AFQT. Experimentation with education-adjusted AFQT measures produced essentially the same inference in the analysis we report below.

(3) Most high school completion occurs at age 18 or 19 after 12 years of continuous schooling. Interruption of secondary schooling is an uncommon event. High school completion after age 19 is almost exclusively through GED attainment. College entry, if it happens at all, follows high school completion with little or no delay. About 90% of all college entry occurs within two years of high school completion.

(4) Family income is a powerful correlate of college entry. Indeed, differences in family income alone can explain most of the White-minority gaps in college entry of high school completers.

(5) Whites and minorities are significantly different in terms of family background, the actual costs of college they face, and a number of other important variables.

We now turn to a multivariate analysis of schooling attainment using a dynamic discrete choice model that extends previous grade transition models estimated in the literature on schooling attainment.

2. An Econometric Model of Schooling Attainment

We extend the econometric models currently used in the literature on the economics of schooling attainment by formulating and estimating a dynamic discrete choice model over ages 15-24 for the people we study. While previous analysts have mostly concentrated on the determinants of "highest grade completed" or whether or not a person attends college for a sample of people who have already completed high school, the point of departure for our research is the recognition that schooling attainment at any age is the outcome of previous schooling choices. The probability that a person enters college depends on high school graduation, which in turn depends on finishing grade 11 and so forth back to the earliest schooling decisions. For minority groups and low-income Whites, high school graduates are select members of the source population and it is important to control for the effects of educational selectivity to isolate *ceteris paribus* effects of tuition and family background on college attendance. Cameron and Heckman (1998) document the empirical importance of controlling for educational selectivity in isolating *ceteris paribus* effects of family background on schooling decisions.

12

Researchers who have studied how family factors affect the highest grade of schooling completed cannot distinguish the effect of family income on the high school attendance decisions from its effect on college entry. Family factors and other influences may affect schooling decisions differentially by age and grade level.

To sort out the influence of family income on college entry for high school completers from its accumulated long-term influence in making people eligible to attend college, conventional methods in the educational attainment literature used by Hauser and Kane are problematic. Our methodology enables us to separate out age-by-age influences in a general way. By analyzing the entire set of age-specific schooling decisions from age 15 through age 24, we are able to parcel out by age the influences of family income and other variables. Using our estimated econometric model, we can then evaluate the consequences of policies that seek to promote college attendance through raising high school graduation.[11]

Let age be denoted by a ($a \in \{\underline{a}, ..., \bar{a}\}$, where $\underline{a}$ is the initial age, and $\bar{a}$ is the highest age). Schooling attainment at age a is $j_a \in J$ (J is a set of possible attainment states over all ages). Agents with schooling status j_a make their choices about schooling at age $a + 1$ from the feasible choice sets C_{a,j_a}. Let $D_{a,j_a,c}$ be 1 if option $c \in C_{a,j_a}$ is chosen by a person of age a with schooling status j_a. Assuming that some choice is made,

$$\sum_{c \in C_{a,j_a}} D_{a,j_a,c} = 1.$$

The model is fundamentally recursive; the choice made at a affects the choice set at age $a + 1$.

While this notation may appear to be cumbersome, it allows us to consider schooling attainment processes that are more general than straight grade progression models

---

[11]Two other problems that past researchers have generally ignored are also handled in our framework. First is the issue of time-varying explanatory variables, such as indicators of the state of the labor market. Previous frameworks are fundamentally atemporal, and accommodate variables that change over time only in arbitrary ways. Second, we control for unobserved characteristics. Our framework builds on the work of Cameron and Heckman (1998), who show that serious biases in the estimated effects of family income on schooling arise when unobserved variables that are persistent over grade transitions are not accounted for.

presented in Bartholomew (1973), Mare (1980) or Cameron and Heckman (1998). For example, schooling attainment may consist of 11 years of formal school and a GED; grade transition probabilities may be age (or time) specific. This model allows us to generalize the notion of a grade progression that dominates the literature on educational attainment and to recognize that a variety of different schooling trajectories are possible.

Assume that agents choose optimally at each age and schooling status j, inclusive of the options for further schooling opened up by attaining status j (Comay, Melnik and Pollatshek, 1973). Then the optimal choice at age a, denoted by "$\hat{\ }$" is

$$\hat{c}_{a,j_a} = \underset{c \in C_{a,j_a}}{\arg\max} \{V_{a,j_a,c}\}$$

where $V_{a,j_a,c}$ is the value of option c at age a for a person with j_a years of schooling. Then $D_{a,j_a,c} = 1$ for $c = \hat{c}_{a,j_a}$ and $D_{a,j_a,c} = 0$ otherwise. The model is fundamentally sequential; the choice set C_{a,j_a} confronting the agent at age a is a consequence of choices made last period. Observe that $j_a = \hat{c}_{a-1,j_{a-1}}$. To avoid notational clutter henceforth we drop the "a" subscript on "j" except where making it explicit clarifies the argument.

For computational simplicity we approximate $V_{a,j,c}$ using a linear in the parameters form:

(1) $V_{a,j,c} = Z'_{a,j}\beta_{a,j,c} + \varepsilon_{a,j,c}$

where $Z_{a,j}$ is a vector of observed (by the econometrician) constraint and expectation variables at age a for a person of schooling attainment j and $\varepsilon_{a,j,c}$ is an unobservable from the point of view of the economic analyst. Heckman (1981), Eckstein and Wolpin (1989) and others advocate this linear-in-the-parameters structure as a starting point for a more general analysis of discrete dynamic choices. The essential idea in this model, as in any sequential model of discrete choice, is that the choice sets $C_{a,j}$ as captured in (1) by the $Z_{a,j}$, $\beta_{a,j,c}$ and $\varepsilon_{a,j,c}$ are determined by previous choices. Econometrically, this creates the possibility at any point in the decision process that the $Z_{a,j}$ conditional on past choices are endogenous.

In this paper we follow the computational simplification for discrete choice processes proposed by Heckman (1981, Appendix) and assume that $\varepsilon_{a,j,c}$ is characterized by a factor

structure:

$$(2) \quad \varepsilon_{a,j,c} = \alpha_{a,j,c}\eta + \nu_{a,j,c}$$

where

(A-1) $\eta \perp\!\!\!\perp \nu_{a,j,c}$ ("$\perp\!\!\!\perp$" denotes independence)

for all a, j, c, and the η are independent across persons, and η is a mean zero, unit variance random variable. In addition, we assume that all of the random variables are independent across people.

We further assume that

(A-2) $\nu_{a,j,c}$ is an extreme value random variable, independent of all other $\nu_{a',j'',c'''}$ except for $a = a', j = j''$ and $c = c'''$.

The extreme value assumption produces a one-factor version of McFadden's (1974) multinomial logit model. Conditional on η, we obtain :

$$(3)\; \Pr(D_{a,j,c'} = 1 \mid Z_{a,j},\eta) = \Pr(\arg\max_{c} V_{a,j,c} = c' \mid Z_{a,j}, \eta) = \frac{\exp\{Z'_{a,j}\beta_{a,j,c'} + \alpha_{a,j,c'}\eta\}}{\sum_{c \in C_{a,j}} \exp\{Z'_{a,j}\beta_{a,j,c} + \alpha_{a,j,c}\eta\}}.$$

As a consequence of the one-factor assumption (A-1), any dependence between $D_{a,j,c}$ and $D_{a',j'',c'''}, a \neq a'$, for the same person conditional on $Z_{a,j}$ and $Z_{a',j''}$ arises from η, the omitted person-specific effect.

One way to eliminate this dependence is to estimate η for each person, or condition it out, using the methods surveyed in Arellano and Honoré (1998). In general, estimating η along with the other parameters of the model produces inconsistent estimates. We estimate the model by assuming that

(A-3) $Z_{a,j} \perp\!\!\!\perp \eta, \qquad \forall a, j \in C_{a,j}$ for all choice sets,

i.e. that the $Z_{a,j}$ are independent of η. This does not imply that the $Z_{a,j}(a > \underline{a})$ *conditional on past choices* are independent of η.[12] Because in general they are not, it is necessary to model the **history of the process** leading up to any transition being analyzed.

---

[12] Cameron and Heckman (1998) demonstrate how conditioning on the history of the life cycle process corrects for the induced dependence between η and $Z_{a,j}$, $a > \underline{a}$, given the history of previous choices.

With these assumptions in hand, we may write down the probability of any schooling history by building up the sequence of age-specific probabilities over the life-cycle. Let $D_{\underline{a},c}$ denote the initial schooling attainment state at age $\underline{a}$; $\sum_{c \in C_{\underline{a}}} D_{\underline{a},c} = 1$, where $C_{\underline{a}}$ is the set of possible initial states; $Z_{\underline{a}}$ is the data that summarizes the initial state. For later purposes we denote $d_{\underline{a},c}$ as the realization from $D_{\underline{a},c}$. For simplicity, we also parameterize the initial state probability using a multinomial logit model:

$$(4) \qquad \Pr(D_{\underline{a},c'} = 1 \mid Z_{\underline{a}}, \eta) = \frac{\exp\{Z_{\underline{a}}'\beta_{\underline{a},c'} + \alpha_{\underline{a},c'}\eta\}}{\sum_{c \in C_{\underline{a}}} \exp\{Z_{\underline{a}}'\beta_{\underline{a},c} + \alpha_{\underline{a},c}\eta\}}.$$

At age $\underline{a}+1$, the agent has choice set $C_{\underline{a}+1,c'}$, and we may write the probability that c is chosen given that c' was the initial state at $\underline{a}$ as $\Pr(D_{\underline{a}+1,c',c} = 1 \mid Z_{\underline{a}+1,c'}, \eta)$ where we make explicit the conditioning on the previous choice (which defines the feasible choice set at age $\underline{a}+1$). Note that $j_{\underline{a}+1} = c'$. The probability of any sequence of life cycle schooling histories $(D_{\underline{a},c'} = 1, D_{\underline{a}+1,c_{\underline{a}},c} = 1, D_{\underline{a}+2,c_{\underline{a}+1},c} = 1, ..., D_{\bar{a},c_{\bar{a}-1},c} = 1)$ given the relevant conditioning sets and η is

$$(5) \qquad \Pr(D_{\underline{a},c'} = 1 \mid Z_{\underline{a}}, \eta) \cdot \Pr(D_{\underline{a}+1,c',c} = 1 \mid Z_{\underline{a}+1,c'}, \eta) \cdot$$
$$\Pr(D_{\underline{a}+2,c_{\underline{a}+1},c} = 1 \mid Z_{\underline{a}+2,c_{\underline{a}+1}}, \eta) \cdots \Pr(D_{\bar{a},c_{\bar{a}-1},c} = 1 \mid Z_{\bar{a}}, c_{\bar{a}-1}, \eta).$$

While notationally somewhat formidable, this expression is just the probability that a person starts in initial state $c' \in J$, progresses to $c_{\underline{a}+1}$ at age $\underline{a}+1$ given the choice set $C_{\underline{a}+1,c_{\underline{a}}}$ produced by the choice at age $\underline{a}$, then progresses to state $c_{\underline{a}+2}$ at age $\underline{a}+1$, then progresses to state $c_{\bar{a}}$ at age $\bar{a}$ given the choice $c_{\bar{a}-1}$ made at age $\bar{a}$-1.

The model we estimate integrates out the η in expression (5) using the distribution $F(\eta)$. An analysis of identification of the factor structure for this model is given in Heckman and Taber (1994). The full likelihood and a discussion of the nonparametric likelihood estimator for $F(\eta)$ are presented in Appendix B. Standard asymptotic distribution theory results apply to this model following the analysis of Chen, Heckman and Vytlacil (1998). In brief, we normalize $\mathbf{Var}(\eta) = 1$, and make one normalization of one factor loading $\alpha_{\underline{a},c^*} = 1$ for one component $c^* \in C_{\underline{a}}$. As is common in discrete choice models, it is necessary to normalize one benchmark state to zero for each choice set ($\beta_{a,j,\bar{c}} = 0$ and $\alpha_{a,j,\bar{c}} = 0$ for

16

benchmark state $\tilde{c}$ for each a, j).

Our model nests a variety of widely-used models as special cases. The grade-progression model of Bartholomew (1973) and Mare (1980) specifies the choice sets facing individuals to be (a) independent of age:

(A-4a) $\qquad C_{a,j} = C_j, \qquad \forall a$

and (b) to possess two elements: either continue to the next grade or not. Letting j be current grade level,

(A-4b) $\qquad C_j = \{j, j+1\}.$

Those models ignore heterogeneity

(A-4c) $\qquad \alpha_{a,j,c} = \alpha_{j,c} = 0 \qquad \forall a, j$

and so do not account for educational selectivity. Cameron and Heckman (1998) maintain (A-4a) and (A-4b) but relax (A-4c) and build models that account for grade-specific effects of a common unobservable variable. Unlike Willis and Rosen (1979), we model schooling decisions through high school graduation and account for endogenous secondary schooling. Unlike Kane (1994), we model the schooling transitions leading up to high school graduation and do not focus solely on high school graduation and college attendance. This is empirically important for understanding Hispanic schooling histories because many Hispanics drop out of school before 11th grade. Cameron and Heckman (1998) document the empirical importance of controlling for educational selectivity in estimating the *ceteris paribus* effect of family variables on grade transitions.

3. Evidence on Educational Selectivity and the Dynamics of Schooling Choices: Estimates from the NLSY Data

In this section, we apply the econometric model of Section 2 to extend the univariate analyses of the Section 1 to jointly estimate the contribution of family income, family background, scholastic ability, tuition costs and opportunities in unskilled labor to schooling attainment by age and grade. Our analysis begins at age 15 and ends at age 24. In this section we also simulate the estimated model to address three questions. (1) Which variables have the most influence on schooling attainment at various ages? (2) Can differences in personal endowments and family characteristics explain gaps in White-minority schooling

17

attainment? (3) Is the estimated influence of family income on college attendance primarily a consequence of long-run family environment or short-term borrowing constraints?

We answer the first question, variable by variable, by equating the distribution of the characteristics for Blacks, Hispanics, and Whites while holding the distribution of the other characteristics at their sample levels, and measuring how high school graduation and college enrollment respond. The second question is answered by simulating schooling outcomes when minorities have the entire bundle of variables faced by Whites. We address the third question by comparing the estimated effects of family resources on schooling when scholastic ability (AFQT) or family background variables are included as explanatory variables and when they are not. We interpret AFQT as the outcome of long-term family and environmental factors produced in part from the long-term permanent income of families. Family background variables have the same interpretation. To the extent that the influence of family income measured at a point in time is diminished by the inclusion of AFQT, we can conclude that long-run family factors crystallized in AFQT scores or in measured family background variables (parental education) are the driving forces behind schooling attainment, and not short-term credit constraints experienced at age 17.

This section begins with a description of how we implement the econometric model discussed in section 2 and the goodness-of-fit tests used to evaluate the estimates produced from it. In a nonlinear model, parameter estimates are difficult to interpret. As a consequence, we focus on simulations of estimated models, which are used to address the questions posed in the introduction to this paper.

Estimating The Baseline Econometric Model

The initial conditions of the model are specified in the following way. Individuals enter our sample at age 15 either in grade 10, grade 9, or grade 8 and below. From their initial grade, they stay in school and move to the next highest grade level or drop out. Thus, the set of possible destination states at age 16 is not just the set of highest grades completed but the possible highest grades completed for the number of possible attendance states (attend or not attend). From age 16, an individual currently attending grade 11 has the option of continuing in school and graduating or dropping out. An individual who dropped

18

out after completing grade 9 at age 16, for instance, has the options of returning to school to complete grade 10, complete high school through GED certification, or not returning to school.[13] The distinction between GED and traditional high school completion is important in accounting for Black, Hispanic, and White schooling differences due to the large number of minorities completing high school through receipt of a GED.

Once a person finishes high school through GED attainment or high school graduation, he may choose to enter a two-year college, a four-year college, or not enter college at all. Once a person enters college, he is no longer followed in our analysis. If a person does not enter college immediately after high school completion, we estimate his chances of college entry until age 24, when our analysis ends. As noted in Section 1, very little college entry occurs after age 24, so this restriction is of no practical consequence.

The number of possible transitions proliferates rapidly as individuals get older. There are few observations for many of these transitions so that it is not possible to estimate the associated parameters $\beta_{a,j,c}$ or $\alpha_{a,j,c}$ with any precision. Some judgement has to be made to limit the number of estimated parameters.

Our strategy is as follows.

(1) For transitions with relatively few (less than 30) observations, we only estimate the intercepts and not the slope parameters in $\beta_{a,j,c}$. Factor loadings for these parameters are set to zero.

(2) We test for the presence of initial age (a), initial state (j) and final destination (c) interactions in the slope coefficients (denoted $\bar{\beta}_{a,j,c}$) and factor loadings ($\alpha_{a,j,c}$).

(3) We test for differences in the slope coefficients and factor loadings among Blacks, Whites and Hispanics for the transitions that are not "rare" as defined in (1), maintaining ethnic/race-specific intercepts for each transition.

---

[13] A small number of people in our sample go directly from school attendance to GED certification. In addition, a small number skip grades and graduate early. Because of the small number of people making these transitions, they are not important for our analysis.

19

Appendix C (available on request) summarizes these tests. Briefly, we find (a) strong evidence of racial and ethnic differences in slope coefficients and factor loadings; (b) that initial age matters greatly in determining schooling transitions and (c) for secondary schooling transitions, initial grade is not an important determinant of the probability of transiting to the next grade. By imposing the restrictions that are not rejected, we alleviate the problem of parameter proliferation and estimate a parsimonious description of life cycle schooling transitions for male youth. We also report that a model that accounts for η, persistent across spells, fits the data better than a model that ignores such persistent heterogeneity.

All of the baseline models and tests are developed for the case where AFQT is excluded from the model. We do this because of the controversial nature of the AFQT variable and because of the potential endogeneity of the variable. This strategy gives us a conventional benchmark framework against which we can compare a model that includes AFQT as a regressor.

We describe the final estimated model by discussing, in the following order, (1) the determinants of the probability of the initial condition; (2) the determinants of secondary school transitions for those who attend school; (3) the schooling decisions for those who have left school; and (4) the college entry decision. In all specifications, the family income and background variables described in Table 8 (highest grade of mother, highest grade of father, number of siblings, and whether or not the person lived with both biological parents at age 14), average wages[14] in the local labor market, college proximity, and two-year tuition costs net of Pell grant subsidies at public colleges are included in transitions with slope coefficients. Year dummies are entered in all specifications. In addition to estimating the determinants of college entry, net four-year tuition were entered as a price of four-year college entry and net two-year tuition was included as a two-year college price (See Appendix A for a more complete discussion of these variables). Finally, all models are estimated both with and without the AFQT score, and all models are estimated separately for Blacks, Hispanics, and Whites.

---

[14]In general, the local unemployment rate was found not to affect schooling choices at any level. See below for more details.

(1) Initial Grade Level. Few individuals leave school before age 15 (less than 1% of Whites and Blacks less than 3% of Hispanics are out of school at age 15). The mode grade level at age 15 is 9, with most of the remaining sample in grades 10 or 8. Hence, initial grade levels are estimated by a multinomial logit (conditional on η) for grades 10, 9, or 8 or less.[15] Further disaggregation of "8 or less" into "grade 8" and "grade 7 or less" was not empirically important as judged by model selection criteria.

(2) Secondary School Transitions for School Attenders. Because exiting school with a GED is a rare event (most GED attainers leave school for at least a short period before GED certification), the probability of moving to this state is estimated only with an intercept term (that is, the slope parameters defined as $\bar{\beta}_{a,j,c}$ and factor loading $\alpha_{a,j,c}$ are set to zero.) In addition, because secondary school transitions are almost nonexistent after age 19, these transitions were also modeled with an intercept only.

A more important restriction concerns the origin educational state on secondary school continuation probabilities. For instance, one might ask whether at age 15, the coefficient vectors (both $\beta_{a,j,c}$ and $\alpha_{a,j,c}$) governing school continuation are the same regardless of whether a student is initially in grade 8 or less, grade 9, or grade 10. This restriction can be tested in two ways. The most restrictive version of the test assumes that all coefficients including the intercept are identical. The second version of the restriction allows an intercept alone to depend on the origin state. Hence, except for dummy variables that represent grade levels at age $a - 1$, the slope coefficients on $Z_{a,j}(\bar{\beta}_{a,j})$ are identical. The weaker hypothesis is not rejected at any age for any racial-ethnic group. Hence, the only interactions with the current state for the estimated model parameters displayed in Table C-5 of Appendix C are intercept terms.

(3) High School Dropouts. High school dropouts face three choices: return to school, obtain a GED, or neither. Because both "return" and "GED" are relatively rare events, these transitions are modeled with a single multinomial logit for all ages 16 through 20 with age effects introduced through a set of indicator variables for age.

---

[15] A small number of individuals, less than 1 percent, were already in grade 11 at age 15 and were deleted from the analysis. See Appendix A for full details of sample-inclusion criteria.

(4) College Entry. Statistical tests reject the hypothesis that the determinants of two-year college entry are the same as those of four-year entry. Indeed, separating college entry into two and four-year entry is important in explaining White-minority differences in schooling attendance.[16]

Goodness-of-Fit and the Importance of Unobserved Heterogeneity

A natural question to ask is whether it is necessary to control for unobserved heterogeneity. Cannot a simpler scheme account for schooling histories that ignores serially correlated unobservables? A Bayesian Information Criterion (BIC) model selection procedure rejects a specification with no heterogeneity for all racial-ethnic groups. (See Appendix C-5, available on request).

A second way to judge the importance of heterogeneity is to examine whether or not introducing it makes a difference in estimated coefficients and on estimated marginal probabilities. The impact of background and economic variables is generally much stronger at each grade once account is taken of the selective nature of schooling attainment status. A similar result is reported in Cameron and Heckman (1998). Controlling for heterogeneity is · especially important for the analysis of Hispanic college transitions. Recall that Hispanic college attenders are a very select group. For the sake of brevity we delete these results.

A third way to gauge the importance of correcting for unobserved heterogeneity, η, is through goodness-of-fit tests. Using a variety of predictions of schooling attainment probabilities at different ages, we find that models with heterogeneity predict better schooling. Therefore, all the simulations conducted below are for a model that includes the heterogeneity correction. Specifications with and without AFQT will figure prominently in our interpretation of the data.[17]

---

[16] See Tables 35 and 37 of Cameron 1996.

[17] It is of some econometric interest to notice AFQT still plays a large role in explaining the dynamics of schooling attainment even after the heterogeneity control is applied. Apparently, the unmeasured factor, sometimes called "ability", does not fully capture the intertemporal correlation in schooling decisions nor does it supplant AFQT in explaining the dynamics of schooling attainment. See Cameron and Heckman (1998) for further evidence on this point.

22

Are There Differences in Racial-Ethnic Schooling Behavior?

A central concern of this paper is assessing the relative importance of endowments, resources and prices facing agents and differentials in racial and ethnic responses to identical constraints and opportunities in explaining racial-ethnic gaps in educational attainment. We take three approaches to making this assessment. First, each model is estimated separately by race, and tests for parameter equality between racial-ethnic groups are conducted age-by-age to pinpoint any differences that exist. Second, we conduct counterfactual simulations to summarize the overall quantitative importance of behavioral differences (parameters) versus endowments (covariates) in explaining racial-ethnic schooling differentials. Third, we present side-by-side comparisons of the effects of individual variables.

The schooling model is estimated separately for Blacks, Hispanics, and Whites (including separate heterogeneity distributions $F(\eta)$). While maintaining separate heterogeneity distributions, the behavioral parameters (all the $\beta_{a,j,c}$ parameter vectors) are tested for equality. In all cases, racial-ethnic equality is rejected, indicating the possibility that differences in response to the same variables play an important role in explaining schooling differences.

One way to decompose a schooling attainment gap into sources due to behavior (parameter differences) and sources due to endowments (covariate differences) is described here. Let $\hat{\beta}_w$ and $\hat{\beta}_b$ denotes estimates of the White and Black parameters vectors, where $\hat{\beta}_w$ and $\hat{\beta}_b$ are inclusive of the $\beta_{a,j,c}$, $\alpha_{a,j,c}$ and $F(\eta)$ estimated for each group and let Z_b and Z_w represent draws from the Black and White covariate distributions, respectively. The probability of completing high school at age 24, say, is the sum over the probability of all paths that lead to GED attainment and high school graduation by that age. For convenience, let $\Pr(\cdot)$ denote this overall probability of high school completion. The following expression shows one way of decomposing the difference in predicted high school graduation rates:

$$(6a) \quad E_{Z_w}\Pr(Z_w\hat{\beta}_w) - E_{Z_b}\Pr(Z_b\hat{\beta}_b)$$

$$= E_{Z_w}\Pr(Z_w\hat{\beta}_w) - E_{Z_b}\Pr(Z_b\hat{\beta}_w) + E_{Z_b}(\Pr(Z_b\hat{\beta}_w) - \Pr(Z_b\hat{\beta}_b))$$

$$= \textbf{gap due to endowment difference} + \textbf{gap due to behavioral difference.}$$

The "endowment difference" is the fraction of the schooling gap attributable to differ-

23

ences in background, local job market conditions, and other variables evaluated at White parameter estimates, while the "behavior difference" is the portion attributable to the divergence in parameter estimates evaluated at the Black distribution of covariates. An alternative way of making the decomposition evaluates the "behavioral difference" at the White covariate distribution and the "endowment difference" at the Black estimates:

(6b) $\qquad E_{Z_w}(\Pr(Z_w\hat{\beta}_w) - \Pr(Z_w\hat{\beta}_b)) + (E_{Z_w}\Pr(Z_w\hat{\beta}_b) - E_{Z_b}\Pr(Z_b\hat{\beta}_b))$

$\qquad$ = **gap due to behavioral difference + gap due to endowment difference**.

The difference between (6a) and (6b) is that the fraction of the schooling gap explained by differential behavior (equating the covariate distributions) is evaluated using the White distribution in (6b) while it is evaluated using the Black distribution in (6a). If the estimated parameter vectors are not identical, then the order in which the decomposition is formed will affect the outcome of it because the estimated parameter vectors will weight the various elements of the covariate vector differently. In principle, one can have two very different estimates of the relative importance of endowment and behavior in explaining racial-ethnic schooling differentials.

Tables 10 through 12 report sample analogues of counterfactual schooling attainment gaps that equate the covariates at both the White and minority distributions:

(7a) $\qquad E_{Z_w}(\Pr(Z_w\hat{\beta}_w) - \Pr(Z_w\hat{\beta}_b))$

$\qquad$ and

(7b) $\qquad E_{Z_b}(\Pr(Z_b\hat{\beta}_w) - \Pr(Z_b\hat{\beta}_b))$.

Expression (7a) is the attainment gap explained by equating endowments at the White distribution (the "behavioral difference" of (7a)), and (7b) is the counterfactual evaluated at the minority distribution (the "behavioral difference" of (6b)). If covariates fully explain schooling differences, then both expressions will be approximately zero. If behavior explains schooling differences, then both expressions should approximately equal the actual schooling gap.

Table 10 presents predicted schooling-attainment gaps and counterfactuals that summarize the quantitative importance of differences in racial-ethnic "behavior" and endowment

and opportunities distributions in accessing schooling differences.[18] These simulations are computed at four levels and shown in panels A through D: the initial state at age 15, high school completion (GED attainment and high school graduation combined), college entry conditional on high school completion (combined two- and four-year entry), and overall college entry (unconditional on high school completion), which is a function of the previous levels. While high school completion and college entry have received a good deal of attention, the importance of family background in explaining early school differences is little studied.

The first row of each panel shows the actual gap between Whites and minorities. Panel A shows that age 15, the gap between Whites and Blacks achieving grade 9 or higher is 16 percentage points. Standard errors of the predictions are given in parentheses, and a "(*)" is used to denote statistical significance at the 5 percent level. The second and third rows of each panel present counterfactual simulations of schooling gaps when the distributions of endowments faced by Whites and minorities are equalized (this includes all family background, family income, and other variables used in the analysis). The predicted gap when White endowments are given to minorities (corresponding to (7a)) is shown in the third row, and the predictions when minority endowments are assigned to Whites (corresponding to (7b)) are presented in the third row. A positive number means that Whites are more likely to achieve the indicated schooling level, while a negative number says the opposite. Finally, the fourth and fifth rows of each panel conduct these simulations when AFQT is included in the model.

Row (1) shows the predicted percentage point gap in completion of grade nine or more at age 15. The White-Black and the White-Hispanic gaps are .16 and .21 respectively. Rows (2) and (3) show that equating endowments eliminates virtually all of the schooling gaps. What is left over is not statistically significantly different from zero. Adding AFQT to the model produces the same results except for Blacks, who are predicted to be significantly

---

[18]In all cases, schooling attainment predicted using estimates and sample covariates were identical to corresponding sample values up to the third significant digit, indicating that predicted values were reliable estimates of actual values.

25

more (6 percentage points) likely than Whites to be at the mode grade or above at age 15.

One potential problem with the use of the AFQT score in these data is that the test was taken in 1980, while the sample was still of high school age. Thus, because AFQT is both a cause and a consequence of schooling there may be a problem with reverse causality. However, for our subsample of the NLSY, AFQT was measured before any member of the sample was eligible for college entry and before high school graduation. Thus, AFQT is predetermined with respect to college entry and high school graduation.[19]

Panel B presents counterfactual simulations of giving minorities the white endowment for high school completion rates by age 24. Of the actual .06 and .14 percentage point gap between White-Black and White-Hispanic high school completion (row (6)), rows (7) and (8) show that although Whites and Hispanics are equalized, Blacks are significantly more likely to complete high school. Adding AFQT favors minorities even more. Blacks have a 7 to 12 percentage point advantage over Whites; Hispanics have a 2 to 9 percentage point advantage.

Minority high school completers lag behind Whites in college entry by 11 and 7 percentage points (row (11)). Assigning minorities White endowments, minorities gain an advantage between 0 and 11 percentage points over Whites, which rises to between 8 and 15 percentage points when AFQT is equalized.

Panel D decomposes the *population* probability of college entry (*i.e.*, not conditional on high school completion). The gap in the unconditional college entry probabilities is a function of the differentials in panels A through C. The predicted gap in college entry for Blacks and Hispanics compared to Whites is 12 and 14 percentage points, respectively. The results for this schooling gap accord with what we have found for the other gaps. Equating family background between Whites and Blacks and Whites and Hispanics raises minority schooling above the White rate. This is especially true when scholastic ability, as measured by the AFQT, is included in the empirical model.

This finding for college attendance is consistent with other results reported in the litera-

---

[19]In addition, our measure of AFQT is age adjusted. Estimates using age - and education-adjusted AFQT are similar to the ones we report in the main tables.

ture, but the other gaps in educational attainment have not been investigated.[20] The effect of equalization of background variables in making college entry higher for Blacks than for Whites has been noticed in a number of highly restricted models of high school completion or college entry (usually in a probit model as the estimated coefficient associated with a binary indicator of racial-ethnic identity). However, previous research does not control for educational selectivity as we do, leaving open the possibility that the effect is an artifact of selection bias. Our analysis reveals that the effect remains even controlling for selection bias.

The effect of equalization on attainment of earlier levels of education has not been studied, nor have more general models for all levels of attainment with separate slope coefficients by race. We reject the hypothesis of equality of response to common opportunities (that is, common slope parameters for Whites and minorities) that is assumed in most studies of ethnic and racial differences in educational attainment. Ours is the most general analysis yet performed on this question because we consider more grade transitions and because we control for educational selectivity which, if uncontrolled, could generate differences in racial and ethnic coefficients.

Table 6 shows large ethnic and racial differences in the choice between two- and four-year college entry. Table 10 displays simulations from an extension of the baseline model that disaggregates the college-going state into attendence at two and four year colleges. The initial conditions and the secondary school transitions are identical to those in the baseline model, so only the simulations corresponding to college entry given high school completion and overall college entry are presented.[21]

Panel A of Table 10 shows substantial differences in the simulated gaps. Whether evaluated at the minority or majority covariate distributions, the counterfactual gaps are much closer to zero than those of Table 9, particularly for Hispanics. Large differences remain,

---

[20]See Hauser, 1991, for instance who uses parental background but not AFQT in reaching the same conclusion we do regarding college enrollment.

[21]The age-specific estimates for each racial-ethnic group are exhibited in appendix Table C-5. Available on request from the authors.

however. Counterfactuals for overall college entry are shown in panel B. For Hispanics, the predicted gap evaluated at the Hispanic covariate distribution falls from a 13 percentage point underprediction to an 11 percentage point overprediction (row three). For Blacks, a 13 percentage point underprediction falls to 8 percentage points overprediction. The finding that Blacks and Hispanics are more likely both to complete high school and to enter college is robust across a number of specifications.[22] Even disaggregating by type of college attended, we still find that adjusting for family background, minorities are *more likely* to attend college, even after controlling for selection.

We shed some light on the issue by examining the individual determinants of school for each of the four schooling levels just analyzed.

Effects of Individual Variables

Table 11 presents evidence on the two questions just posed. It analyzes the sources of schooling differences between Whites and minorities by decomposing the gap in high school graduation and college enrollment rates into the contribution made by each explanatory variable. It also presents evidence on the robustness of family income and other family factors when AFQT is included as an explanatory variables. Because racial and ethnic groups may vary in their sensitivity to differences in family income or tuition policies, the schooling attainment predictions displayed in the table are based on estimates of the econometric model that are made separately for each racial-ethnic group.

Panel A of Table 11 presents counterfactual simulation results for completing 9th grade by age 15, the initial condition (or first schooling attainment state) for our model. Panel B presents simulation results for completing high school (by graduation or exam certification), Panel C presents counterfactual simulation results for entry into college conditional on completing high school and Panel D presents counterfactual simulation results for entering college not conditioning on high school attainment. This final simulation measures the net effects of background on college entry operating through schooling completion at all prior stages. The last row in each table shows the actual White-minority gap in attainment

---

[22] Disaggregation of the schooling and dropout states into work and schooling states produce the same conclusions. These results are available on request from the authors.

28

which in all cases is very close to the gap predicted from the model.

Rows (1) through (8) present the changes in the schooling gaps when the variable named in the left-hand column is adjusted to the White level while holding the other variables fixed at sample values. For example, the number in column (1) and row (1) of panel A shows that if the four components of family background listed in rows (1a) through (1d) are adjusted for Blacks to White levels, then the Black rate of completing ninth grade completion rate is only 9 percentage points lower than the White rate rather than the 16 percentage point gap actually found in the data.[23] Rows (1a) through (1d) show the change in the gap when the individual components of family background are equated. Row (2) shows the change when minority family income is adjusted to White levels, and row (3) displays the same change when county average wages are equalized (see Tables 7 and 8 for definitions and group means). Next, row (4) shows the effect of adjusting tuition and for college proximity, and row (5) shows the effect of equalizing ability test scores. Finally, rows (6) through (8) show the combined effects of various incremental simulations. Columns (1) and (2) show the predicted gap between Whites and Blacks and Whites and Hispanics, respectively. Columns (3) and (4) show the corresponding calculation when AFQT score is included in the set of explanatory variables as a proxy for long-term family and environmental influences.

From columns (1) and (2) in each panel, we reach several important conclusions. First, row (2) shows that adjusting for family income alone largely eliminates the White-minority gaps in high school completion and much of the gap in college entry. This confirms the univariate analysis presented based on Table 6. For high school completion, all of the White-Black gap would completely vanish as would most of the White-Hispanic gap (compare rows (2) and (9)). A similar conclusion holds for college entry, though the effect is not quite

---

[23] In these simulations, only the means of the variables are equated between the two groups. Consider the adjustment for sibling size. This is accomplished by finding the difference in the mean number of siblings. Note that because these models are nonlinear, there are other ways of making these simulations, such as using the marginal distributions of White attributes but preserving the original covariance structure, but these more elaborate methods have little impact on the simulations. We report the simplest and most easily replicable results.

as dramatic. Equating family income raises Black and Hispanic college enrollment 6.3 and 5.4 percentage points. The actual gap for both groups is 13 percentage points.

Before concluding that family income is the whole story in explaining racial and ethnic schooling differentials, it is important to examine rows (1a) through (1d) which show that a similar story can be told for family background factors. The effects of equating family background (holding family income and other explanatory variables constant) are particularly strong for attending college. For Blacks and Hispanics, they explain respectively 11 and 12 percentage points of a 13 percentage point gap between Blacks and Whites.

Local labor market variables representing the opportunities available to persons with little education are statistically significant in our estimated model but play only a modest role in explaining differences in schooling continuation decisions. They contribute little to explaining White-minority differences (row (3)). The reason is that average differences in local labor market conditions among the groups are very small (see the means in Table 8). Hence, even though their effect on all levels of schooling attainment is somewhat important, their effects are neutral across racial groups and explain little of the racial schooling attainment gaps.

Equating tuition and college proximity (row (4)) does little to narrow the White-minority gap. In fact, adjusting for tuition and college proximity increases the gap between Whites and minorities for both college entry (both unconditional and for high school graduates) and high school completion.[24] The reason for this is apparent from the means in Table 9. Blacks and Hispanics face lower average tuition and have more geographic proximity to college than Whites. Equating minority to White levels lowers college participation for Blacks and Hispanics.[25]

Row (6) in each panel of the table shows the combined effect of adjusting Black (column

---

[24]College tuition and proximity were included in the explanatory variables for high school completion.

[25]College entry may be a primary motivator of high school graduation. Thus, higher college costs could lead to lower high school graduation rates. In our joint work, we find little evidence for such an effect. Both tuition and proximity variables are small in magnitude and statistical significance for all groups. For college entry, by contrast, both college tuition and proximity are statistically significant and numerically important for all race groups, though adjustment widens the White-minority gap slightly.

(1)) and Hispanic (column (2)) levels of both family background and family income to White levels. Family background and income are the main determinants of all four attainment statuses that we study. Other variables explain little of the White-minority schooling gaps. Comparing rows (6) and (9) shows that family background and family income together completely explain White-minority schooling gaps. These factors are powerful predictors of the four attainment statuses we study. They explain most of the gap in ninth grade schooling attainment at age 15. For the higher schooling attainment statuses, in all cases except one (Hispanic high school completion), adjusting family background and family income to White levels over-predicts the schooling attainment gap. Blacks and Hispanics are more likely to complete high school and attend college compared to Whites with the same levels of family background and income. The same conclusion applies to row (7), which shows the change in the gap when all minority variables listed in rows (1)-(4) are equated to White levels.

Columns (3) and (4) repeat the same simulation exercises except scholastic ability as measured by AFQT is included as an explanatory variable. Comparing the family income and background effects in row (2) when AFQT is in the model (columns (3) and (4)), and when it is not, demonstrates that the effect of the family income variables for high school completion, college attendance given completion, and college attendance is substantially weakened by the inclusion of AFQT scores. What is noteworthy is that the effects of including AFQT on weakening the estimated family income effect is largest on the later schooling transitions. Family income plays an important role in high school dropout and completion decisions even conditioning on AFQT, especially for Blacks.

Equalization in schooling attainment appears in a somewhat less dramatic fashion when family background variables other than income are included in the model (see row (6)). Thus our conclusions do not rely solely on estimates based on the AFQT. Long-term family factors and not family income account for much more of the college enrollment gap. This does not say that long-run family income does not matter in explaining college attendance. However, it does say that short-run liquidity constraints experienced at the college going ages are much less important than the long-run factors in promoting college attendance.

31

Equating family background and income raises Hispanic high school completion and college entry for high school graduates by 12 and 13 percentage points, respectively, when AFQT is not included. The corresponding gaps when AFQT is included are only .2 and 3 percentage points, respectively. If minority AFQT alone were adjusted to White levels, then college entry would rise by 18 percentage points for Blacks and 16 percentage points for Hispanics (row (5) columns (3) and (4) in Panel D). The effects of conditioning on AFQT in the college entry equation for high school completers are 15 and 12 percentage points for Blacks and Hispanics, respectively. Regardless of income and family income background, at the same AFQT level Blacks and Hispanics enter college at rates that are substantially *higher* than the White rate. The predictions for high school completion are similarly dramatic. The role of AFQT in explaining racial and ethnic schooling differences is thus seen to be very important. It is long-run factors that promote scholastic ability that explain most of the measured gaps in schooling attainment and not the short-run credit constraints faced by students of college going age that receive most of the attention in popular policy discussions. These long run factors that promote college readiness are · proxied by AFQT. Even if we exclude AFQT from the analysis, parental background factors play essentially the same role as AFQT, although the effects are weaker.

The next two tables confirm the results in Table 11, but in a different way. The tables show estimated average derivatives of the probability of completing various levels of schooling by age 24 with respect to the covariates. Results are given separately for Blacks, Hispanics, and Whites. Standard errors are given in parentheses. Table 12 shows estimates when AFQT is not included in the model, and Table 13 shows estimates when it is. Comparing corresponding entries for family income effects in panel C of Tables 12 and 13 shows that AFQT eliminates the effect of family income on college entry for high school completers, but it does not eliminate the effect of family income on earlier schooling transitions. Family income still plays a role at earlier grades, albeit a diminished one when AFQT is included in the model. If family income is interpreted as a source of short-run credit constraints, credit constraints are more important on the high school dropout and completion decision than in the college enrollment decision.

32

Table 14 recasts the major findings reported in this section in yet another way. Rather than showing how schooling gaps change in response to adjustments in the explanatory variables, it shows the predicted change in schooling attainment in response to a substantial $10,000 rise in family income on college entry for each racial-ethnic group, both with and without AFQT included in the prediction equations. Comparing rows (1) and (2) of panel A shows that family income effects of college enrollment are reduced to about one-third of their size when AFQT is included in the model. (Similar results are found for the effect of the same increase in family income on college enrollment not conditioning on high school graduation. Again, the effects of family income at earlier transitions are stronger). Panels B and C of the table make an additional point. Even when AFQT is not included among the explanatory variables, family income is only a trivial determinant of two-year college entry. Thus, comparing rows (1) and (2) of the panel reveals that the measured effect of family income is largely due to its effect on four-year entry and not to its effect on two-year college entry. Including AFQT among the explanatory variables (panel C) does not reverse this conclusion; it only weakens the estimated effect of family income on four-year college entry. This evidence reveals that controlling for ability, the estimated relationship between income and college entry is mainly a relationship between income and entry into four-year college.

It might be argued that our measure of family income is riddled with measurement error and that our finding of a weak current income effect conditional on long-term factors is a consequence of this. Elsewhere (Cameron and Heckman, 1998), we address this problem and find that when we use the average of family income over a few years, we find a still weak effect of family income on schooling. We perform similar estimation for the model reported here. Using two and three year averages of family income, we find results very similar to those we report in this paper.[26]

It is important to notice that family income plays a more important role on earlier

---

[26] These simulations are available on request from the authors. Multiple measures of family income are only available for a substantially smaller subsample of the data. For that reason we report the results for one year income in this paper.

schooling transitions compared to its effect later ones. This evidence suggests that it is family income at earlier ages and not later ones that matters in explaining college attendance. Its effect, however, is on college readiness (high school completion) and not directly on college attendance for high school graduates.[27] If family income effects conditional on long-term factors are a measure of short-term credit constraints, such constraints are more important at earlier ages rather than later ones.

The Effects of College Tuition

Table 15 presents simulations from another version of the model[28] which considers the effects of a rise in both two-year and four-year gross tuition costs by $1000 on two- and four-year college entry for high school graduates for each racial-ethnic group. We estimate three separate specifications: (1) a model of college entry including year effects alone (no family background or AFQT effects); (2) a model that adds family background and other variables (the baseline specification used in this paper), and (3) a model that adds AFQT to the specification in (2). All specifications control for unobserved heterogeneity allowing for separate distributions in each ethnic/race group. The simulations based on his model show both own and cross effects: that is, the effect of two-year tuition on two-year entry (own effect) and four-year entry (cross effect).

Controlling for family background greatly weakens estimated tuition effects. In specifications with AFQT included as a regressor there are virtually no four-year tuition effects. Two-year tuition effects are large and statistically significant, though a substantial portion of the decline in two-year attendance is due to enrollment in a four-year college instead. Four-year tuition effects are inconsequential and sometimes perversely positive, though statistically insignificant.

Table 16 shows the effect of a $1000 increase in tuition on college entry probabilities (two- and four-year schools aggregated) for several different models.[29] A baseline set of

---

[27] This finding is consistent with the evidence of Duncan, Brooks-Gunn, Yeung and Smith (1998).

[28] These results are discussed in detail in Cameron (1996, Table 39).

[29] That is, the table shows how high school graduates would respond to changes in tuition policy. As noted above, college tuition has a negligible influence on secondary school completion choices. The results for college entry not conditioning on high school graduation are very similar to the estimates reported

34

simulations for each race group are shown in row (1) of panel A. The estimates are for entry at two- and four-year colleges combined using the estimates reported in Table 16. The simulated effects for Whites, Blacks, and Hispanics are shown in columns (1), (2), and (3) respectively. These estimates are close to the median estimate (after an inflation-adjustment) presented by Leslie and Brinkman (1986) in their survey of over 20 studies the effect of tuition on college enrollment.

Consistent with the disaggregated results shown in Table 15, including family background variables (as defined in Table 7) greatly diminishes the estimated effect of tuition on schooling. Adding AFQT further weakens the estimated effect. Nonetheless, the effects of tuition on enrollment are substantial even controlling for AFQT. However, given the pattern of smaller Black responses to tuition, the recent rise in college tuition works to *reduce* Black-White differences in college enrollment. The more negative effect of tuition on college enrollment for Hispanics accounts for only a small component of the difference between Hispanic and White college enrollment rates. For all groups, the estimated effect of the tuition increase operates mainly through its effect on attendance at community colleges where there is greater sensitivity to tuition. There is no effect of an increase in college tuition on enrollment in four-year schools.

Panel B shows how estimated responses to tuition vary by family income quartile (these numbers were calculated from a model that included income-tuition interaction terms). For each demographic group, the top quartile is much less responsive to tuition increases than are the other quartiles although these effects are never statistically significantly different from the tuition effects for other quartiles for each demographic group using conventional levels of statistical significance. Such differences in response to tuition by income class have been noted in the literature (Kane, 1998) and have been interpreted as evidence that low-income persons are more responsive to tuition because of borrowing constraints. However, one cannot reject the hypothesis that the tuition coefficients are the same across income classes for each demographic group. Controlling for ability, the estimated effects of

---

in the text and for the sake of brevity are not reported. These tables are discussed at greater length in Cameron (1996, Table 39).

tuition on enrollment diminish, especially for Blacks (see the estimates reported in Panel C). Again, one cannot reject the null of equality of the coefficients on tuition across income quantiles. Panel D reports the response to the tuition increase across AFQT quartiles. For all demographic groups, one cannot reject the hypothesis of equality of the estimated effects of tuition on college going across income levels.

Table 17 reports the sensitivity of estimated tuition effects to inclusion of regional variables and to adjustment for Pell Grants. Consistent with the evidence reported in Kane (1994) introduction of geographical dummy variables weakens estimated tuition coefficients. Adjusting tuition for eligibility for Pell Grants, however, has only a minor effect on the reported estimates. A $1000 increase in Pell Grants entitlements produces less than a 1 percent increase in enrollments as opposed to about a 6 to 8 percent response from a comparable change in tuition. Pell Grants offset tuition costs for low-income people and should have effects of equal but opposite magnitude to tuition. This difference in responses is consistent with evidence that many Pell Grant-eligible individuals do not apply for the grant because they lack the necessary scholastic preparation for college.

Orfield (1992) and Kane (1997) speculate that poor people may not have reliable information about their own Pell Grant eligibility. Heckman, Smith and Wittekind (1997) show that among persons eligible for the JTPA program targeted toward the poor, the poorest and least educated are less likely to be aware of Pell Grants. Thus family background factors play a substantial role in making people aware of their Pell Grant benefits. However, it seems unlikely to us that children from poor families who have persevered through high school would be unaware of their eligibility for grants and loans. It seems more likely that able and college ready students are aware of their eligibility for grants and act on the information. Indeed, Heckman, Smith and Wittekind (1997) find that low-income high school graduates are much more aware of their eligibility for Pell Grant programs than are high school dropouts.[30]

---

[30]They report that the log odds ratio of a high school graduate with no college education being aware of the Pell Grant program is twice as great as the log odds ratio for a high school dropout. The Orfield argument might be salvaged by claiming that persons who were unaware of Pell Grants in high school

Table 18 reports a test of one version of the credit constraint hypothesis that is commonly stated in the literature. It postulates that if students are severely credit constrained and family income limits their attendance, a $1000 rise in tuition should have equal and opposite effects on schooling decisions as a $1000 rise in family income.[31] Panel A reports results when AFQT is excluded from the model. Panel B reports results when it is included. The estimated effects of family income on schooling are an order of magnitude smaller than the estimated effects of tuition when AFQT is excluded from the model. When AFQT is included, family income effects fall by a factor of three-fourths while estimated tuition effects are barely affected. When AFQT is excluded, for Whites and Hispanics, one can decisively reject the hypothesis of equality of absolute values at conventional significance levels. For Blacks, the hypothesis is rejected at the 10% level. When AFQT is included, the estimated family income effects become so imprecisely determined that one cannot reject the hypothesis of equality, although the point estimates are dramatically different. The estimated income effects also become numerically very small. The key to this result is that estimated family income effects are not statistically significantly different from zero. Tuition and family income play fundamentally different roles in college attendance decisions. The hypothesis of no binding credit constraints at the stage of the life cycle where college decisions are being made is consistent with this evidence.

Supporting Evidence From Other Studies

Shea's (1996) research provides some support for our analysis. He argues that in order to identify the contribution of credit constraints to schooling choice, one needs to look at how unexpected changes in family income influence educational attainment and hence the earnings of children. Using the Panel Study of Income Dynamics data, he estimates the correlation between children's income and variations in father's labor income due to predictable and unpredictable components. (An instance of the latter is job loss through plant

---

failed to graduate because they were unaware of the grants, overestimated the costs of college and failed to complete high school because of their underestimate of the true rate of return.

[31]This widely invoked intuitive model ignores the important point that price effects of tuition should make income and tuition effects different even if there are credit constraints.

shut down; an instance of former is father's education). Shea finds that predictable components of family income are positively correlated with the child's income, but components of family income due to "luck" are uncorrelated with child income. He finds no evidence that unpredictable components affect children's income and hence he finds no evidence of a strong role for short-term credit constraints.[32]

Other indirect evidence is provided by Altonji and Dunn (1996). If short-term credit constraints do not hinder college entry, and access to credit markets is available to everyone, there should be no relation between the returns to education and family income provided parental consumption motives for the schooling of their children are unimportant. Once a student graduates from high school, he or she should invest until the return from another dollar spent on education is the same as the return on physical assets. Altonji and Dunn find that rates of return are more or less constant across individuals and do not vary systematically by family income.[33] Thus, their evidence supports evidence reported in Table 6 that short-run credit market constraints are not a significant determinant of schooling and college choices.

Another piece of evidence comes from Cameron and Heckman (1998) who analyze the determinants of grade by grade schooling attainment for five cohorts of American males. While policy makers and economists have focused on the measured effects of family income on college entry, we find that family income and family background factors are powerful determinants of choices at all stages of the schooling process from the decision to complete elementary school through entry into graduate school. There is a stable behavioral relationship between family background and schooling continuation across all levels of schooling, from elementary school through entry into graduate school. This is consistent with the evidence reported in Table 6. Analysts who have only studied the relationship between college

---

[32] A major problem with Shea's analysis is that he does not justify which components of family income are predictable and which are not. In addition, he does not separate the effects of luck on younger children from the effects on older chilren. His evidence suggests that unexpected lottery winnings would have no effect on schooling choices regardless of their time of arrival.

[33] If there were a strong consumption motive for education, the rate of return to schooling should be lower for children from higher income families.

choice and family income have ignored the fact that family income and other measures of family background, such as parental education, are also strong determinants of decisions to complete the ninth grade, and to attend and graduate high school where tuition costs are effectively zero. An appeal to borrowing constraints is not required to explain the relationship between family income and college attendance decisions.

The importance of family background in determining educational choices is found in other policy environments and other countries as well. The collection of papers in Blossfeld and Shavit (1993) studies data from a variety of countries in different stages of political and economic development with profoundly different institutional organizations for education. They document roughly the same basic pattern found in the U.S. data: family background and family income are important determinants of schooling choices from the earliest grade levels to the highest irrespective of whether college tuition costs paid by students are large or small.

4. Summary and Conclusions

This paper examines the determinants of college entry. The strong correlation between college attendance and family resources is widely interpreted as evidence that short-term borrowing constraints impede enrollment. We argue that the importance of short-term credit constraints is greatly exaggerated. It is the long-term influence of family income and family background as captured by our measure of ability, or equivalently by parental education, that best explains the correlation. Family income matters but its greatest influence is on forming the ability and college-readiness of children and not in financing college education. Family income is more important in explaining earlier grade transitions than later ones, suggesting that money spent on tuition policy aimed at high school graduates does not target the right population.

We apply this analysis to examine racial-ethnic differences in schooling. We find that controlling for family background, minorities are more likely than Whites to graduate high school and attend college. Again, it is long-term factors that mainly account for this relationship, not short-term cash constraints that can easily be fixed by Pell Grants or other

transfer programs offered to children late in their life cycle of adolescent development. It is early differences in resources and not later ones that matter more. Tuition and opportunities in unskilled labor markets play only a minor role in accounting for majority-minority differences in college enrollment.

We raise a number of questions about the empirical and intellectual foundations of current government income-subsidy programs designed to promote college attendance. The edifice in place is currently very generous to minorities and to children from poor families. A main conclusion from our work is that to raise college attendance and success in college, policy should focus on ensuring that more students graduate from high school and obtain the skills and motivation required to perform successful college work. Government policies such as Pell Grants and other tuition subsidies focus on getting high school graduates into college, but our evidence suggests that the scope for such policies is minimal because most of the problem of disparity in schooling attainment among racial, ethnic and income groups arises at earlier points in the life cycle of children from poor families.

References

[1] Altonji, J. and T. Dunn, 1996. "Returns to Education and the Family," Review of Economics and Statistics, Vol. LXXVIII(4), 692-704.

[2] Arellano, M. and B. Honoré 1998, "Panel Data: Some Recent Developments," forthcoming in J. Heckman and E. Leamer, eds., Handbook of Econometrics, Vol. 5, Amsterdam: North Holland.

[3] Bartholomew, D.J., 1973. Stochastic Models For Social Proceses. 2d Edition, Wiley.

[4] Blossfeld, H. and Y. Shavit, 1993. Persistent Inequality: Changing Educational Attainment in Thirteen Countries, Boulder, Co: Westview Press, 1993.

[5] Callan, P., 1997."Stewards of Opportunity: American's Public Community Colleges," Daedalus, 126(4), 95-112.

[6] Cameron, S., 1996. "Three Essays on the Demand for Education." Dissertation, University of Chicago, Department of Economics.

[7] Cameron, S. and J. Heckman, 1998. "Life Cycle Schooling Decisions: Models and Evidence." Journal of Political Economy, Vol. 106(2), 262-311, (...rst presented at the University of Wisconsin, June, 1990).

[8] _____ 1996. "Dynamics of Educational Attainment For Blacks, Whites and Hispanics," expanded version of this paper, University of Chicago, Department of Economics.

[9] _____ 1994. "The Determinants of High School Graduation and College Attendance." In Training and the Private Sector, ed Lisa M. Lynch, University of Chicago Press.

[10] _____ 1993. "The Nonequivalence of High School Equivalents." The Journal of Labor Economics 11(1, pt. 1) (January): 1-47.

[11] _____ 1992. "Comment on Hauser," in Studies of Supplies and Demand in Higher Education, ed. by Charles Clotfelter and Michael Rothschild, University of Chicago Press, 105-119.

[12] Card, D., 1996, "Schooling, Earnings and Ability Revisited," in S. Polachek, ed., Research in Labor Economics, 14, Greenwich, CT, JAI Press.

[13] Cavallo, A., J. Heckman, J. Hsee, and E. Vytlacil, 1998. "Understanding The GED," unpublished book manuscript, University of Chicago, Department of Economics.

[14] Chen, X., J. Heckman and E. Vytlacil, 1998, "Identi...cation and $\sqrt{N}$ Estimation of Semiparametric Panel Binary Variable Models For Single Spells With A Factor Structure," unpublished manuscript, University of Chicago, under revision.

[15] Clotfelter, C., R. Ehrenberg, M. Getz and J. J. Siegfried, 1991. Economic Challenges In Higher Education, Chicago, University of Chicago Press.

[16] College Board 1997. "Trends in Student AID: 1986-1996," Washington, D.C.: The College Board.

[17] Comay, Y, M. Melnik and R. Pollatshek, 1972. "The Option Value of Education and the Optimal Path For Investment in Human Capital", International Economic Review; 14(2), 421-435.

[18] Domencich, T. and D. McFadden, 1975. Urban Travel Demand. North Holland, Amsterdam.

[19] Duncan, Greg, Brooks-Gunn, Jeanne, Yeung, W. Jean and Smith, Judith, 1998. "How Much Does Childhood Poverty A¤ect the Life Chances of Children?", American Sociological Review, Vol. 63, pp. 406-423.

[20] Eckstein, and Wolpin, 1989, "The Speci...cation and Estimation of Dynamic Stochastic Discrete Choice Models," Journal of Human Resources, 24:562-598.

[21] Hansen, L., 1989. "Impact of Student Financial Aid on Access." In The Crisis in Higher Education. Joseph Froomkin, ed. New York: Academy of Political Science.

[22] Hauser, R., 1993. "Trends in College Attendance among Blacks, Whites, and Hispanics." In Studies of Supply and Demand in Higher Education. C. Clotfelter and M. Rothschild, eds. University of Chicago Press.

[23] _____ 1991. "The Decline in College Entry Among African Americans: Findings in Search of Explanations." Manuscript, March.

[24] _____ and D. Featherman, 1976. "Equality of Schooling: Trends and Prospects." Sociology of Education 49, 99-120.

[25] Heckman, J., 1995. "Lessons From the Bell Curve," Journal of Political Economy, 103(5), 1091-1120.

[26] _____ 1981. "Statistical Models For Discrete Panel Data." In Structural Analysis of Discrete Data, C. Manski and D. McFadden, eds. MIT Press.

[27] Heckman, J., J., Smith, and Wittekind, M., 1997. "Awareness of Social Programs," unpublished manuscript, University of Chicago.

[28] Heckman, J. and B. Singer, 1984. "A Method of Minimizing The Impact of Distributional Assumption in Econometric Models For Duration Data." Econometrica, 52(2), 271-320.

[29] Heckman, J. and C. Taber, 1994. "Identi...cation in Binary Stochastic Process Models", unpublished manuscript, University of Chicago.

[30] Heckman, J. and J. Walker, 1990. "The Relationship Between Wages and Income And the Timing and Spacing of Births." Econometrica, 58(6), 1411-1441.

[31] Kane, 1998, "Public Subsidies For Higher Education: Overview and Outline for Reform," in Financing College Tuition: Government Policies and Educational Priorities, M. Kosters, ed., American Enterprise Institute, 51-53.

[32] Kane, T., 1997. "College Entry By Blacks since 1970: The Role of College Costs, Family Background, and the Returns to Education." Journal of Political Economy 102:878-911.

[33] _____ 1995. "Rising Public College Tuition and College Entry: How Well Do Public Subsidies Promote Access to College?" NBER Working Paper 5164, April, 1995

[34] _____ and C. Rouse, 1995. "The Labor Market Returns To Two and Four Year College: Is A Credit A Credit and Do Degrees Matter? " American Economic Review.

[35] Katz, L., and Murphy, K., 1992. "Changes in Relative Wages, 1963-1987: Supply and Demand Factors," Quarterly Journal of Economics; 107(1): 35-78.

[36] _____ Kosters, Marvin, 1998. Financing College Tuition: Government Policies and Educational Priorities, American Enterprise Institute, Washington, D.C.

[37] Krueger, A., 1997. "What's Up With Wages?," November, Goldman Sachs U.S. Economic Research, 1-18.

[38] Lazear, E.,1977. "Education: Consumption vs. Production?," Journal of Political Economy, 85(3), 569-598.

[39] Leslie and Brinkman, 1986. The Economic Value of Higher Education, MacMillan Publishing Company, New York.

[40] Lindsey, B., 1995. Mixture Models: Theory, Geometry and Applications, Institute of Mathematical Statistics, Alexandria, Va.

[41] McFadden, D. 1974. "Conditional Logit Analysis of Qualitative Choice Behavior," P. Zarembka, ed., Frontiers In Econometrics, New York: Academic Press.

[42] McPherson, M., and Schapiro, M.O., 1991. "Does Student Aid A¤ect College Enrollment? New Evidence on a Persistent Controversy," American Economic Review, 81(1), 309-18.

[43] Mare, R.D., 1980. "Social Background and School Continuation Decisions." Journal of The American Statistical Association 75 (June), 295-305.

[44] Mortenson, T., 1988. "Pell Grant Program Changes and Their E¤ects on Applicant Eligibility," American College Testing Program Research Report, No. 88-2.

[45] National Center for Education Statistics (NCES), 1997. The 1997 Digest of Education Statistics.

[46] Or…eld, G., 1992. "Money, Equity, and College Access." Harvard Educational Review 62(Fall): 337-72.

[47] Rao, C. R., 1965. "On Discrete Distributions Arising Out of Methods of Ascertainment," in Classical and Contagious Discrete Distributions, ed. by G. P. Patil. Calcutta: Statistical Publication Society, 320-333

[48] Read, T. and N. Cressie, 1988. Goodness of Fit Statistics For Discrete Multivariate Data. Springer, Berlin.

[49] Reich, R.,1991. The Work of Nations: Preparing Ourselves for 21st-Century Capitalism. New York: A. A. Knopf.

[50] Sewell, W. and R. Hauser, 1975. Education, Occupation and Earnings: Achievement In The Early Career, New York, Academic Press.

[51] Shea, J., 1996. "Does Parent's Money Matter?," University of Maryland Working Paper.

[52] St. John, Edward, P., 1993. "Untangling the Web: Using the Price-Response Measures in Enrollment Projections," Journal of Higher Education, 64(6), 675-695.

[53] Taubman, P., 1977. Kinometrics: Determinants of Socioeconomic Success Within and Between Families. North Holland.

[54] _____1989. "Role of Parental Income in Educational Attainment." A.E.R. Papers and Proceedings 79 (May):57-61.

[55] U. S. Department of Education, 1996. The Condition of Education, 1996, Washington, U.S. Government Printing O¢ce.

[56] White, H., 1982. "The Econometrics of Misspeci...ed Maximum Likelihood Models." Econometrica 50.

[57] Willis, R. and S. Rosen, 1979. "Education and Self-Selection," Journal of Political Economy 87 (October):S9-S40.

Figure 1. High School Completion and College Entry among 21 to 24 Year Old Men

A. High School Completion (Includes Equivalency Degrees)

Race – Ethnicity: ——— White ——— Black --------- Hispanic
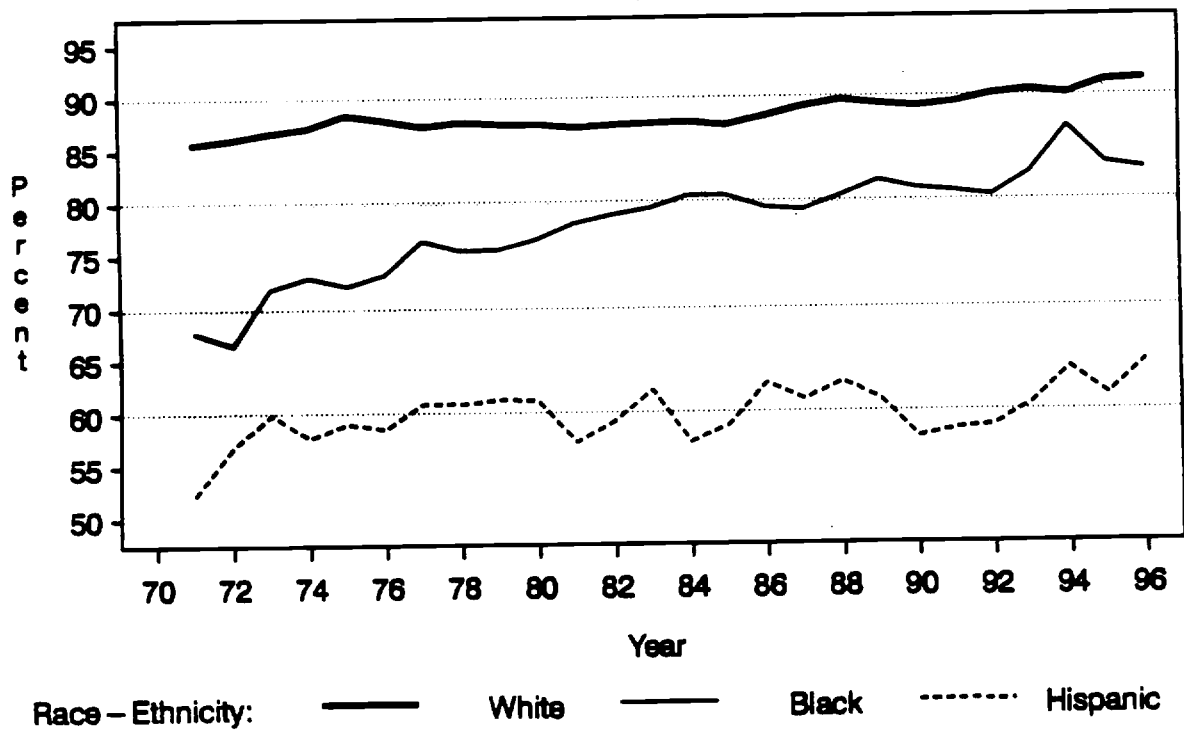
B. College Entry
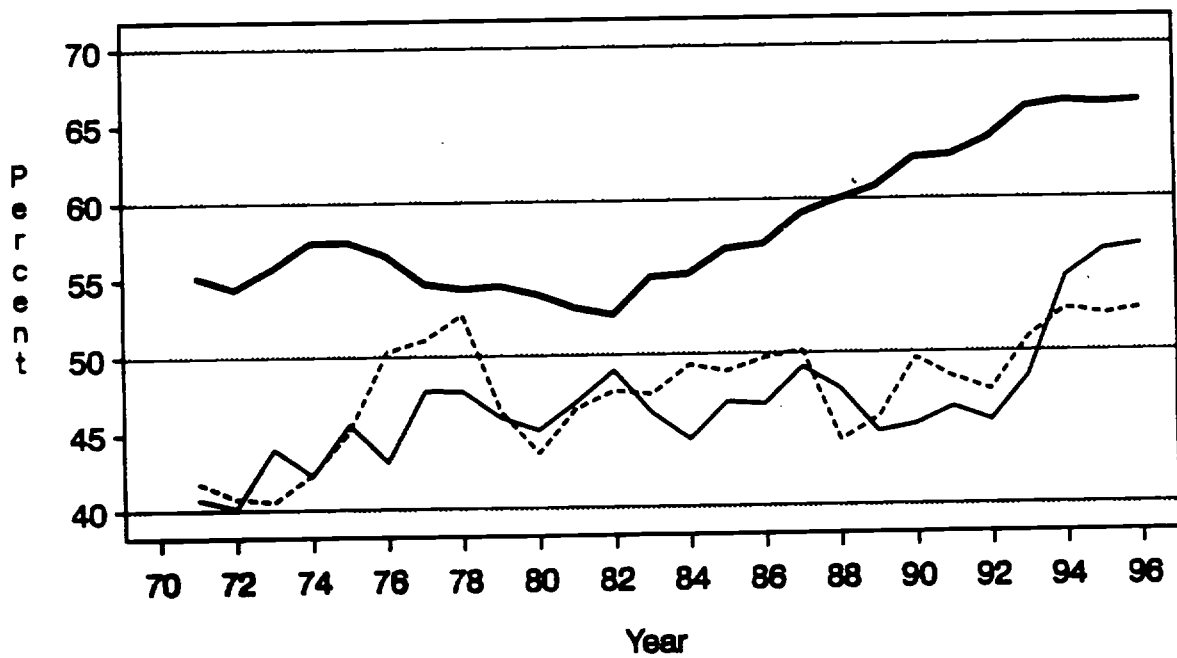
Race — Ethnicity: ——— White ——— Black ········· Hispanic

Notes. — These series are three — year moving averages of the raw data. Racial — ethnic groups are defined mutually exclusively.

Source. — Authors' calculations from 1971 to 1996 March CPS data files.

Figure 2. College Participation by 18 to 24 Year Old High School Completers by Family Income Quartile
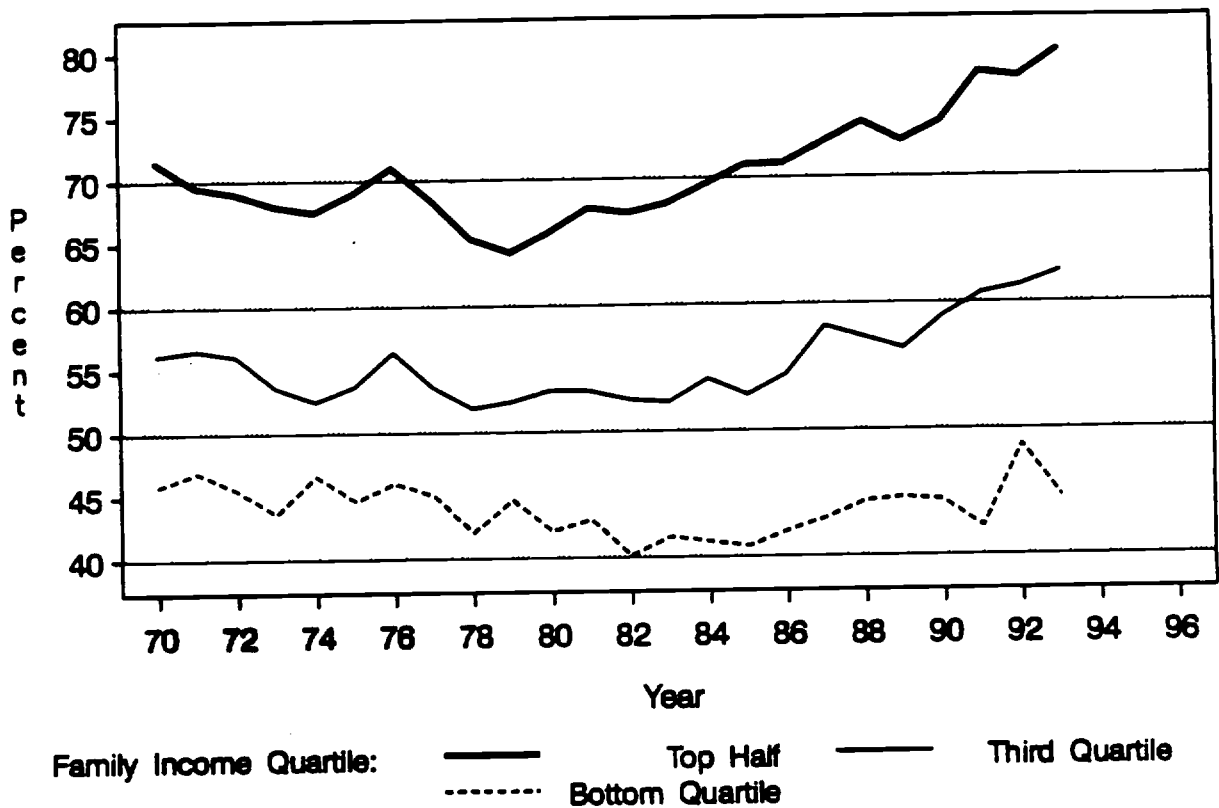
Source.—The 1971 to 1989 numbers were calculated from CPS P—20 school reports, and the 1990—1993 figures represent authors' calculations from October CPS data files.

Note.—High School completion includes equivalency degrees

Figure 3A. Highest Grade Completed at Age 15

Percent

9.8	28.4	52.3	9.5	13.6	30.0	47.0	9.3	3.6	19.0	68.5	8.9

7 — 8 9 10 +   7 — 8 9 10 +   7 — 8 9 10 +

├—— Black ——┤   ├— Hispanic —┤   ├—— White ——┤
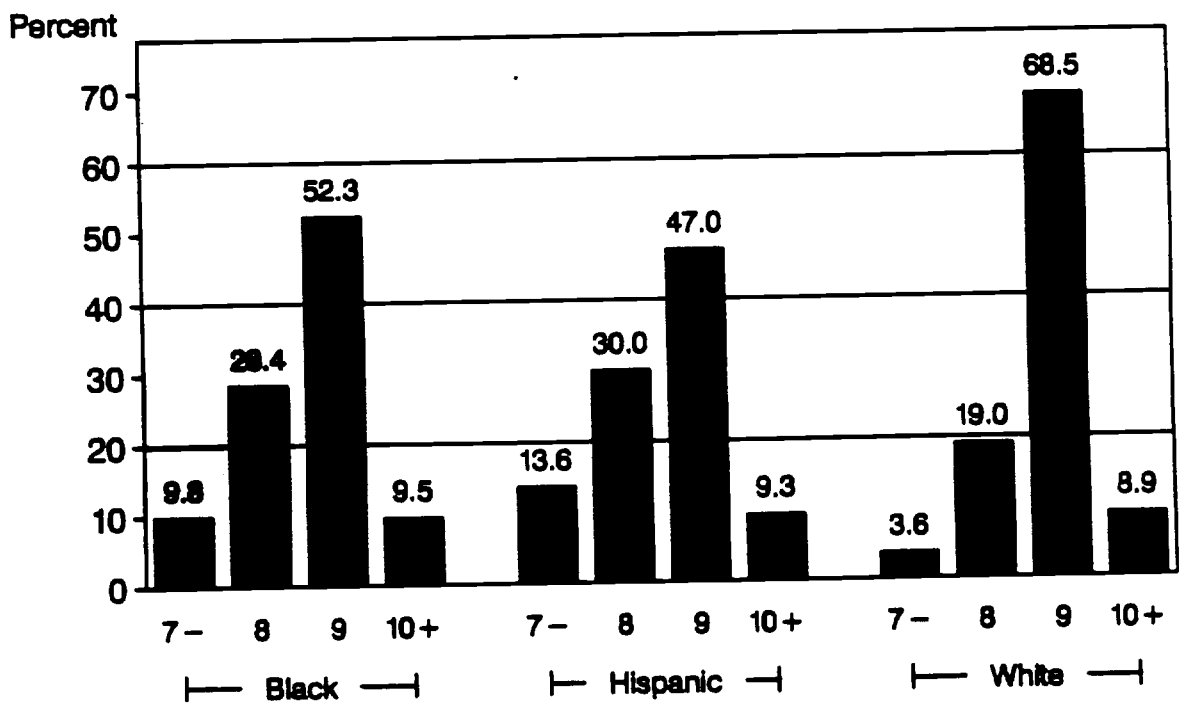
DEFINITIONS: "7 —" denotes grade 7 or less, and "10 +" denotes grade 10 or more.

Figure 3B. Highest Grade Completed at Age 24

Percent
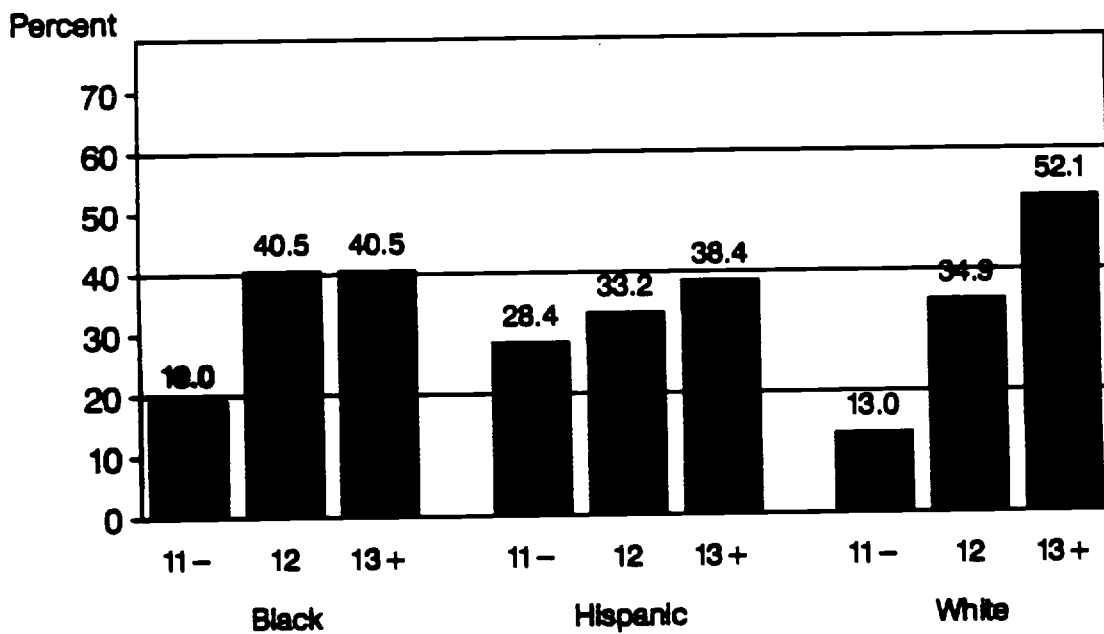
DEFINITIONS: "11—" denotes grade 11 or less, and "13+" denotes
college attendance.

Source. National Longitudinal Survey of Youth.

Figure 4. Schooling Attainment by Age

A. High School Graduation by Age (excludes equivalency degree receipt)

Race – Ethnicity: —— Black —— Hispanic - - - - - - White
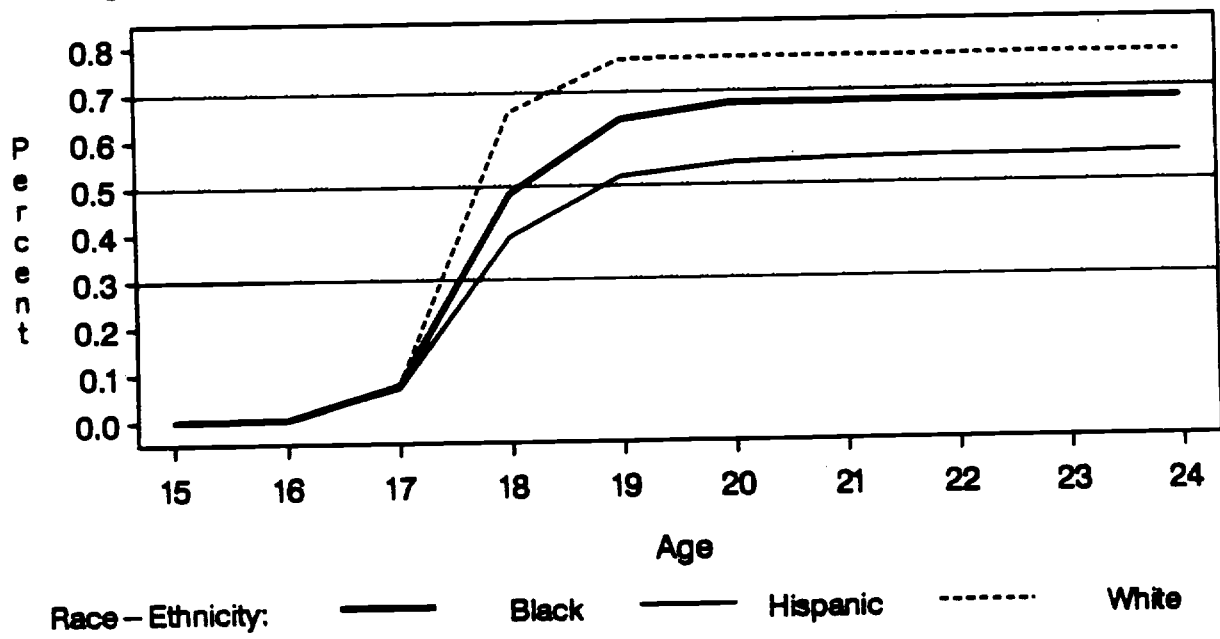
B. Equivalency Degree Attainment by Age

Race: ————— Black   ——— Hispanic   ·········· White

C. College Entry by Age
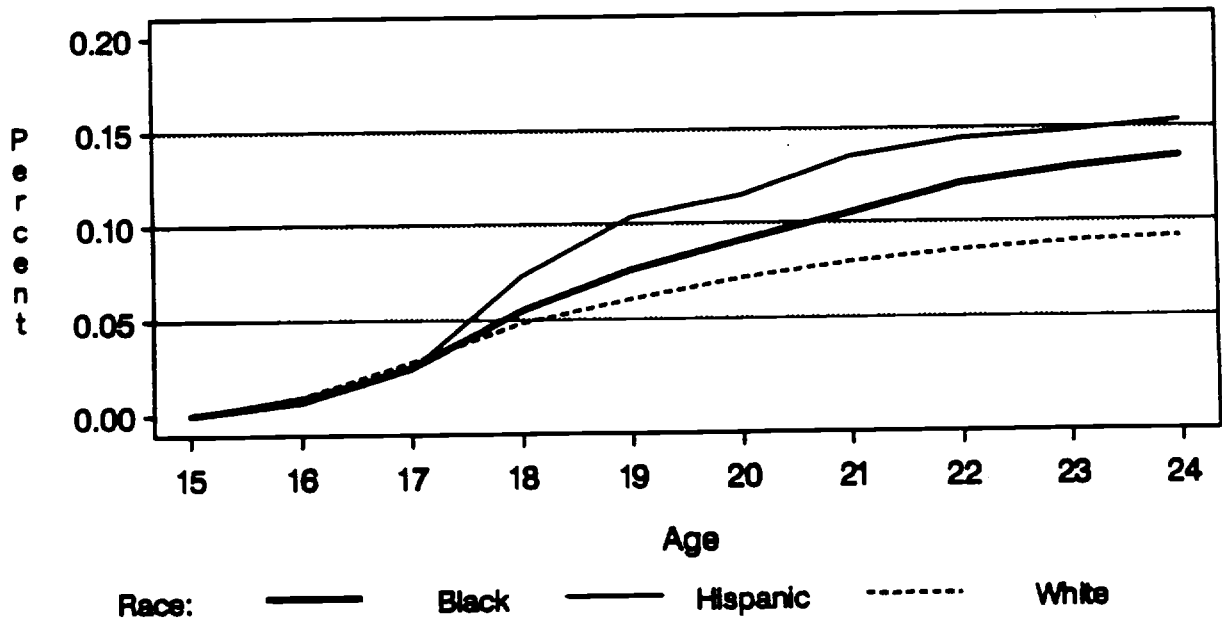
Race – Ethnicity: ▬▬▬ Black ▬▬▬ Hispanic - - - - - - - White
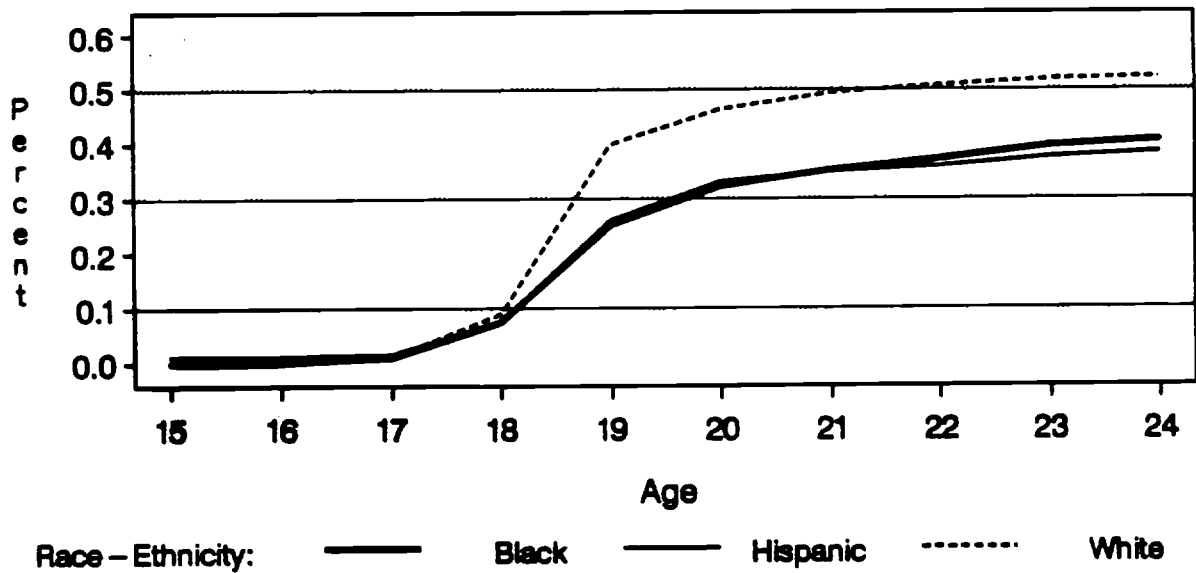
Note. – The college – entry numbers are not conditioned on high school completion.

Source. – The National Longitudinal Survey of Youth.

Table 1

High School Completion Status at Age 24
(standard errors of the means)

Blacks		Hispanics		Whites	
Total Completion Rate (1)	GED as % of Total (2)	Total Completion Rate (3)	GED as % of Total (4)	Total Completion Rate (5)	GED as % of Total (6)
.81 (.02)	17%	.72 (.02)	23%	.87 (.01)	11%

Table 2

College Entry by Type of College First Attended for
High School Graduates and GED Recipients at Age 24
(standard errors of the means)

A. High School Graduates								
Blacks			Hispanics			Whites		
4-Year (1)	2-Year (2)	Any (3)	4-Year (4)	2-Year (5)	Any (6)	4-Year (7)	2-Year (8)	Any (9)
(1).33(.02)	.23(.02)	.56(.02)	.29(.03)	.31(.03)	.60(.03)	.38(.01)	.25(.01)	.63(.01)
B. GED Recipients								
(2) .08(.02)	.15(.04)	.23(.03)	.10(.03)	.17(.03)	.27(.03)	.12(.03)	.22(.03)	.34(.03)

TABLE 3

Age 15 Grade Level and Schooling Outcomes at Age 24

| | Schooling Level at Age 24 | | | | |
Grade at Age 15	Dropout (1)	GED Only (2)	High School Only (3)	College Attendance* (4)	Age 15 Grade Distribution (5)
A. Blacks					
7 or Less	.60	.16	.17	.07	10%
8	.26	.18	.41	.15	28%
9	.09	.09	.45	.37	52%
10	.05	.14	.39	.42	10%
					100.0%
B. Hispanics					
7 or Less	.67	.19	.13	.01	14%
8	.34	.18	.32	.16	30%
9	.16	.11	.38	.33	47%
10	.12	.06	.32	.50	9%
					100.0%
C. Whites					
7 or Less	.80	.10	.08	.02	4%
8	.26	.15	.37	.22	19%
9	.07	.05	.39	.49	69%
10	.03	.10	.34	.53	9%
					100.0%

*College entrant may be a high school graduate or GED holder.

Note.—Columns (1)-(4) in each row sum to 1 and represent the distribution of age 24 schooling outcomes given the initial grade level specified in the right column.

Table 4

Percentage of Individuals who Ever Return to Regular
Secondary School Attendance After Having Dropped Out

	A. High School Graduates		
	Blacks (1)	Hispanics (2)	Whites (3)
Ever Dropout and Return	3%	6%	2%
	B. GED Recipients		
Ever Dropout and Return	42%	36%	39%
	C. Dropouts		
Ever Dropout and Return	12%	8%	6%

Note. A dropout is defined as an individual who leaves school for a least eight consecutive weeks during
the regular school year. A return is defined as at least eight consecutive weeks of attendance following a·
spell of nonattendance.

Table 5

Delay in College Entry: Distribution of Years Between High School
Completion and College Entry

A. High School Graduates			
Years Since Graduation	Black	Hispanic	White
Less than 1	73%	82%	82%
1-2	13%	9%	9%
2-3	6%	4%	3%
More than 3	8%	5%	7%
TOTAL	100%	100%	100%
B. Cell Recipients			
Less than 1	48%	53%	45%
1-2	22%	21%	20%
2-3	9%	9%	13%
More than 3	21%	17%	22%
TOTAL	100%	100%	100%

Note. Individuals completing high school after age 21 are excluded. GED recipients comprise 7%, 11%, and 6% of Black, Hispanic, and White college entrants respectively.

TABLE 6

College Entry by Age 24 for High School Completers (Graduates and GED Completers)
by Type of College First Attended and Quartile of the White Family Income Distribution
(standard errors of the mean in parentheses)

	4-Year (1)	2-Year (2)	Any (3)	% of Total Group (4)
A. Overall College Entry Proportions				
(1) Blacks	.29 (.02)	.21 (.02)	.50 (.02)	100%
(2) Hispanics	.25 (.02)	.28 (.02)	.53 (.03)	100%
(3) Whites	.35 (.01)	.26 (.01)	.61 (.01)	100%
B. Entry for those in Top 25% of White Family Income Distribution				
(4) Blacks	.47 (.08)	.30 (.07)	.77 (.07)	5%
(5) Hispanics	.39 (.08)	.18 (.06)	.58 (.08)	10%
(6) Whites	.50 (.03)	.27 (.02)	.77 (.02)	25%
C. Second 25% of White Family Income Distribution				
(7) Blacks	.43 (.07)	.17 (.05)	.60 (.05)	9%
(8) Hispanics	.37 (.07)	.35 (.07)	.72 (.06)	14%
(9) Whites	.36 (.03)	.26 (.02)	.62 (.03)	25%
D. Third 25% of White Family Income Distribution				
(10) Blacks	.30 (.04)	.22 (.04)	.51 (.05)	17%
(11) Hispanics	.27 (.04)	.28 (.05)	.55 (.05)	22%
(12) Whites	.28 (.03)	.25 (.03)	.53 (.03)	25%
E. Bottom 25% of White Family Income Distribution				
(13) Blacks	.25 (.02)	.22 (.02)	.47 (.02)	70%
(14) Hispanics	.19 (.03)	.30 (.02)	.49 (.03)	54%
(15) Whites	.27 (.03)	.25 (.03)	.52 (.03)	25%
F. Predicted Attendance if Minority Family Income Distributions were Equal to the White Distribution*				
(16) Blacks	.36 (.04)	.23 (.03)	.59 (.04)	100%
(17) Hispanics	.30 (.05)	.28 (.04)	.58 (.05)	100%

*These predictions represent the total proportion of minorities who would attend college if minorities had the same family income distribution as Whites and the minority-specific college entrance probabilities given in panels B through E.

Source: National Longitudinal Survey of Youth

TABLE 7

Definitions of Variables Used in The Schooling-Transition Analysis

Number of Siblings	Number of living siblings.
Family Income	A two-year average of total family income measured at consecutive years between age 14 and 18. Includes all components of income (salary, public assistance, and so forth). In 1994 dollars.
Highest Grade Father	Highest grade completed in years by father when respondent was 14.
Highest Grade Mother	Highest grade completed in years by mother when respondent was 14.
Broken Home	Absence of one or both parents from the respondent's household at age 14.
Southern Residence at Age 14	Whether the respondent lived in the Southern census region at age 14.
Urban Residence at Age 14	Whether the respondent lived on an urban area at age 14.
County Average Annual Earnings	The average annual earnings per job in the county of residence as measured by earnings in low-skill industries (retail and wholesale trade, services, and so forth). Measured annually at all ages.
County Unemployment Rate	Annual unemployment rate in country of residence. Measured annually at all ages.
College Tuition	Two-year or four-year public college tuition for in-state students in the individual's county (if available) or state of residence. In 1994 dollars. Measured annually at all ages. In 1994 dollars.
Pell Grant Award	Imputed amount of Pell grant award for both two- and four-year public institutions in the county or state. In 1994 dollars.
Pell Grant Award	Imputed amount of Pell grant award for both two- and four-year public institutions in the county or state. In 1994 dollars.
AFQT Score	Score on the Armed Forces Qualification Test.
College Proximity	Whether any public two- or four-year school exists in the county of residence. Measured at all ages and broken out by two- or four-year type for some analyses

Table 8

Mean Characteristics by Race Group
(Standard errors of the mean in parentheses)

	Blacks (1)	Hispanics (2)	Whites (3)
Number of Siblings[*]	4.7 (0.09)	4.5 (0.12)	2.9 (0.05)
Highest Grade Father	9.3 (0.09)	7.9 (0.17)	12.2 (0.08)
Highest Grade Mother	10.7 (0.08)	7.8 (0.16)	11.9 (0.06)
Family Income[**]/1000	25.4 (0.05)	30.3 (0.07)	44.8 (0.06)
Broken Home	0.43 (0.02)	0.27 (0.02)	0.13 (0.01)
Urban Res. Age 14	0.82 (0.01)	0.93 (0.01)	0.73 (0.01)
Southern Res. Age 14	0.57 (0.02)	0.26 (0.02)	0.28 (0.01)
County Annual Average Wage[**]/1000 at Age 18	21.0 (0.12)	22.3 (0.15)	20.1 (0.32)
County Unemployment Rate at Age 18	6.0 (0.12)	6.9 (0.16)	6.3 (0.12)
Four-Year College State Average Tuition** at Age 18	1705 (17.3)	1528 (14.3)	1850 (13.6)
Two-Year College State Average Tuition** at Age 18	798 (10.6)	580 (17.0)	888 (8.9)
Two-Year College County Average Tuition** at Age 18	780 (12.5)	572 (21.4)	833 (11.2)
Four-Year College Pell Grant**	1260 (27.5)	1315 (22.8)	548 (12.1)
College Proximity	0.89 (.10)	0.92 (.01)	0.82 (.01)
AFQT Score	46.4 (1.1)	54.4 (1.9)	71.8 (.62)

Table continue on next page

‘The large number of siblings is a consequence of size-biased sampling in NLSY. See the NLSY Handbook for a description of the sampling design.

‘‘Reported in 1994 dollars. Definitions are given in Table 7. Tuition figures are for in-state students at public institutions.

TABLE 9
Can Covariate Differences Explain Racial-Ethnic Schooling Gaps?
Simulated Differences in Predicted White-Minority
Schooling Attainment Probabilities

A. Complete Grade 9 or More by Age 15 (Initial Condition)	White-Black Gap (1)	White-Hispanic Gap (2)
(1) Actual White-Minority Gap	.16 (.02)†	.21 (.02)†
Exclude AFQT		
(2) Predicted Gap when Minorities have White Covariates*	-.02 (.03)	.03 (.06)
(3) Predicted Gap when Whites have Minority Covariates**	-.05 (.04)	.02 (.07)
Include AFQT		
(4) Predicted Gap when Minorities have White Covariates*	-.05 (.03)†	-.01 (.05)
(5) Predicted Gap when Whites have Minority Covariates**	-.08 (.03)†	-.01 (.07)
B. High School Completion Gap (Includes GED Attainment)		
(6) Actual White-Minority Gap	.06 (.01)†	.14 (.02)†
Exclude AFQT		
(7) Predicted Gap when Minorities have White Covariates*	-.04 (.04)	.02 (.05)
(8) Predicted Gap when Whites have Minority Covariates**	-.07 (.04)†	-.00 (.05) .
Include AFQT		
(9) Predicted Gap when Minorities have White Covariates*	-.07 (.03)†	-.02 (.04)
(10) Predicted Gap when Whites have Minority Covariates**	-.12 (.03)†	-.09 (.04)†
C. College Entry Probabilities Given High School Completion		
(11) Actual White-Minority Gap	.11 (.02)†	⁻.07 (.02)†
Exclude AFQT		
(12) Predicted Gap when Minorities have White Covariates*	-.00 (.02)	-.04 (.04)†
(13) Predicted Gap when Whites have Minority Covariates**	-.05 (.03)†	-.11 (.04)†
Include AFQT		
(14) Predicted Gap when Minorities have White Covariates*	-.08 (.02)†	-.06 (.03)†
(15) Predicted Gap when Whites have Minority Covariates**	-.12 (.02)†	-.15 (.04)†
D. Population College Entry Gap (Unconditional on High School Completion)		
(16) Actual White-Minority Gap	.12 (.02)†	.14 (.02)†
Exclude AFQT		
(17) Predicted Gap when Minorities have White Covariates*	-.04 (.04)	-.04 (.05)
(18) Predicted Gap when Whites have Minority Covariates**	-.10 (.023)†	-.16 (.04)†
Include AFQT		
(19) Predicted Gap when Minorities have White Covariates*	-.04 (.03)	-.04 (.04)
(20) Predicted Gap when Whites have Minority Covariates**	-.10 (.03)†	-.16 (.04)†

† Significant at the 10 percent level.

* These numbers represent the predicted gap in schooling attainment (the white probability minus the minority probability) when minority schooling attainment is predicted using the White covariate distribution. More precisely, let

the probability of high school graduation, for instance, be represented by $Pr(X_i;\beta_j)$,where i and j denote Black,

Hispanic, or White covariates and parameter estimates. For Blacks, the numbers in rows (7) and (9) of column (10)

correspond to $Pr(X_w;\beta_w) - Pr(X_w;\beta_b)$ where the subscripts b and w denote Black and White respectively. The

first term in the difference corresponds to predicted White schooling and the second term shows predicted Black
schooling if Blacks had the White covariate distribution. A negative entry in the table indicates that the minority

probability is higher than the white probability.

** These numbers show the predicted gap in schooling attainment when Whites are given either the Black (or Hispanic) covariate endowments: $Pr(X_w; \beta_w) - Pr(X_b; \beta_w)$.

Notes: Aggregate predicted probabilities do not differ from their sample counterparts in the first two significant digits. Standard errors were calculated using 500 draws from the relevant estimated parameter distributions.

Table 10

Can Covariate Differences Explain Racial-Ethnic College Attendance Gaps?
Same Specification as the Model of Table 10 with AFQT Except College Attendance
is Disaggregated into Two-Year and Four-Year Attendance
(The specification of the secondary-school transitions is the same as in Table 10)

A. College Entry Probabilities given High School Completion		
	Whites-Blacks Gap	Whites-Hispanics Gap
(1) Actual Gap	.11	.07
(2) Predicted Gap when Minorities have White Covariates	-.01	-.01
(3) Predicted Gap when Whites have Minority Covariates	-.06	-.09

B. Population College Entry Gap (Unconditional on High School Completion)		
(1) Actual Gap	.13	.03
(2) Predicted Gap when Minorities have White Covariates	-.03	.-02
(3) Predicted Gap when Whites have Minority Covariates	-.08	-.12

C. College Entry Probabilities given High School Completion		
(1) Actual Gap	.11	.07
(2) Predicted Gap when Minorities have White Covariates	-.06	-.05
(3) Predicted Gap when Whites have Minority Covariates	-.14	-.12

D. Population College Entry Gap (Unconditional on High School Completion)		
(1) Actual Gap	.13	.13
(2) Predicted Gap when Minorities have White Covariates	-.12	-.08
(3) Predicted Gap when Whites have Minority Covariates	-.16	.-15

Note: See the notes to Table 9

TABLE 11
The Change in the Predicted White-Minority Schooling Gap when Minority Explanatory Variables are Equated to White Levels
(Standard Errors in Parentheses)
The Actual White-Minority Schooling Gap is Given in the Last Row of Each Panel

A. Change in minority probability of being in grade 9 or higher at age 15

	Without AFQT Score		With AFQT Score	
	Blacks (1)	Hispanics (2)	Blacks (3)	Hispanics (4)
(1) Equating All Family Background Components	.07 (.021)†	.08 (.022)†	.03 (.022)	.02 (.021)
(1a) Number of Siblings	.03 (.009)†	.03 (.012)†	.02 (.010)†	.01 (.013)
(1b) Highest Grade of Father	.04 (.021)†	-.01 (.028)	.01 (.022)	-.03 (.029)
(1c) Highest Grade of Mother	.01 (.005)†	.06 (.021)†	.00 (.007)	.04 (.023)†
(1d) Broken Home	-.02 (.010)	-.003 (.006)	.001 (.010)	.003 (.007)
(2) Equating Family Income	.09 (.021)†	.12 (.021)†	.08 (.027)†	.04 (.020)†
(3) Equating Local Average Wages	.007 (.004)†	-.01 (.008)	.005 (.005)	-.02 (.011)†
(4) Equating Tuition and College Proximity	-.01 (.005)†	-.02 (.008)†	-.01 (.006)†	-.02 (.009)†
(5) Equating AFQT Scores	na	na	.16 (.034)†	.17 (.027)†
(6) Equating 1 and 2	.14 (.023)†	.18 (.030)†	.10 (.025)†	.06 (.028)†
(7) Equating 1, 2, 3, and 4	.13 (.021)†	.17 (.021)†	.09 (.023)†	.04 (.027)†
(8) Equating 1, 2, 3, 4, and 5	na	na	.21 (.027)†	.20 (.026)†
(9) Actual Gap between Whites and minorities	.16	.21	.16	.21

B. Change in minority probability of high school completion at age 24
(high school graduation and GED attainment combined)

	Without AFQT Score		With AFQT Score	
	Blacks (1)	Hispanics (2)	Blacks (3)	Hispanics (4)
(1) Equating All Family Background Components	.06 (.014)†	.05 (.021)†	.03 (.015)†	-.01 (.024)
(1a) Number of Siblings	.02 (.007)†	.03 (.010)†	.01 (.007)†	.01 (.010)
(1b) Highest Grade of Father	.03 (.015)†	-.01 (.032)	.02 (.016)	-.04 (.029)†
(1c) Highest Grade of Mother	.01 (.005)	.03 (.020)	-.002 (.005)	.02 (.022)
(1d) Broken Home	.01 (.008)	-.005 (.006)	.007 (.007)	-.001 (.005)
(2) Equating Family Income	.07 (.016)†	.08 (.018)†	.05 (.018)†	.01 (.019)
(3) Equating Local Average Wages	.01 (.004)†	.001 (.008)	.008 (.004)†	-.01 (.010)
(4) Equating Tuition and College Proximity	-.006 (.003)†	-.001 (.008)	-.007 (.003)†	.009 (.008)
(5) Equating AFQT Scores	na	na	.11 (.019)†	.16 (.019)†
(6) Equating 1 and 2	.12 (.013)	.12 (.021)†	.06 (.016)†	.002 (.026)
(7) Equating 1, 2, 3, and 4	.12 (.013)	.13 (.023)†	.06 (.017)†	.008 (.030)
(8) Equating 1, 2, 3, 4, and 5	na	na	.15 (.014)†	.16 (.026)†
(9) Actual Gap between Whites and Minorities	.06	.14	.06	.14

TABLE 11 (continued)

C. Change in minority college entry probabilities at age 24 conditional on high school completion

	Without AFQT Score		With AFQT Score	
	Blacks (1)	Hispanics (2)	Blacks (3)	Hispanics (4)
(1) Equating All Family Background Components	.10 (.027)†	.11 (.026)†	.08 (.021)†	.05 (.014)†
(1a) Number of Siblings	.03 (.012)†	.03 (.012)†	.02 (.011)†	.01 (.009)
(1b) Highest Grade of Father	.08 (.028)†	.03 (.032)	.06 (.020)†	.02 (.029)
(1c) Highest Grade of Mother	.003 (.008)	.05 (.025)†	-.005 (.006)	.02 (.022)
(1d) Broken Home	-.01 (.011)	.01 (.009)	-.002 (.009)	.01 (.008)†
(2) Equating Family Income	.05 (.023)†	.03 (.013)†	.004 (.020)	-.02 (.019)
(3) Equating Local Average Wages	.004 (.006)	.04 (.013)†	.002 (.003)	.03 (.010)†
(4) Equating Tuition and College Proximity	-.03 (.006)†	-.05 (.016)†	-.02 (.004)†	-.05 (.016)†
(5) Equating AFQT Scores	na	na	.15 (.028)†	.12 (.022)†
(6) Equating 1 and 2	.14 (.027)†	.13 (.023)†	.08 (.027)	.03 (.023)
(7) Equating 1, 2, 3, and 4	.12 (.029)†	.12 (.030)†	.06 (.028)	.01 (.030)
(8) Equating 1, 2, 3, 4, and 5	na	na	.20 (.029)†	.13 (.030)†
(9) Actual Gap between Whites and Minorities	.11	.07	.11	.07

D. Change in minority college entry probabilities at age 24 (unconditional on high school graduation)

	Without AFQT Score		With AFQT Score	
	Blacks (1)	Hispanics (2)	Blacks (3)	Hispanics (4)
(1) Equating All Family Background Components	.13 (.027)†	.11 (.026)†	.07 (.022)†	.03 (.026)
(1a) Number of Siblings	.04 (.011)†	.04 (.013)†	.02 (.010)†	.01 (.012)
(1b) Highest Grade of Father	.09 (.026)†	.017 (.021)	.06 (.020)†	-.01 (.026)
(1c) Highest Grade of Mother	.005 (.008)	.055 (.028)†	-.005 (.006)	.02 (.023)
(1d) Broken Home	-.004 (.011)	-.009 (.007)	.003 (.010)	.01 (.007)
(2) Equating Family Income	.08 (.023)†	.06 (.210)†	.03 (.022)	-.007 (.018)
(3) Equating Local Average Wages	.01 (.005)†	.03 (.013)†	.003 (.005)	.02 (.013)
(4) Equating Tuition and College Proximity	-.03 (.006	-.03 (.015)†	-.02 (.005)†	-.04 (.013)†
(5) Equating AFQT Scores	na	na	.18 (.029)†	.19 (.023)†
(6) Equating 1 and 2	.19 (.027)†	.17 (.030)†	.10 (.028)†	.03 (.028)†
(7) Equating 1, 2, 3, and 4	.16 (.029)†	.17 (.032)†	.09 (.029)†	.01 (.030)
(8) Equating 1, 2, 3, 4, and 5	na	na	.27 (.033)†	.21 (.035)†
(9) Actual Gap between Whites and Minorities	.12	.14	.12	.14

† Significant at the 10 percent level.

Notes: The simulations were calculated by shifting the mean of the explanatory variables for minorities to the mean of the variables for Whites. For example, the number in row (2) of panel D represents the rise in college entry if all Black families were given an income transfer equal to the mean difference in Black and White family incomes. The standard errors shown above were calculated using 500 draws from the estimated parameter distributions.

Table 12
Estimated Average Derivatives of Age 24 Schooling Attainment Probabilities

A. Complete Grade 9 or Higher by Age 15			
	Blacks	Hispanics	Whites
AVERAGE DERIVATIVES:			
Family Income (10,000's of $)	.052 (.013)†	.081 (.016)†	.031 (.006)†
Number of siblings	-.018 (.005)†	-.021 (.008)†	-.013 (.004)†
Highest Grade Completed Father	.012 (.007)†	-.003 (.006)	.011 (.005)†
Highest Grade Completed Mother	.009 (.005)†	.017 (.005)†	.010 (.004)†
College Tuition ($1000's)	-.127 (.045)†	-.111 (.038)†	-.043 (.025)†
Local Average Wage ($1000's)	-.008 (.004)†	.006 (.005)	.001 (.001)†
10% INCREASE IN THE CHANCES OF:			
Broken Home	.004 (.003)	.002 (.004)	-.000 (.002)
College in County	.002 (.002)	-.002 (.003)	.005 (.003)†

B. Complete High School by Age 24 (includes GED attainment)			
Family Income (10,000's of $)	.041 (.010)†	.056 (.014)†	.020 (.004)†
Number of Siblings	-.011 (.004)†	-.020 (.007)†	-.005 (.004)†
Highest Grade Completed Father	.013 (.006)†	-.002 (.006)	.018 (.005)†
Highest Grade Completed Mother	.004 (.004)†	.008 (.005)†	.011 (.003)†
College Tuition ($1000's)	-.133 (.041)†	-.012 (.032)	.025 (.018)
Local Average Wage (1000's of $)	-.014 (.004)†	-.001 (.003)	-.008 (.004)†
10% INCREASE IN THE CHANCES OF:			
Broken Home	-.002 (.002)	.001(.003)	-.004 (.002)†
College in County	.004 (.004)	-.002 (.004)	-.001 (.001)

C. Enter College by Age 24 Conditional on High School Completion			
Family Income (10,000's of $)	.032 (.012)	.016 (.008) .	.018 (.006)†
Number of Siblings	-.020 (.007)	-.017 (.009)	-.024 (.007)†
Highest Grade Completed Father	.031 (.010)	.007 (.008)	.023 (.008)†
Highest Grade Completed Mother	.002 (.007)	.011 (.006)	.044 (.006)†
College Tuition ($1000's)	-.036 (.021)	-.064 (.032)	-.057 (.024)†
Local Average Wage (1000's of $)	-.006 (.004)	-.019 (.008)	-.011 (.055)†
10% INCREASE IN THE CHANCES OF:			
Broken Home	.004 (.004)	-.010 (.006)	.005 (.004)
College in County	.025 (.005)†	.035 (.007)	.023 (.002)†

D. Enter College by Age 24 (not conditioned on high school–population effects)			
Family Income (10,000's of $)	.048 (.013)†	.042 (.015)†	.027 (.006)†
Number of Siblings	-.022 (.006)†	-.023 (.008)†	-.024 (.006)†
Highest Grade Completed Father	.032 (.009)†	.004 (.007)	.031 (.007)†
Highest Grade Completed Mother	.003 (.007)	.013 (.006)†	.046 (.005)†
College Tuition ($1000's)	-.088 (.033)†	-.055 (.030)†	-.035 (.024)
Local Average Wage (1000's of $)	-.010 (.005)†	-.014 (.005)†	-.013 (.004)†
10% INCREASE IN THE CHANCES OF:			
Broken Home	.002 (.004)	-.007 (.005)	.003 (.004)
College in County	.022 (.004)†	.022 (.005)†	.020 (.003)†

+Significant at the 10 percent level.

Notes. The standard errors shown above were calculated using 500 draws from the estimated parameter distributions.

TABLE 13
Estimated Average Derivatives of Age 24 Schooling Attainment Probabilities
AFQT is Included as an Explanatory Variable in All Transitions

A. Complete Grade 9 or Higher by Age 15			
	Blacks	Hispanics	Whites
AVERAGE DERIVATIVES:			
Family Income (10,000's of $)	.041 (.015)†	.021 (.015)	.019 (.006)†
Number of Siblings	-.013 (.005)†	-.007 (.007)	-.009 (.004)†
Highest Grade Completed Father	.001 (.007)	-.010 (.007)	.002 (.005)
Highest Grade Completed Mother	.003 (.006)	.010 (.006)†	-.000 (.004)
Local Average Wage ($1000's)	-.020 (.019)	.050 (.024)†	.023 (.017)
College Tuition ($1000's)	-.132 (.052)†	-.114 (.038)†	-.038 (.029)
AFQT Score-10	.086 (.009)†	.101 (.010)†	.058 (.004)†
10% INCREASE IN THE CHANCES OF :			
Broken Home	.001 (.004)	-.002 (.004)	-.001 (.002)†
College in County	.003 (.004)	.002 (.004)	.005 (.003)

B. Complete High School by Age 24 (includes GED attainment)			
AVERAGE DERIVATIVES:			
Family Income ($10,000's)	.035 (.010)†	.008 (.013)	.010 (.005)†
Number of Siblings	-.007 (.004)†	-.005 (.006)	.001 (.004)
Highest Grade Completed Father	.006 (.006)	-.010 (.006)†	.006 (.004)†
Highest Grade Completed Mother	-.001 (.004)	.005 (.005)†	.006 (.004)†
Local Average Wage ($1000's)	-.039 (.017)†	.025 (.020)	-.011 (.011)
College Tuition ($1000's)	-.140 (.037)†	.024 (.018)	.020 (.017)
AFQT Score-10	.084 (.006)†	.084 (.008)†	.043 (.004)†
10% INCREASE IN THE CHANCES OF :			
Broken Home	-.002 (.002)	.001 (.003)	-.001 (.002)
College in County	.001 (.002)	-.003 (.003)	-.005 (.002)†

C. Enter College by Age 24 Conditional on High School Completion			
AVERAGE DERIVATIVES:			
Family Income ($10,000's)	.005 (.012)	-.012 (.013)	.008 (.006)
Number of Siblings	-.012 (.006)†	-.008 (.008)	-.019 (.006)†
Highest Grade Completed Father	.022 (.008)†	.002 (.007)	.013 (.006)†
Highest Grade Completed Mother	-.004 (.006)	.004 (.007)	.032 (.005)†
County Tuition ($1000's)	-.027 (.021)	-.074(.035)†	-.045 (.019)†
Local Average Wage ($1000's)	-.008 (.010)	-.043 (.022)†	-.017 (.012)
AFQT Score-10	.088 (.005)†	.080 (.013)†	.073 (.006)
10% INCREASE IN THE CHANCES OF :			
Broken Home	-.000 (.002)	-.010 (.006)†	.001 (.003)
College in County	.024 (.004)†	.032 (.009)†	.020 (.002)†

D. Enter College by Age 24 (not conditioned on high school--population effects)			
AVERAGE DERIVATIVES:			
Family Income ($10,000's)	.019 (.014)	-.005 (.013)	.010 (.006)†
Number of Siblings	-.012 (.005)†	-.008 (.008)	-.016 (.006)†
Highest Grade Completed Father	.021 (.007)†	-.002 (.007)	.014 (.006)†
Highest Grade Completed Mother	-.004 (.005)	.005 (.006)	.031 (.005)†
College Tuition ($1000's)	-.077 (.029)†	-.056 (.038)†	-.030 (.018)†
Local Average Wage ($1000's)	-.020 (.015)	-.020 (.024)	-.019 (.017)
AFQT Score-10	.112 (.009)†	.113 (.012)†	.094 (.006)†
10% INCREASE IN THE CHANCES OF :			
Broken Home	-.001 (.003)	-.007 (.004)†	.001 (.003)
College in County	.021 (.004)†	.019 (.008)†	.016 (.003)†

† Significant at the 10 percent level.
Notes. The standard errors shown above were calculated using 500 draws from the estimated parameter distributions.

Table 14

The Effect of a $10,000 Increase in Family Income on the Chances of Two-Year
and Four-Year College Entry by High School*
(standard errors of estimates in parentheses)

	Blacks (1)	Hispanics (2)	Whites (3)
A. Combined Two- and Four-Year College Effect (same as reported in panel C of Tables 12 and 13)			
(1) Excluding AFQT	.032 (.012)†	.016 (.008)†	.018 (.006)†
(2) Including AFQT	.005 (.012)	-.012 (.013)	.008 (.006)
B. Including AFQT and Disaggregating Two-Year and Four-Year Effects			
(3) Two-Year College	.008 (.008)	.005 (.005)	-.002 (.005)
(4) Four-Year College	.024 (.008)†	.011 (.004)†	.020 (.005)†
C. Including AFQT and Disaggregating Two-Year and Four-Year Effects			
(5) Two-Year College	-.002 (.008)	-.008 (.007)	-.003 (.005)
(6) Four-Year College	.007 (.008)	-.004 (.010)	.010 (.005)†

*High School graduates and GED attainers combined.
†Significant at the 10 percent level.

Table 15

Own-effects and Cross-effects of a $1000 rise in Two-Year Tuition, Four-Year Tuition, and Both on Enrollments at Two-Year and Four-Year Colleges For High School Completers[*]

	Blacks			Hispanics			Whites		
	Enrollment Change at 2-Year 4-Year Combined			Enrollment Change at 2-Year 4-Year Combined			Enrollment Change at 2-Year 4-Year Combined		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
A. Exclude Family Background, Exclude AFQT									
(1) 2-Year Tuition Rise	-.10† (.02)	.03† (.01)	-.07† (.02)	-.11† (.03)	.03† (.01)	-.08† (.03)	-.15† (.02)	.07 (.02)	-.09† (.02)
(2) 4-Year Tuition Rise	.01† (.01)	-.04† (.00)	-.03† (.01)	.01 (.02)	-.03 (.01)	-.02 (.02)	.01† (.01)	-.09† (.02)	-.08† (.02)
(3) Rise in Both	-.09† (.02)	-.01† (.01)	-.10† (.02)	-.10† (.02)	-.00 (.02)	-.10† (.03)	-.14† (.03)	-.02† (.01)	-.17† (.03)
B. Include Family Background, Exclude AFQT									
(1) 2-Year Tuition Rise	-.07† (.02)	.01 (.02)	-.06† (.02)	-.09† (.03)	.01 (.01)	-.08† (.03)	-.11† (.03)	.05† (.01)	-.06† (.02)
(2) 4-Year Tuition Rise	.01 (.01)	.01 (.01)	.02 (.01)	.01 (.01)	.01 (.01)	.02 (.01)	.02 (.01)	-.02 (.01)	-.00 (.01)
(3) Rise in Both	-.06† (.02)	.02† (.01)	-.04† (.02)	-.08† (.03)	.02 (.01)	-.06† (.03)	-.09† (.03)	.03† (.01)	-.06† (.03)
C. Include Family Background and AFQT									
(1) 2-Year Tuition Rise	-.08† (.02)	.02 (.02)	-.06† (.02)	-.08† (.03)	.01 (.01)	-.07† (.03)	-.10† (.02)	.05† (.01)	-.05† (.02)
(2) 4-Year Tuition Rise	.03 (.02)	-.00 (.02)	.03 (.02)	.01 (.01)	-.00 (.00)	.00 (.01)	-.01 (.01)	.01 (.01)	-.00 (.01)
(3) Rise in Both	-.05† (.02)	.02 (.02)	-.03 (.03)	-.07† (.03)	.00 (.01)	-.07† (.04)	-.11† (.02)	.06† (.02)	-.05† (.02)

Table continue on next page

†Significant at the 10 percent level.

Notes.--**High** school completers are composed of GED recipients and high school graduates combined. The numbers reported above in columns (3), (6), and (9) or row (3) in panels A and B correspond to the tuition effects reported in panel C or Tables 11 and 12.

Table 16

Effects of a $1,000 Increase in Gross Tuition (Both Two- and Four-Year) on the College Entry
Probabilities of High School Completers By Family Income Quartile and By AFQT Quartile

	Blacks (1)	Hispanics (2)	Whites (3)
A. Overall Gross*Tuition Effects			
(1) No explanatory variables except tuition in the model	-.10	-.10	-.17
(2) Baseline specification (see note at table base, includes family income and background, and so forth)	-.04	-.06	-.06
(3) Adding AFQT to the row (2) specification	-.03	-.06	-.05
B. By Family Income Quartiles (Panel A Row (2) Specification)			
(4) Top Quartile	-.01	-.04	-.04
(5) Second Quartile	-.03	-.05	-.06
(6) Third Quartile	-.07	-.08	-.07
(7) Bottom Quartile	-.05	-.08	-.06
(8) Joint Test of Equal Effects Across Quartiles (P-values)	.23	.66	.49
C. By Family Income Quartiles (Panel A Row (3) Specification			
(9) Top Quartile	-.02	-.02	-.02
(10)Second Quartile	.00	-.05	-.06
(11)Third Quartile	-.05	-.09	-.07
(12)Bottom Quartile	-.04	-.07	-.04
(13)Joint Test of Equal Effects Across Quartiles (P-values)	.45	.49	.34
D. By AFQT Quartiles (Panel A Row (3) Specification plus tuition-AFQT interaction terms)			
(14)Top Quartile	-.02	-.03	-.03
(15)Second Quartile	-.01	-.05	-.06
(16)Third Quartile	-.03	-.07	-.06
(17)Bottom Quartile	-.03	-.05	-.05
(18)Joint Test of Equal Effects Across Quartiles (P-values)	.84	.68	.60

*Gross tuition is the nominal sticker-price of college and excludes scholarship and loan support.
Notes.--These simulations assume both two-year and four-year college tuition increase by $1,000
for the population of high school completers. The baseline specification used in row (2) of panel A and
rows (4) through (7) of panel B includes controls for family background, family income, average income,
average wages in income interactions, estimated Pell grant award eligibility, and dummy variables, that
indicate the proximity of two- and four-year colleges. Definitions of the variables are located in Table 4.
The panel D specification adds AFQT and an AFQT-tuition interaction to the baseline specification.

Table 17

Average Effects of a $1000 Increase in Tuition on High School
Completion and Combined Two- and Four-Year College Entry
(Own-Effect as a % of the Total Effect is in Parentheses)

	Blacks (1)	Hispanics (2)	Whites (3)
A. Gross Tuition (Aggregates Two- and Four-year College, Exclude Region Dummies)			
Overall College Entry High	-.051 (80%)	-.070 (96%)	-.041 (116%)
School Completion	-.021	-.004	.015
B. Gross Tuition (Aggregate Two-and Four-year College, Include Region Dummies)			
Overall College Entry High	-.036 (93%)	-.052 (104%)	-.030 (119%)
School Completion	-.006	.005	.010
C. Gross Tuition (Disaggregate Two-and Four-year College, Exclude Region Dummies)			
Overall College Entry High	-.067 (87%)	-.082 (98%)	-.048 (112%)
School Completion	-.021	-.004	-.015
D. Gross Tuition (Disaggregate Two- and Four-year College, Exclude Region Dummies, Exclude Tuition and Secondary School Equations)			
Overall College Entry	-.060 (na)	-.080 (na)	-.054 (na)
High School Completion	na	na	na
E. Gross Tuition (Disaggregate Two- and Four-year College, Exclude Region Dummies, Include AFQT Score in Initial Condition and all Transition Equations)			
Overall College Entry	-.056 (82%)	-.071 (96%)	-.071 (96%)
High School Completion	-.017	-.006	-.006
F. Net Tuition (Disaggregate Two-and Four-year College, Exclude Region Dummies)			
Overall College Entry	-.026 (70%)	-.051 (109%)	-.051 (109%)
High School Completion	-.015	.011	-.011

Notes: 'na' stands for not applicable. Tuition is measured as gross tuition as the average tuition over all two-year colleges in the country of residence or as a state-level average if there are no two-year colleges in the county. In Panel E, 'Net Tuition' is defined as gross tuition less the amount for imputed Pell grants. Pell grant values were imputed using the algorithm reported in Mortenson (1989a). This variable may be negative as Pell grants may exceed gross tuition costs as the calculation of Pell grant awards considers board and room and other expenses of college.

Table 18

Are Family Income Effects and Tuition Effects of Equal Magnitude?
P-values of Chi-square Test of Equal Magnitude of a $1000 Rise in Tuition and a $1000 Rise in Family
Income
(standard errors of estimates in parentheses)

	Blacks (1)	Hispanics (2)	Whites (3)
A. Excluding AFQT			
(1) $1000 Rise in Family Income	.0035(.0012)†	.0026(.0011)†	.0028(.0009)†
(2) $1000 Rise in Tuition	-.04 (.02)	-.06 (.03)	-.06 (.02)
(3) P-values of Equality Test	.08	.03	.02
B. Including AFQT			
(4) $1000 Rise in Family Income	.0008(.002)	-.001(.001)	.0009(.007)
(5) $1000 Rise in Tuition	-.03(.02)	-.06(.04)	-.05 (.02)
(6) P-values of Equality Test	.18	.08	.07

Available on Request

Appendix A. Data and Sample-Inclusion Criteria

Introduction

This appendix contains information on how we construct our sample as well as a discussion of how we construct the choice set of schooling facing persons, how we measure local tuition and college proximity and local labor market conditions. We also discuss the data on the Armed Forces Quali...cation Test (AFQT). The discussion begins with a brief discussion of the NLSY data.

1. Background on the NLSY Data

The micro data we use are from the 1979-1989 waves of the National Longitudinal Survey of Youth (NLSY). The NLSY includes a randomly chosen sample of 6,111 U.S. youths and a supplemental sample that includes 5,296 Black, Hispanic, and non-Black, non-Hispanic economically disadvantaged young people. Interviewees have been surveyed annually since the initial wave of the survey in 1979, when sample members all ranged between age 14 and 21. Our sample is restricted to males who were in the random sample or the Black and Hispanic supplemental samples. Combining Blacks and Hispanics from the random sample with the minority supplemental samples yields a total of 1461 Blacks and 939 Hispanics. The random sample alone has 2437 non-Black, non-Hispanic youth. For brevity, we refer to this group as the "White" sample. These samples are all statistically representative of their populations.

2. Sample Inclusion Criteria

Information on schooling, county and state of residence, and family income is limited or nonexistent for many sample members so we limit ourselves to individuals age 13 to 16 in January of 1978, when monthly attendance records commenced in the NLSY. This gives us about 60% of each sample. In addition, approximately 6 to 8 percent of each sample were excluded from the analysis for several reasons. First, individuals not attending school full-time at age 15 were excluded (.2% of Whites, .8% of Blacks, and 2.2% of Hispanics —see Table 39 of Cameron [1996] for discussion of these numbers). Next, approximately 1 percent of each sample who reported having completed grade 11 or higher at age 15 were

also eliminated (grade 9 is the mode highest grade completed). Next, about 3% of each sample is dropped because schooling records were seriously incomplete or contradictory. Finally, individuals who missed more than one interview between their initial 1979 interview and their age 21 interview (which occurs between 1982 and 1986) were also excluded to ensure that we have complete records on school attendance on our sample during the most important schooling years, (we lose between 3 and 5 percent of each sample due to this criterion). It should be also noted that about 2 percent of individuals who were included in each initial sample attrited from the data sometime between age 21 and 24.

These restrictions bring our sample sizes to 915 Blacks, 686 Hispanics, and 1417 Whites. The summary statistics shown in Tables 1 to 6 are calculated from these samples. For the multivariate analysis, approximately 13 percent of each sample was lost due to missing values in family income, AFQT, or another variable. (For a discussion of these deletion decisions, see Section A.4 below).

One ...nal limitation was placed on the fraction of observations used in the multivariate analysis. In order to guarantee that AFQT test scores were not in‡uenced by college attendance, a small fraction of the sample was dropped - those who took the test after completing high school. The ASVAB test battery, from which the AFQT is derived (see below), was administered in the summer of 1980. Individuals who were in the oldest cohort of our sample and who were also in grade 10 at age 15 would have been eligible to enter college in the fall of 1979 if they attended school without interruption. About 3% of each sample were dropped after age 17 to ensure that AFQT is measured prior to college attendance. The oldest cohort of the NLSY data comprise roughly 30% of each sample, and approximately 10% of that cohort is in grade 10 at age 15 (see Figure 3A). The remaining AFQT scores are age-adjusted (linear age e¤ects are removed).

3. Data on Schooling Choices

Monthly school attendance is measured in the NLSY beginning in January of 1978. Before that date, information on schooling attendance is limited, though individuals are asked about any prior spells of nonattendance and when they left school if they are not attending in the ...rst survey in 1979. Together with information on highest grade completed

and diplomas (whether or not a high school diploma or a GED diploma was obtained and when), a continuous schooling history was constructed. The data on highest grade completed shows a number of obvious inconsistencies, which require hand edits and some programming to correct. For college attenders, information is gathered in each interview on the type of college (two- or four-year) attended (for up to 3 colleges if more than one was attended since the last interview) and the beginning and ending dates of attendance. Attendance in a grade in this paper is de…ned as two or more months in the grade.

Age is measured in mid-October of each year to approximate age enrollment cuto¤s for Kindergarten. This convention also facilitates comparison of the NLSY to October CPS school enrollment supplements, which are used as a check on the NLSY data. The school year is assumed to last from September 1 through May 31 of each year. A person who reports attending for at least one month of this period is considered a part-time attender. A dropout is classi…ed as someone who does not attend at all.

4. Data on Family Background, Family Income, and AFQT

Family background is measured by the highest grade completed of a person's mother and father, parental family income, the number of living siblings, and whether the respondent came from a broken home at age 14 (that is, did not live with both biological parents). Information on parental education, broken home, and number of siblings were derived from the young people themselves and was reported for most observations in the data (less than 2% of each sample have missing information on one of these variables). Information on parental income, however, came from parents, and was not available for a person who was no longer considered to be dependent on his or her parents. A dependent is de…ned by the NLSY as a person living at home (a type "A" interview) or not at home but considered a dependent while living in a dorm or military barrack (type "B" interview). For this second group, parental family income is obtained from the respondent and not his or her parents. A person living in his or her own apartment or house (even if at college) is deemed to be independent, and no steps are taken to gather parental family income. Thus family income is generally not known for older members of the NLSY. For our samples, family income was missing in only about 6% of the leases, and over 90% of the remaining cases

had observations on family income from more than one interview. For these cases, a two-year average was constructed for family income around age 15 or 16. The two-year average is a slightly better measure than a one-year observation on family income when used in statistical analyses. (Averaging out measurement errors seems to help).

AFQT is a weighted sum of four tests (focusing on reading skills and numeracy) of the ten-part Armed Services Vocational Aptitude Battery (ASVAB), which was administered to NLSY respondents in the summer of 1980. About 7% of our sample did not complete the ASVAB, and so the AFQT score is not available for them. Altogether, missing data on the AFQT, family income, and other variables together eliminate about 13% of the samples from the analyses of the text.

5. Data on Local Labor Market Conditions

The NLSY Geocode data has present annual unemployment rates (of prime age males) in the county of residence for each year. However, this variable has little impact on schooling choices. We also merged into the NLSY a supplementary a data set from the Bureau of Economic Analysis containing detailed annual measures of labor market conditions (average hourly wages and employment) by industry. These data are collected mainly from state unemployment insurance records.

For secondary school transitions (through GED attainment), we construct a measure of average earnings per job in the unskilled sector as an opportunity cost of schooling. Since our data are by major industries and not occupations, we use the average earnings per job in the service and retail trade industries to proxy what a high school dropout would earn.

For analyzing transition into college, the set of industries for which average opportunities wages are constructed is expanded to include manufacturing, construction, mining and extraction, and transportation and public utilities.

6. Tuition Data and Pell Grant Eligibility

Annual records on tuition, enrollment, and location of all public two- and four-year colleges (including universities) in the United States were constructed from the Department of Education's annual HEGIS and IPEDS "Institutional Characteristics" surveys. By matching location with a person's county of residence, we were able to determine the presence of

50

two-year colleges and four-year colleges, in the county of residence. Tuition measures were taken as enrollment-weighted averages of all two- or four-year colleges in a person's county of residence (if available) or at the state level if not available (in which case the indicator for presence of college was set to zero). Choice sets and tuition costs, were de...ned as those relevant at age 17 or 18, to avoid the problem that people who actually attend college might live in the same county as the college.

For some of the analysis, the amount of Pell Grant eligibility was needed. These variables were imputed using parental family income and number of siblings using the appropriate annual formulae summarized in Mortenson (1988).

Problems with the CPS Data

In‡uential studies by Hauser (1991) and Kane (1994) use CPS data to investigate the role of tuition costs, ...nancial aid, family background, and family income in explaining the time series of black-white college attendance. A major problem with using these data to correlate family income with college entry is that young persons who do not live at home and who do not live in group quarters (i.e., dormitories if they attend college) are assigned their own income in the survey rather than that of their parents. This limitation has given rise to a convention in the CPS-based determinants of schooling literature— estimating the determinants of college-entry on samples restricted to dependent children who are high school graduates. It also gives rise to a focus on college enrollment rather than on college graduation, despite the greater importance of the latter in determining career outcomes. The CPS dependency link between youth and their parents becomes much weaker for youth making post-secondary schooling decisions beyond initial entry decisions since it is unusual for students to stay at home for many years after leaving high school.

This convention creates two distinct problems: (a) excluding non-dependents means the sample is no longer random and may not represent family-income e¤ects on schooling participation for the population more generally and (b) statistical problems arise because the factors that in‡uence the decision to live as a dependent may also govern college attendance decisions.[1] All said, CPS analysts underestimate (in absolute value) the e¤ect on

---

[1]Conditioning on a choice variable (dependency status) generates a potential simultaneous equations

schooling attainment because any variable, such as family income, that moves college attendance and dependency status together will overstate (in absolute value) the true e¤ect on variables that have opposing e¤ects on attendance and dependency. For example, if higher family income is associated with both higher dependency and higher college attendance by children, then the estimated e¤ect of family income on college attendance will be smaller than the true population e¤ect if one only studies samples of dependents.

Cameron and Heckman (1992) demonstrate the impact of this measurement problem on estimating the impact of family income on college attendance and show that strong e¤ects of family income are obscured by this convention, particularly for Blacks.

Appendix B
The Likelihood Function

Using the notation introduced in Section 2, we write out the likelihood function for the model we estimate in this paper. We use the notation "$d_{a;j;c}$" for the realized value of $D_{a;j;c}$. (For the initial condition, this is abbreviated to $d_{a;c}$). Associated with each choice set $C_{a;j}$, we have a probability over the destination states that can be accessed from initial state $a;j$, which we sometimes denote by $a;j_a$ to recognize that the educational attainment state at age a is characterized by an age-speci…c set of origin states. We may write the probability of the history

$$(D_{a;c} = d_{a;c}; D_{a+1;c_a;c} = d_{a+1;c_a;c}; ::::; D_{\bar{a};c_{\bar{a}_i\ 1};c} = d_{\bar{a};c_{\bar{a}_i\ 1};c})$$

conditional on $Z_a; Z_{a+1;c_a}; ::::; Z_{\bar{a};c_{\bar{a}_i\ 1}}$ and $\acute{}$ as

$$(B-1) \quad \prod_{c2C_a} [Pr(D_{a;c} = d_{a;c} \mid Z_a; \acute{})]^{d_{a;c}} \quad \prod_{c2C_{a+1;c_a}} [Pr(D_{a+1;c_a;c} = d_{a+1;c_a;c} \mid Z_{a+1}; \acute{})]^{d_{a+1;c_a;c}}$$

$$\text{¢¢¢} \prod_{c2C_{\bar{a};c_{\bar{a}_i\ 1}}} [Pr(D_{\bar{a};c_{\bar{a}_i\ 1};c} = d_{\bar{a};c_{\bar{a}_i\ 1};c} \mid Z_{\bar{a};c_{\bar{a}_i\ 1}}; \acute{})]^{d_{\bar{a};c_{\bar{a}_i\ 1};c}}$$

where

---

problem since dependency status likely is a¤ected by the same unobservables governing college attendance decisions. By using dependency as a "casual" variable, CPS analysts likely produce biased estimates of the impact of socioeconomic variables on college attendance.

$$c_{\underline{a}} = \text{Arg} \max_c d_{\underline{a};c}; \qquad c \, 2 \, C_{\underline{a}}$$
$$c_{\underline{a}+1} = \text{Arg} \max_c d_{\underline{a}+1;c_{\underline{a}};c}; \qquad c \, 2 \, C_{\underline{a}+1;c_{\underline{a}}}$$
$$\vdots$$
$$c_{\bar{a}} = \text{Arg} \max_c d_{\bar{a};c_{\bar{a}_i \, 1};c} \qquad c \, 2 \, C_{\bar{a};c_{\bar{a}_i \, 1}}$$

where Argmax is applied the second subscript of $d_{a;b}$. This selects the element in each choice set which is chosen at each age because $d_{a;b} = 1$ if b is selected; $d_{a;b} = 0$ otherwise for each age and each choice set. The random e¤ects likelihood integrates out ´ with respect to the distribution $F(´)$. The functional forms for the probabilities are given by equation (3) for the transition probability and by equation (4) for the initial condition.

Following Heckman and Singer (1984), we approximate $F(´)$ by a discrete distribution with mass points:

(B-2) $(P_i; ´_i)_{i=1}^I$

where $P_i \, ¸ \, 0$ is the probability associated with the mass point $´_i$, and $\sum_{i=1}^I P_i = 1$. This is the representation for the nonparametric maximum likelihood estimator of $F(´)$. (Lindsay (1995)), and $I \, ! \, 1$ as sample size (N) gets large ($N \, ! \, 1$). Chen, Heckman and Vytlacil (1998) present su¢cient conditions for the NPMLE to produce $\sqrt{N}$ consistent estimates of the parametric portion of the model. Alternatively, we may assume that the true model for F is a ...nite mixture (I is ...xed and bounded $I \, \cdot \, I'$): Under the latter assumption, we can produce conventional ($N^{1=2}$) errors for the estimated parameters using the information matrix.

Following the suggestions in Heckman (1981), we can estimate the model recursively determining (B-2) o¤ the probability for the initial conditions. Alternatively, we can estimate the model jointly for initial conditions and subsequent transitions. In this paper, after considerable experimentation, we ...nd that $I = 2$ describes the data. This low dimensionality has been found in many studies of mixture models. (See e.g. Heckman and Walker, 1990). Setting $´_1 = 0$ and $´_2 = 1$, we estimate P_1 (and $P_2 = 1 \, ¡ \, P_2$) and to obtain a pre-speci...ed variance for ´ we multiply by constant k. We pick k so that $Var(´) = 1$, a normalization needed to identify the factor structure and slope coe¢cients. For each choice

probability structure associated with the choice set $C_{a;j_a}$ we need to normalize one $\bar{}_{a;j_a;c}$ to some value to identify the remaining parameters. Let c_a^{π} be the normalization choice; then $\bar{}_{a;j_a;c^{\pi}}$ and $®_{a;j_a;c^{\pi}}$ are constrained to zero within each choice set.

Appendix C
Tests of Model Specification

Each transition probability in the general model exposited in the main body of the paper (Section 2) is parameterized by a separate coefficient vector (both $\beta_{a,j,c}$ and $\alpha_{a,j,c}$) for each age, each origin state, and each destination in the $C_{a,j}$ choice set as well as an unobserved heterogeneity component η governed by distribution $F(\eta)$. As noted in Section 1. most individuals follow a standard path through school. They complete grade 9 at age 15 (the mode grade) and continue through school without interruption until they terminate their schooling. Few individuals return to secondary school after quitting, although many get a GED. Many possible transition paths are rare, including some that are never observed. Few individuals are still attending secondary school at age 19, say, and attendance at that age is even rarer for individuals who have only completed grade 8. Thus, many possible parameter vectors cannot be estimated because of data scarcity.

Even if there were plenty of data to estimate all possible transitions, it is not clear that it is necessary to estimate separate age-specific and origin-specific parameters for all transitions. Such generality in model specification comes at the cost of potential inefficiency. Thus it is of interest to apply statistical testing procedures to find a more parsimonious specification.

We take two paths toward model parsimony. The first deals with our treatment of rare events. For transitions made by few individuals in these data (fewer than 30 observations). the slope parameters associated with the regressors Z_t generally cannot be estimated with any precision. In such cases, the transition probability is modeled to depend only on an intercept term; all the elements of $\beta_{a,j,c}$ except the intercept are assumed to be zero and $\alpha_{a,j,c}$ is also set to zero. This decouples these transitions from the heterogeneity distribution while at the same time accounting for all of the observed sample paths.

The second path toward parsimony is to test restrictions on $\beta_{a,j,c}$ and to impose them if they are not rejected. For each demographic and ethnic group, four types of restrictions are tested in the following order: (1) Interactions with the origin state $\beta_{a,j,c} = \beta_{a,j',c}$; $\alpha_{a,j,c} = \alpha_{a,j',c}$; $j \neq j'$. (For example, whether the determinants of the transition to completing the next highest grade at age 15 are governed by the same parameters whether a student starts in grade 9 or grade 10). (2) Age interactions $\beta_{a,j,c} = \beta_{a',j,c}$, $a \neq a'$; $\alpha_{a,j,c} \neq \alpha_{a',j,c}$. (For instance, whether or not the determinants of continuing from grade 10 to 11 are identical at age 16 and age 17 for school attenders.) (3) Destination state interactions $\beta_{a,j,c} = \beta_{a,j,c'}$; $\alpha_{a,j,c} = \alpha_{a,j,c'}$; $c \neq c$ (whether the determinants of entry into two-year college are the same as entry into four-year college, for instance). (4) Once these tests are used

1

to select models, we test equality of the α and β parameters across race and ethnic groups (e.g. whether the transition-specific parameters are the same for Blacks, Hispanics, and Whites).

These tests are implemented only for the slope coefficients (denoted $\bar{\beta}_{a,j,c}$) and the factor loadings ($\alpha_{a,j,c}$) where intercept parameters can be freely specified. We test the following hypotheses in this order:

(C-1a) $\bar{\beta}_{a,j,c} = \bar{\beta}_{a,j',c}$, $\alpha_{a,j,c} = \alpha_{a,j',c}$, $j \neq j', \forall a, j, j', c$

(C-1b) $\bar{\beta}_{a,j,c} = \bar{\beta}_{a',j,c}$, $\alpha_{a,j,c} = \alpha_{a',j,c}$, $a \neq a', \forall a, a', j, c$

For college going, we also test equality of the destination state (two- and four-year colleges).

(C-1c) $\bar{\beta}_{a,j,c} = \bar{\beta}_{a,j,c'}$ $\alpha_{a,j,c} = \alpha_{a,j,c'}$, $c \neq c', \forall a, j, c, c'$

If these hypotheses are not rejected, we impose equality of slope coefficients allowing intercept terms to be freely specified. We allow the intercept terms to be freely specified and define them as $\gamma_{a,j,c}$.

Findings From The Model Specification Tests.

This section presents the sequence of tests that yield the final specification used to conduct the simulations reported in this paper. Our findings are organized into four sections corresponding to: (1) tests on initial conditions; (2) tests on secondary school transitions for those attending school; (3) tests on transitions for those who have dropped out of secondary school; and (4) tests on transitions into college from high school completion.

Initial Conditions or Initial Grade Level

The initial conditions at age 15 consist of three states: grade 10, grade 9, or grade 8 or less. Individuals who are not in school at age 15 are deleted from the sample (less than 1 percent of the total sample for all ethnic groups), and individuals who completed more than 10 years of school (less than 1 percent of the total sample) were also eliminated. Two tests were performed on initial grade probabilities. First, a test of slope coefficient and factor loading equality was performed to determine whether grade 8 less should be disaggregated in two separate states (grade 8 and grade 7 or less). (Test C-1a). Equality was not rejected for any racial-ethnic group (see row (1) of Table C-1).[1] Second, tests to gauge whether minority initial grade probabilities are governed by the same parameters as those of Whites were also rejected (row (2) of Table C-2). Likelihood values for the different models are reported in Table C-3.

Secondary School Transitions for School Attenders

From age 16 through 24, the choice set for secondary school transitions for individuals actively attending school may include the transition to full-year school attendance and

---

[1]Note that a test of parameter equality is only a test of a necessary condition for aggregation. When two states are aggregated in this way, the two specific-state error terms ($\nu_{a,j,c}$) for grade 8 and grade 7 or less are collapsed into one error term representing grade 8 or less. We do not test this additional restriction in this paper as it raises new conceptual problems, not analyzed in the existing statistical literature.

completion of the next highest grade, dropping out, partial attendance, attendance and dropping out to obtain the GED in the same school year. and full-year attendance while repeating the current grade. The final two states are rare events and these transitions are characterized using an intercept parameter only. Attendance in secondary school after age 19 is also rare, so the probability of those events is also modeled using only an intercept. The contrast state[2] is "drop out."

At age 16 through 18, two types of tests are performed. The first is a test of origin-state interactions in the slope parameters and factor loadings (C-1a). It is performed for each transition at age 16 through 19. For age 16 and 17, a p-value of tests for slope and factor loading parameter equality across the three most relevant grade levels (the mode grade and the grades above and below) are reported in rows (2) and (3) of Table C-1.[3] The hypothesis is not rejected. At age 18, the mode transition is school attendance to high school graduation (grade 12), so the tests were performed only for interactions in the transitions to the next highest grade from attending grades 10 and 11 only (row (4)). The equality restrictions are not rejected at all ages at the 5 percent level.[4] Likelihood values are reported in Table C-3.

Next, we test for age-interactions in the slope coefficients (C-1b). We test equality at adjacent ages (age 16 coefficients equal to age 17 coefficients, age 17 coefficients equal to age 18 coefficients, and age 18 equal to age 19). For all racial-ethnic groups , (C-1b) is strongly rejected except for age 18 and 19. (See rows (5)-(7) in Table (C-1)). It is important to point out, though, that the mode age of secondary school completion is 18, and relatively few individuals are in school after this age. Thus for school continuation, age but not grade matters in determining grade transitions from one age to the next.

We do not test hypothesis (C-1c) for the secondary school transitions. Finally, after imposing all of the restrictions accepted from the previous procedure, we test equality of slope coefficients and factor loading between Whites and minorities. Racial-ethnic equality restrictions are also rejected (row (3) of Table C-2) even when dummy variables for race/ethnicity are entered in each transition. Parameter estimates for the model of initial conditions are displayed in panel A of Table C-4 (grade 8 or less is left-out or contrast state).

Parameter estimates of the final equation for secondary school transitions (excluding

---

[2] As noted in Section 2, in multinomial discrete choice models, it is standard to set to zero the parameters of one alternative for model identification. In this paper and in the literature, this alternative is referred to as the "constrast" state.

[3] The tests were also performed pairwise on the parameters of adjacent states with no change in the conclusions.

[4] The transition from grade 11 to graduate high school was the only transition that showed any evidence of origin-specific interactions particularly for Whites, with a p-value of .07.

3

constants) are presented in panel B (age 16), C(age 17), and D(age 18 and 19) in Table C-4 (recall that beyond age 19, secondary school transitions are captured by intercept terms only).

Secondary School Transitions for Dropouts

Dropouts are assumed to face three choices: return to school and complete the next grade, remain out of school, and obtain a GED Equivalency degree. Compared to secondary school transition for school attenders, returning to school and GED attainment are relatively low-probability occurrences, and the samples are not sufficiently large to permit testing for origin-states interactions. Thus, a dummy-variable specification is used where binary indicator variables are used to represent the relevant additive states in the state space (highest grade completed for those not attending school). In addition, age effects are assumed to operate. Thus, a multinomial logit model (conditional on the random effect) with common slope parameters and factor loading terms represents these choice probabilities from age 16 through 20. There is virtually no secondary schooling attendance after age 20, and the mode age of GED completion is 18. (Cameron, 1996). Hence, all transitions for dropouts between age 21 and 24 are modeled using dummy variables. Parameter estimates of the final equation (excluding constants) are presented in panel E (age 16) of Table C-4.

College Attendance

Individuals can enter college at one age from four possible states at the previous age: high school attendance and just graduating, high school graduate not previously attending, GED holder attending secondary school the previous year, and GED holder not attending secondary school in the previous year. Before any testing is performed, several restrictions are placed on the specification to eliminate parameters that cannot be estimated because of data scarcity. First, the mode college entrant completes high school at age 18 and immediately enters college during his 19th year. The youngest high school graduates and GED recipients in our sample complete school at age 17 and are eligible for college entry at age 18. These young graduates make up less than 8 percent of the entire sample, so full age interactions are not estimated for college entrants at age 18 and 19. Instead, the probabilities are modeled by a single transition equation with a dummy variable to indicate age. Second, as little college entry occurs after age 20, the college entry changes at age 20 through 24 are also estimated by a single multinomial logit with dummy variables to represent age. Third, regardless of attendance status in the previous years GED completers make up only a small fraction of college entrants. High school graduates who are not attending school also comprise a tiny fraction of college entrants. The samples are too small to permit testing a full set of origin-state interaction terms, so we assume that college entry probabilities depend on origin state only through a dummy variables representing high school degree (GED or high school graduation) and attendance status (in secondary school last year or not). A similar specification is used to estimate entry at age 18 and 19 and

4

for age 20 through 24. Starting with this specification. three tests of parameter equality are performed and all are rejected. First. we conduct a test to determine whether the determinants of entry into two-year college are the same as those of entry into four-year college. (Hypothesis (C-1c)). Rows (9) and (10) of Table C-1 shows that these tests are all rejected with p-values of less than .005 (reported as .00 in the table). Cameron (1996) shows that, compared to the standard specification in the literature, which treats college entry as a single choice, disaggregating college entry into two-year and four-year college entry greatly improves the ability of the model to explain differences in White-minority college entry rates. Second, the hypothesis of age effects for the slope parameters for the two college entry equations (age 18 or 19 and age 20 through 24) is rejected for all racial-ethnic groups at the 5 percent level (row (11) of Table C-1). Finally, equality of slope coefficients college entry for Whites and minorities is also rejected (row (4) of Table C-2).

Do We Need to Account for Heterogeneity?

Throughout the paper, we account for correlated heterogeneity across transitions. We have already discussed this issue in the text using a variety of goodness of fit tests. Table C-5 presents an additional justification for doing so. We use a BIC test as a guide to determining whether the fit of the model is improved by adding the additional heterogeneity parameters. As reported in row (4), BIC favors a two point heterogeneity model over a no heterogeneity or three point model. An *ad hoc* likelihood ratio test supports the same conclusions.

A Note about Two other Model Extensions

Before concluding this subsection, we mention two other generalizations of the model. discussed in detail in Cameron (1996). First, because schooling attendance is measured over a one-year interval, one could imagine that a person could attend part-time during the year - either dropping out during part of the year and returning or beginning the school year and leaving before it is over. The empirical unimportance of dropping out and return during the period of secondary schooling are discussed in section 1 (see also the data description in Appendix A). Nevertheless, Cameron (1996) tests whether the determinants of "part-year" attendance status is identical to "not attend" and finds that they are. It should be noted that the way most dropouts leave school is by attending through the end of one school year and not returning the next year. The next most common event is to start a school year and then leave in the middle. Few dropouts return to school in the middle of a year (see the tables in Cameron 1996, especially Tables 26 and 27). Hence, in this paper we adopt the convention that individuals attend and complete a grade as opposed to not complete (which includes part-year attendance).

In an effort to find a model that can explain White-minority gaps schooling gaps by endowments rather than "tastes" (see the discussion in Section 4 of the text), we estimate another model. School non-attendance states were disaggregated into two states: "dropout

5

and work" and "dropout and not work," where "work" is four weeks or more of working. The results of this analysis can also be found in Cameron (1996) along with some summary statistics of the relative size of the work and not-work status (see Tables 29 and 36). Thus, a person drops out of secondary school into a work or non-work state, a dropout returns to school or obtains a GED from work or non-work states, and a high school completer enters college from the work or non-work states. This model does little to improve our ability to explain White-minority schooling gaps. Disadvantageously, the variance of predicted schooling also increases substantially. An additional disadvantage is the substantial increase in the number of states and slope parameters that could no longer be estimated. This lack of improvement in the predictive power of the model should come as no surprise, because as previously noted, most students continue uninterrupted from age 15 until completion. In addition, even though we use a lazy man's definition of work to classify an individual as working in a year (four weeks of work or more), only about half of high school dropouts work in a year until age 18, when the fraction rises to about 60 percent (see Table 39 of Cameron 1996).

Summary

Four general conclusions can be sustained about the importance of various interactions in the data. First, the data show little support for hypothesis C-1a: origin states don't matter for secondary transitions. For instance, for secondary school attenders at age 16 the determinants of the transition to next highest grade is the same for individuals in grades 9, 10, or 11 (the intercept term was unrestricted). Second, by contrast, age effects (C-1b) are found to be important for both secondary school attendance and college entry. The determinants of school continuation for attenders are statistically different at age 16, 17, and 18 (there is very little attendance after age 18) at the 10 percent level. Third, disaggregation of the college going state into two-year and four-year colleges is empirically important. Fourth, equality of slope coefficients between Whites and Blacks as well as Whites and Hispanics is rejected for all schooling states. This finding is summarized in row (1) of Table C-2,which shows p-values of a joint test of racial-ethnic equality of all slope coefficients and factor loadings excluding intercept terms. Given the modest size of our samples, this finding is strong evidence against the notion that differences in endowments alone can explain White-minority schooling gaps.

6

Table C-1

P-Values of Chi-Square Test Statistics of Various Model Specification Tests

	Blacks (1)	Hispanics (2)	Whites (3)

A. Initial Conditions* How many states are needed to represent initial school state distribution. (13 slope + 1 factor leading parameters = 14 degrees of freedom for each test below)
H_0: (C-1a)

	Blacks (1)	Hispanics (2)	Whites (3)
(1) Can Grades 8 and 7 or Less Be Pooled?	.50	.69	.27
(2) Can Grades 8 and 9 be Pooled?	.00	.02	.00
(3) Can Grades 9 and 10 be Pooled?	.13	.01	.03

B. State-of-Origin Interactions in the Slope
Parameters of the School Continuation Decision For School Attenders*
H_0: (C-1a)

	Blacks (1)	Hispanics (2)	Whites (3)
(4) Age 16	.49	.51	.38
(5) Age 17	.32	.60	.23
(6) Age 18	.19	.38	.04

C. Age Aggregation of Secondary School Transitions*
H_0: (C-1b)

	Blacks (1)	Hispanics (2)	Whites (3)
(7) Equality of Age 16 and Age 17 Transitions	.01	.00	.00
(8) Equality of Age 17 and Age 18 Transitions	.00	.00	.00
(9) Equality of Age 18 and Age 19 Transitions	.09	.18	.29

D. Equality of Determinants of Two- and Four-year College Entry*:
(Note: Entry is disaggregated into two-and four-year college entry
(28 degrees of freedom for each test)
H_0: (C-1c)

	Blacks (1)	Hispanics (2)	Whites (3)
(10) Equality of Age (18-19) and Age 20 Transitions	.00	.15	.04
(11) Age 20 and Age 21-24 Transitions	.30	.61	.19
(12) Equality of Age (18-19) and Age (20-24) Transitions (tested after imposing that determinants at age 20 = age(21-24)	.00	.11	.01

E. Equality of Determinants of Two-and Four-year College Entry,* (Can two-year and four-year entry be pooled into a single college entry state)?

	Blacks (1)	Hispanics (2)	Whites (3)
(13) Equality of Two-and Four-Year College Determinants	.01	.05	.00
(14) Equality of Two- and Four-Year College Determinants for Age 20 to 24 Transition	.02	.03	.02

Table continues on next page

Note 1. Age 18 and 19 entry was aggregated into one equation up to an age 18 indicator because lack of data prevented fully interacted estimation of an age 18 college entry equation. Entry at age 21 through 24 was also aggregated into one equation with age indicator variables to account for age effects.

Table C-2

Tests of Racial-Ethnic Equality of Slope Coefficients
P-Values of Chi-Square Tests

	Tests of Equality of ($\beta_{a,j,c}$ and $\alpha_{a,j,c}$) Across Groups	
	Whites & Blacks (1)	Whites & Hispanics (2)
(1) Entire Model	.00	.00
(2) Initial conditions only	.01	.00
(3) Secondary-school transitions only	.00	.00
(4) College-entry transitions only	.00	.03

Note: The model is includes dummy variables to indicate racial-ethnic identity (Black, Hispanic, or White, in all transitions. The tests are standard likelihood ratio tests with a Chi-square distribution .

Table C-3

Negative Log-Likelihood Values Associated with Various Departures from the Baseline Model
See Table C-1 for P-values of the Chi-Square Test Statistics
See the main text and appendices for the baseline specification

	Blacks (1)	Hispanics (2)	Whites (3)
Negative Log-Likelihood Values of Baseline Specification	2543.6	1538.1	4198.4
A. Initial Conditions*			
(1) Can Grades 8 and 7 or Less Be Pooled?	2537.0	1532.7	4190.1
(2) Can Grades 8 and 9 be Pooled?	2593.6	1560.1	4265.6
(3) Can Grades 9 and 10 be Pooled?	2553.5	1552.2	4211.2
B. State-of-Origin Interactions in the Slope Parameters of the School Continuation Decision For School Attenders*			
(4) Age 16	2557.8	1548.9	4213.4
(5) Age 17	2559.1	1550.1	4230.0
(6) Age 18	2560.0	1554.3	4219.8
C. Age Aggregation of Secondary School Transitions*			
(7) Equality of Age 16 and Age 17 Transitions	2558.5	1240.4	4230
(8) Equality of Age 17 and Age 18 Transitions	2575.9	1576.7	4156
(9) Equality of Age 18 and Age 19 Transitions	2558.0	1557.1	4206.2
D. Age Interactions in college entry (Note: entry is disaggregated into two- and four-year college entry)			
(10) Equality of Age (18 & 19) and Age 20 Transitions	2555.5	1542.2	4210.2
(11) Equality of Age 20 and Age (21-24) Transitions	2551.5	1544.0	4207.1
(12) Equality of Age (18-19) and Age (20-24) Transitions	2561.9	1548.0	4208.3
E. Equality of Determinants of Two-and Four-year College Entry*			
(13) Equality of Two-and Four-Year College Determinants for Age 18 and 19 Transition	2560.8	1555.3	4232.9
(14) Equality of Two- and Four-Year College Determinants for Age 20 to 24 Transition	2560.5	1551.1	4211.8

*Notes: See notes to Table C-1.