

APPENDIX A: SAMPLE DEFINITIONS AND TREATMENT OF MISSING DATA

To study black-white differences in educational and economic outcomes, efforts were made to construct comparable samples using longitudinal survey data. For the study of labor market outcomes, four samples were defined. These same samples were also used in the study of educational attainment outcomes, together with two additional samples used for the analyses described in appendix E. For the study of educational achievement outcomes, four samples were defined. Each of these samples, and the variables within them, are discussed in this appendix.

In these samples, key analysis variables are missing for a large proportion of cases. Excluding these cases from the analysis would greatly reduce the sample size and could potentially bias the results.¹ Missing data for key variables were imputed using multiple imputation. Section I describes this approach briefly, while subsequent sections provide further details.

I. MULTIPLE IMPUTATION OF MISSING DATA

The simplest way to deal with missing data is to drop cases with missing values from the analysis. This approach can lead to two problems; the sample size is greatly reduced, leading to less precise estimates, and the estimates will be biased if the cases with missing data are not a random sample of all cases.

Multiple imputation addresses both of these problems. By substituting imputed values for the missing data, the entire sample may be used in the analysis. Careful modeling of the relationships among variables yields imputed values that reflect differences between complete cases and cases with missing values. Multiple imputation has the added advantage of dealing with the uncertainty introduced through imputation. Rather than substituting a single value for each missing value, several plausible values are chosen. Estimates are produced using each of the different imputed values, and the variability among the estimates is incorporated into the standard errors of the estimates. These calculations are described in detail in appendix C, and in Schafer (1997).

¹ In fact, after performing multiple imputation to account for missing data, and redoing our analyses, we obtained similar results as when we used only complete cases, suggesting that any bias from missing data was small.

Imputed data may be obtained using an algorithm known as “data augmentation” (Tanner and Wong 1987). Data augmentation is an iterative process that alternates between estimating the parameters of the distribution of the data, and simulating random variables from this distribution which are substituted for the missing data. These two steps are repeated many times, until the process converges in distribution. At this point the values substituted for the missing data are independent draws from the predictive distribution of the missing data, and may be treated as imputed values.

II. DATA USED IN THE ANALYSES OF LABOR MARKET OUTCOMES

A. Samples

Four samples were used to compare the labor market outcomes of black and white young adults. Two of the samples were defined from school-based surveys: the National Longitudinal Study of the High School Class of 1972 (NLS:72), first conducted in 1972, with follow-up surveys in 1973, 1974, 1976, 1979, and 1986; and the High School and Beyond (HSB) survey of high school sophomores, first conducted in 1980, with follow-up surveys in 1982, 1984, 1986, and 1992.

The other two samples were defined from a household-based survey, the National Longitudinal Survey of Youth (NLSY), first administered in 1979, with annual follow-up surveys thereafter. The first NLSY sample was constructed to be comparable to the NLS:72 sample of high school seniors; the second NLSY sample was constructed to be comparable to the HSB sample of high school sophomores. Individuals in the NLSY were identified as being in grade 12 between 1976 and 1982, and in grade 10 between 1974 and 1980.² To distinguish the two NLSY samples, we refer below to the sample of high school seniors as the NLSY senior sample and to the sample of high school sophomores as the NLSY sophomore sample.³

To allow individuals a considerable length of time to complete their schooling and enter the labor market, employment and earnings were observed several years after high school. Labor market outcomes for the two samples of high school seniors were taken from follow-up surveys occurring 7 years after grade 12, that is, in 1979 for the NLS:72, and between 1983 and 1989 for the NLSY senior sample.

Analysis samples were restricted to cases with known sex and Census region, and which were identified as either non-Hispanic black or non-Hispanic white. The key comparison in this study, black/white differences in the labor market outcomes, relies on reported race, so Hispanics, Asians, Pacific Islanders, and individuals who did not report their race were excluded from the sample.⁴

² Because NLSY participants were sampled in 1979 on the basis of age, rather than on the basis of grade 10 or grade 12 attendance, certain individuals may be missing from the NLSY samples who were present in the NLS:72 or HSB samples. For example, individuals who immigrated to the United States after 1979 and were high school sophomores in 1980 or high school seniors between 1980 and 1982 would be excluded from the sample, as would individuals who were high school sophomores in 1974-1978 or high school seniors in 1976-1978 who either had died before 1979 or were not age 14-21 in 1979. This lack of coverage by the NLSY means that comparison between the NLSY samples and the NLS:72 and HSB samples should be interpreted with caution.

³ There is considerable overlap between the two NLSY samples; 82 percent of the individuals in the NLSY sophomore sample also appear in the NLSY senior sample.

⁴ Individuals with unknown race were combined with other races (Hispanic, Asian, or Pacific Islanders) for the purposes of multiple imputation.

B. Variable Definitions

The following labor market outcome variables were defined in each dataset:

1. *Labor force participation status*—employed or looking for work (defined for all U.S. civilians not living in institutions)
2. *Unemployment status*—looking for work (defined for labor force participants only)
3. *Annual earnings*—the natural log of earnings (in 1992 dollars) for the past calendar year (defined for persons with positive earnings only)
4. *Hourly wage*—the natural log of the hourly wage (in 1992 dollars) for the current job (defined for employed persons only)

An hourly wage variable could not be constructed for the HSB sample since the 1992 survey did not inquire about earnings per hour of work.

The analyses of labor market outcomes used the following variables to describe the different backgrounds of young adults:

1. *Race*—indicator for black young adults (defined for non-Hispanic blacks and non-Hispanic whites only; Hispanic, Asian, or Pacific Islander individuals were excluded from the sample, as were persons of unknown race)
2. *Sex*—indicator for female young adults
3. *Educational achievement* and *educational achievement squared*—an average of mathematics and reading achievement scores, with each component and the overall average normalized to have mean zero and standard deviation one for all individuals (regardless of race) in the same grade in the same year in each sample
4. *Educational attainment* and *educational attainment squared*—number of years of completed schooling
5. *Work experience* and *work experience squared*—cumulative weeks employed divided by 52

High school completion status was not analyzed for the NLS:72 or NLSY senior datasets, since this outcome was not clearly defined in the 1979 NLS:72 sample. Educational achievement scores were normalized separately for each of the class years represented in the NLSY samples (1976 through 1982 for the NLSY senior sample, and 1974 through 1980 for the NLSY sophomore sample), since the underlying achievement tests (the arithmetic, mathematics, paragraph comprehension, and word knowledge sections of the Armed Services Vocational Aptitude Battery) were administered to the entire NLSY sample (aged 15 to 23) in 1980.⁵ For the analyses of annual earnings, the measure of work experience excluded weeks employed since the start of the prior calendar year, since these weeks were included in the outcome variable.

⁵ Note that some individuals in the NLSY took the ASVAB after completing high school and/or attending college. These observations were included in the sample to keep the sample size reasonably large. Because postsecondary educational experiences may affect one's educational achievement relative to one's grade cohort, the educational achievement measures in the NLSY samples differ in important ways from the achievement measures in the NLS:72 and HSB samples, which were administered prior to high school completion.

C. Sample Sizes

Table A.1 presents sample sizes for the four analysis samples. Of the 22,652 persons in the NLS:72 who were high school seniors in 1972, 90 percent (20,273) were included in the analysis of labor market outcomes. About 10 percent (2,374) of the sample were not identified as either non-Hispanic black or non-Hispanic white; an additional 5 cases were missing core background information (sex or Census region). Fifty-seven percent of the 20,273 included cases were missing one or more key analysis variables, which were imputed using the procedures outlined in section D.⁶

Of the 7,962 persons in the NLSY who were high school seniors between 1976 and 1982, 93 percent (7,424) were included in the analysis of labor market outcomes. About 5 percent (396) of the sample were not identified as either non-Hispanic black or non-Hispanic white; an additional 2 percent (142) were missing sex or Census region. Of the cases used in the analysis, 27 percent were missing one or more key analysis variables.⁷

Labor market outcomes for the two samples of high school sophomores were taken from follow-up surveys occurring 12 years after grade 10, that is, between 1986 and 1992 for the NLSY sophomore sample, and in 1992 for the HSB sophomore sample. Of the 9,709 persons in the NLSY who were high school sophomores between 1974 and 1980, 93 percent (8,998) were included in the analysis of labor market outcomes. About 5 percent (528) of the sample were not identified as either non-Hispanic black or non-Hispanic white; an additional 2 percent (183) were missing information on sex or Census region. Missing analysis variables were imputed for 24 percent of the cases included in the analysis.⁸

Of the 14,825 persons in the HSB who were high school sophomores in 1980, 77 percent (11,375) were included in the analysis of labor market outcomes. About 23 percent (3,450) of the sample were not identified as either non-Hispanic black or non-Hispanic white; all of the remaining cases had data on sex and Census region. Forty-five percent of included cases had missing values for one or more analysis variables.⁹

D. Multiple Imputation of Missing Data

Of the cases identified as either non-Hispanic black or non-Hispanic white, with known sex and Census region, between 30 and 57 percent were missing one or more key analysis variables. We address this problem through the use of multiple imputation, which allows the use of all available data for each case, even if some key variables are missing.

Missing values were imputed for each of the analysis variables listed in section I.B, as well as for the parental socioeconomic status variable used in the analysis of educational attainment. Two of the variables to be imputed, presence of earnings and employment/labor force participation status, are categorical. Presence of earnings was imputed separately from amount of earnings because a missing value for log of earnings

⁶ Item nonresponse for the sample of blacks and whites in NLS:72 ranged from 0.05 percent for the employment status variable to 32.00 percent for annual earnings.

⁷ Item nonresponse for the blacks and whites in the NLSY senior sample ranged from 0.26 percent for the SES variable to 23.44 percent for annual earnings.

⁸ Item nonresponse for the blacks and whites in the NLSY sophomore sample ranged from 0.34 percent for the SES variable to 33.27 percent for annual earnings.

⁹ Item nonresponse for the sample of blacks and whites in the HSB ranged from 9 to 58 percent for the SES variable to 27 to 31 percent for annual earnings.

Table A.1 Total sample size and excluded cases (labor market/attainment datasets)

	NLS:72	NLSY-Sr	NLSY-So	HSB
Total cases	22,652	7,962	9,709	14,825
Black	3,119	2,031	2,483	2,238
White	17,159	5,535	6,698	9,137
Other/not reported	2,374	396	528	3,450
Percent missing sex or region	0.19	2.22	2.38	1.34
Black	0.03	1.62	1.57	0
White	0.02	1.97	2.15	0
Other/not reported	1.60	8.84	9.09	5.77
Total cases included in analysis	20,273	7,424	8,998	11,375
Black	3,118	1,998	2,444	2,238
White	17,155	5,426	6,554	9,137
Percent of cases with missing values	54.8	27.1	24.0	44.5
Black	65.3	30.7	19.5	57.2
White	52.9	25.8	25.8	41.3

amount does not distinguish between cases with missing earnings and cases with zero earnings. Employment/labor force participation status has four possible values: civilian employed, unemployed, not in the labor force, and military. However, the small number of military cases made it impossible to treat the military category separately in the multiple imputations. In particular, several subgroups lack cases with military status and no earnings. In order to avoid the estimation problems that arise with empty categories, military cases were combined with employed civilian cases in the multiple imputations. The analyses use the original four-category employment status variable. The military category contains only cases known to be in the military, and all of the imputed cases are assigned to one of the other three categories.¹⁰

The remaining variables were treated as continuous. Although the analysis treats educational attainment as a categorical variable with four categories, sparseness in several categories made it impossible to include this variable in the imputation model. Instead, the continuous variable years of education was imputed, and cases with missing educational attainment status were assigned to an educational attainment category based on imputed years of education.

To improve the imputations, the model included the sample weight along with averages by school and race of all the continuous variables to be imputed.¹¹ These variables were intended to capture much of the difference among schools and regions within each sample.

Since it was necessary to impute missing values for both categorical and continuous variables in the labor market/attainment datasets, a model for mixed continuous and

¹⁰ This should have a negligible effect on the results. For example, in NLS: 72, only 2.2% of cases with known employment status are in the military, and only 10 cases have missing employment status, indicating that it is very unlikely that any cases with missing employment status were actually in the military.

¹¹ The model for the two NLSY datasets does not include average work experience by school and race due to convergence problems when estimating imputed values. This variable is included in the models for HSB and NLS:72, implying that the imputations are somewhat more precise for these two datasets than for the NLSY datasets.

Table A.2 Sample sizes by race/sex category (labor market/attainment datasets)

	NLS: 72	NLSY-Sr	NLSY-So	HSB
Total	20,273	7,424	8,998	11,375
White males	8,650	966	1,232	4,520
White females	8,505	1,032	1,212	4,617
Black males	1,369	2,670	3,274	1,099
Black females	1,749	2,756	3,280	1,139

categorical data was chosen to approximate the relationships among variables. This model, known as the general location model, is described in Schafer (1997), and assumes that the marginal distribution of the categorical variables is multinomial, and the continuous variables have a multivariate normal distribution conditional on the categorical variables. MIX, a library of S-PLUS routines created by Joe Schafer, was used to implement this model.

Missing values were imputed separately for four subgroups: white/non-Hispanic males, white/non-Hispanic females, black/non-Hispanic males, and black/non-Hispanic females.¹² Persons of unknown race or sex were excluded from the imputation. The sample sizes for the four data subgroups are presented in table A.2.

We used the MIX library of S-Plus subroutines to create multiple imputed datasets. For each of the four subsets of each dataset, we first attempted to determine how many iterations were required for convergence of the data augmentation algorithm. Schafer and Olsen (1998), describing this algorithm in greater detail, report that data augmentation “nearly always converges in fewer cycles than does EM.” We ran the EM algorithm several times from a variety of starting values, assessing whether the algorithm converged to a unique solution and how many iterations were required for convergence. We then ran at least that many steps of data augmentation before imputing to be sure that the data augmentation algorithm had converged.

The multiple imputation models were well behaved, converging to a unique value no matter which starting points were used. In almost every subset the EM algorithm converged in fewer than 500 iterations, and it often converged in fewer than 100 iterations. For these subsets we used the results from the EM algorithm as starting values for the data augmentation algorithm, imputing after 500 steps. The EM algorithm required between 500 and 700 iterations to converge for white males in NLS:72 and NLSY-So, and for black males in NLS:72. For these subsets we imputed after 1,000 steps of data augmentation. We then ran four more independent series of data augmentation for each subset, each time using the final values of the previous run as starting values and imputing after 500 (or 1,000) steps.

Finally, for each of the five multiple imputations we merged the four datasets together, creating a single file that included observations from all four sex and race categories. We

¹² Missing data were imputed separately for cases with other race or unknown race for HSB and NLS:72. Missing data were not imputed for other/unknown race cases for NLSY because the small number of such cases led to convergence problems. Since cases with other/unknown race were not used in the analyses described in this report, this appendix focuses primarily on imputation for non-Hispanic blacks and non-Hispanic whites.

merged each imputed dataset onto the original data file and recoded variables that had been standardized to facilitate the imputations.

E. Characteristics of Cases with Missing Data

If there are systematic differences between complete cases and cases with missing data, estimates based on only complete cases may be biased. In particular, if black-white differences were more (or less) pronounced among complete cases than among cases with missing data, and estimates of black-white differences were based only on complete cases, the results would be misleading. A comparison of complete cases and cases with missing values reveals several differences, suggesting that an analysis restricted to complete cases would be biased.

Tables A.3 through A.7 show summary statistics on the key variables in the analysis of educational attainment and labor market outcomes for complete cases and cases with at least some missing variables. We imputed missing data to address the problem of missing values, and the tables include the imputed values. *Note that, because observations with missing values of some variables often have valid values of other variables, the reported means for observations with missing data include both valid and imputed values.* In every instance, the imputation of missing values utilized any valid information for the observation in question.

For characteristics of black and white young adults, differences between complete cases and cases missing one or more variables tended to be similar across the four datasets. The most striking pattern is that respondents with no missing data are more well-off than those for whom data had to be imputed. SES, educational achievement, college attendance and completion, work experience and earnings are higher for the complete cases than for the cases with missing data, for both blacks and whites. The majority of these differences are statistically significant.

Several other differences exist between complete cases and cases with missing data. For example, the percentage black is higher among cases with missing data, and this difference is statistically significant for three of the four datasets. In NLS:72, the percentage female is higher for cases with missing data, for both blacks and whites. However, among black young adults in HSB, the percentage female is higher for complete cases. Some differences in Census region appear as well, although there is no evidence of a consistent pattern across datasets.

The key concern is whether black-white differences vary according to the presence of missing data. Tables A.3 through A7 reveal several characteristics for which this is the case:

- In NLS: 72, blacks are better off relative to whites among cases with missing data, as compared with black-white differences among complete cases. Black-white differences in percentage attending and completing college, earnings, percentage female, and percentage in the Midwest Census region were significantly smaller among cases with missing data. Among cases with missing data, blacks were more likely than whites to have attended college, participate in the labor force, and have earnings, while the opposite was true for complete cases.
- The pattern is not as clear in the NLSY senior sample. Cases with missing data show significantly larger black-white differences in socioeconomic status (SES) and percentage with earnings, and significantly smaller black-white differences in percentage completing college.

Table A.3 Comparison of black and white young adults in NLS:72, by presence of all variables

Sample definition and descriptive variables	Not missing any variables		Missing one or more variables		Difference-of-means (missing - not missing)		Statistical significance
	Mean	Std. Error	Mean	Std. Error	Value	Std. Error	
Population of black and white children	(N=9160)		(N=3417)				
Percentage black	9.2	0.0	11.5	0.0	2.3	0.8	**
Population of white young adults	(N=8079)		(N=2899)				
Percentage female	44.6	0.7	62.2	0.0	17.5	1.3	**
Percentage attending high school in the Northeast	25.7	1.7	26.5	0.0	0.7	1.4	
Percentage attending high school in the Midwest	32.7	1.9	35.5	0.0	2.8	1.8	
Percentage attending high school in the South	25.4	1.6	20.7	0.0	-4.7	1.1	**
Percentage attending high school in the West	16.2	1.4	17.3	0.0	1.1	1.4	
Socioeconomic status	0.109	0.014	0.043	0.018	-0.066	0.017	**
Educational achievement	0.2	0.0	0.0	0.0	-0.2	0.0	**
Percentage that attended college	65.6	0.8	55.3	0.0	-10.3	1.3	**
Percentage that completed college	29.4	0.7	20.8	0.0	-8.6	1.1	**
Average years of work experience	5.1	0.0	3.8	0.1	-1.3	0.1	**
Percentage that are in the labor force	92.5	0.3	60.8	1.3	-31.7	1.3	**
Natural log of wage	2.5	0.0	2.3	0.0	-0.2	0.0	**
Percentage that have earnings	96.4	0.2	69.6	1.2	-26.8	1.2	**
Earnings in thousands	21.0	0.2	13.5	0.4	-7.5	0.4	**
Population of black young adults	(N=1081)		(N=518)				
Percentage female	54.4	1.7	64.7	2.8	10.2	3.1	**
Percentage attending high school in the Northeast	16.6	2.7	11.3	2.2	-5.4	2.3	**
Percentage attending high school in the Midwest	13.1	3.0	27.5	4.8	14.4	4.7	**
Percentage attending high school in the South	64.3	3.6	53.8	4.5	-10.5	4.3	**
Percentage attending high school in the West	6.0	1.3	7.5	2.0	1.5	2.0	
Socioeconomic status	-0.583	0.026	-0.589	0.035	-0.006	0.038	
Educational achievement	-0.9	0.0	-1.1	0.0	-0.2	0.1	**
Percentage that attended college	58.7	2.0	58.5	2.6	-0.3	3.2	
Percentage that completed college	18.2	1.4	16.1	1.8	-2.1	2.2	
Average years of work experience	4.7	0.1	3.3	0.1	-1.3	0.1	**
Percentage that are in the labor force	91.4	1.0	74.3	2.3	-17.1	2.5	**
Natural log of wage	2.3	0.0	2.3	0.0	0.0	0.0	
Percentage that have earnings	93.8	0.8	72.3	2.4	-21.5	2.5	**
Earnings in thousands	17.6	0.4	13.4	1.0	-4.2	1.0	**
Difference between black and white children							
Percentage female	9.8	1.8	2.5	2.9	-7.3	3.3	**
Percentage attending high school in the Northeast	-9.1	2.9	-15.2	2.8	-6.1	2.7	**
Percentage attending high school in the Midwest	-19.6	3.4	-8.0	5.1	11.6	4.9	**
Percentage attending high school in the South	38.9	3.6	33.0	4.5	-5.8	4.3	
Percentage attending high school in the West	-10.2	1.8	-9.8	2.6	0.4	2.4	
Socioeconomic status	-0.692	0.028	-0.632	0.039	0.060	0.042	
Educational achievement	-115.2	3.8	-111.3	5.4	3.9	6.0	
Percentage that attended college	-6.8	2.1	3.2	2.8	10.0	3.4	**
Percentage that completed college	-11.2	1.5	-4.6	2.1	6.5	2.5	**
Average years of work experience	-0.5	0.1	-0.5	0.1	0.0	0.1	
Percentage that are in the labor force	-1.1	1.1	13.5	2.6	14.6	2.7	**
Natural log of wage	-0.1	0.0	0.1	0.0	0.2	0.0	**
Percentage that have earnings	-2.6	0.8	2.7	2.6	5.3	2.7	**
Earnings in thousands	-3.4	0.5	-0.1	1.1	3.3	1.1	**

** Difference-of-means between two groups is statistically significant at $p \leq .05$.

NOTE: Standard errors have been adjusted for complex survey designs. Numbers may not sum to totals due to rounding. Sample sizes are smaller than those reported previously because observations with a sampling weight of zero were excluded from the tabulations. Multiple imputation used to account for missing data.

SOURCE: U.S. Department of Education, National Center for Education Statistics, High School and Beyond (HSB); and U.S. Department of Labor, Bureau of Labor Statistics, National Longitudinal Survey of Youth (NLSY).

Table A.4 Comparison of black and white young adults in NLSY, 7 years after grade 12, by presence of all variables

Sample definition and descriptive variables	Not missing any variables		Missing one or more variables		Difference-of-means (missing - not missing)		Statistical significance
	Mean	Std. Error	Mean	Std. Error	Value	Std. Error	
Population of black and white children	(N=5412)		(N=1458)				
Percentage black	12.2	0.0	18.1	0.1	5.9		1.2
Population of white young adults	(N=4027)		(N=976)				
Percentage female	48.8	0.9	51.2	0.1	2.4		2.7
Percentage attending high school in the Northeast	24.7	4.4	19.0	0.1	-5.7		2.1
Percentage attending high school in the Midwest	34.9	4.8	36.6	0.2	1.7		2.5
Percentage attending high school in the South	25.3	4.0	27.4	0.1	2.0		2.0
Percentage attending high school in the West	15.1	3.1	17.0	0.1	1.9		1.8
Socioeconomic status	0.113	0.032	0.020	0.048	-0.093		0.045
Educational achievement	0.2	0.0	0.0	0.0	-0.2		0.0
Percentage that attended college	58.2	1.4	47.6	0.1	-10.7		2.2
Percentage that completed college	26.8	1.3	12.5	0.0	-14.3		1.7
Average years of work experience	5.5	0.0	3.4	0.1	-2.1		0.1
Percentage that are in the labor force	94.4	0.4	44.3	2.4	-50.2		2.3
Natural log of wage	2.2	0.0	2.2	0.3	0.0		0.3
Percentage that have earnings	96.7	0.4	47.5	2.1	-49.2		2.2
Earnings in thousands	18.3	0.3	7.6	0.7	-10.7		0.7
Population of black young adults	(N=1385)		(N=482)				
Percentage female	51.6	1.5	50.6	2.8	-1.1		3.3
Percentage attending high school in the Northeast	16.9	3.9	15.1	3.1	-1.8		2.7
Percentage attending high school in the Midwest	17.0	3.8	22.2	4.7	5.3		2.7
Percentage attending high school in the South	60.9	5.3	56.2	5.4	-4.7		3.6
Percentage attending high school in the West	5.2	1.3	6.5	1.7	1.2		1.4
Socioeconomic status	-0.520	0.040	-0.738	0.040	-0.219		0.046
Educational achievement	-1.0	0.0	-1.1	0.1	-0.2		0.1
Percentage that attended college	52.0	1.9	42.3	2.8	-9.7		3.1
Percentage that completed college	13.9	1.1	4.7	1.2	-9.2		1.5
Average years of work experience	4.6	0.1	2.3	0.1	-2.3		0.1
Percentage that are in the labor force	94.7	0.6	40.8	2.9	-53.9		2.9
Natural log of wage	2.0	0.0	1.9	0.2	-0.1		0.2
Percentage that have earnings	95.9	0.6	31.8	2.8	-64.1		3.0
Earnings in thousands	13.7	0.4	3.5	0.5	-10.1		0.6
Difference between black and white children							
Percentage female	2.8	1.8	-0.6	3.7	-3.4		4.3
Percentage attending high school in the Northeast	-7.8	5.4	-3.9	4.4	3.9		3.1
Percentage attending high school in the Midwest	-17.9	5.5	-14.3	6.2	3.6		3.3
Percentage attending high school in the South	35.6	5.7	28.8	6.1	-6.8		3.9
Percentage attending high school in the West	-9.9	3.1	-10.5	3.9	-0.7		2.3
Socioeconomic status	-0.633	0.053	-0.758	0.061	-0.126		0.064
Educational achievement	-119.2	4.4	-115.6	7.3	3.6		7.7
Percentage that attended college	-6.3	2.3	-5.3	3.6	1.0		3.9
Percentage that completed college	-12.8	1.7	-7.7	1.9	5.1		2.3
Average years of work experience	-0.8	0.1	-1.1	0.2	-0.3		0.2
Percentage that are in the labor force	0.2	0.7	-3.5	3.6	-3.7		3.6
Natural log of wage	-0.2	0.0	-0.3	0.2	-0.2		0.2
Percentage that have earnings	-0.9	0.7	-15.7	3.5	-14.9		3.7
Earnings in thousands	-4.7	0.5	-4.1	0.9	0.6		1.0

** Difference-of-means between two groups is statistically significant at $p \leq .05$.

NOTE: Standard errors have been adjusted for complex survey designs. Numbers may not sum to totals due to rounding. Sample sizes are smaller than those reported previously because observations with a sampling weight of zero were excluded from the tabulations. Multiple imputation used to account for missing data.

SOURCE: U.S. Department of Education, National Center for Education Statistics, High School and Beyond (HSB); and U.S. Department of Labor, Bureau of Labor Statistics, National Longitudinal Survey of Youth (NLSY).

Table A.5 Comparison of black and white young adults in NLSY observed 12 years after grade 10, by presence of all variables

Sample definition and descriptive variables	Not missing any variables		Missing one or more variables		Difference-of-means (missing - not missing)		Statistical significance
	Mean	Std. Error	Mean	Std. Error	Value	Std. Error	
Population of black and white children	(N=6834)		(N=450)				
Percentage black	14.0	0.0	15.4	0.1	1.4	1.6	
Population of white young adults	(N=4866)		(N=315)				
Percentage female	49.6	0.8	44.9	0.2	-4.7	3.4	
Percentage attending high school in the Northeast	22.3	4.0	24.1	0.3	1.8	2.6	
Percentage attending high school in the Midwest	35.2	4.8	34.7	0.3	-0.5	3.7	
Percentage attending high school in the South	26.3	4.0	24.7	0.3	-1.6	3.4	
Percentage attending high school in the West	16.2	3.3	16.5	0.2	0.3	2.7	
Socioeconomic status	0.099	0.030	0.033	0.075	-0.066	0.077	
Educational achievement	0.2	0.0	-0.1	0.1	-0.3	0.1	**
Percentage that completed high school	87.0	0.7	82.8	0.1	-4.2	2.2	
Percentage that attended college	55.9	1.3	43.9	0.2	-11.9	3.5	**
Percentage that completed college	25.8	1.2	14.3	2.7	-11.5	2.6	**
Average years of work experience	7.9	0.1	7.4	0.2	-0.5	0.2	**
Percentage that are in the labor force	85.8	0.7	83.5	2.7	-2.3	2.7	
Percentage that have earnings	87.4	0.6	80.8	2.8	-6.7	2.8	**
Earnings in thousands	20.1	0.4	17.6	1.2	-2.5	1.2	**
Population of black young adults	(N=1968)		(N=135)				
Percentage female	49.7	1.2	47.8	4.6	-1.9	4.7	
Percentage attending high school in the Northeast	16.7	3.5	22.8	5.6	6.1	4.0	
Percentage attending high school in the Midwest	19.2	4.1	21.7	5.5	2.6	4.0	
Percentage attending high school in the South	57.8	5.1	49.5	6.7	-8.3	5.2	
Percentage attending high school in the West	6.2	1.4	5.9	2.4	-0.3	2.5	
Socioeconomic status	-0.561	0.032	-0.593	0.073	-0.033	0.076	
Educational achievement	-1.0	0.0	-1.1	0.1	-0.1	0.1	
Percentage that completed high school	79.0	1.0	72.9	4.2	-6.2	4.3	
Percentage that attended college	46.5	1.6	48.8	4.8	2.3	4.9	
Percentage that completed college	12.2	0.9	5.7	1.9	-6.5	2.0	**
Average years of work experience	6.3	0.1	5.2	0.3	-1.1	0.3	**
Percentage that are in the labor force	81.4	0.9	76.1	3.7	-5.3	3.7	
Percentage that have earnings	81.1	1.1	70.3	4.8	-10.8	4.7	**
Earnings in thousands	14.3	0.4	13.9	1.8	-0.4	1.8	
Difference between black and white children							
Percentage female	0.1	1.4	2.9	5.6	2.8	5.8	
Percentage attending high school in the Northeast	-5.6	4.8	-1.3	7.0	4.3	4.8	
Percentage attending high school in the Midwest	-16.0	5.6	-12.9	7.3	3.1	5.4	
Percentage attending high school in the South	31.6	5.6	24.8	7.5	-6.7	6.1	
Percentage attending high school in the West	-10.0	3.2	-10.6	4.6	-0.6	3.5	
Socioeconomic status	-0.660	0.044	-0.626	0.104	0.034	0.109	
Educational achievement	-1.2	0.0	-1.0	0.2	0.1	0.2	
Percentage that completed high school	-7.9	1.2	-9.9	4.9	-2.0	5.0	
Percentage that attended college	-9.4	2.0	4.8	6.1	14.2	6.1	**
Percentage that completed college	-13.6	1.5	-8.6	3.3	5.0	3.3	
Average years of work experience	-1.6	0.1	-2.2	0.4	-0.6	0.4	
Percentage that are in the labor force	-4.4	1.2	-7.4	4.5	-3.0	4.5	
Percentage that have earnings	-6.3	1.3	-10.4	5.5	-4.2	5.5	
Earnings in thousands	-5.8	0.6	-3.7	2.3	2.1	2.3	

** Difference-of-means between two groups is statistically significant at $p \leq .05$.

NOTE: Standard errors have been adjusted for complex survey designs. Numbers may not sum to totals due to rounding. Sample sizes are smaller than those reported previously because observations with a sampling weight of zero were excluded from the tabulations. Multiple imputation used to account for missing data.

SOURCE: U.S. Department of Education, National Center for Education Statistics, High School and Beyond (HSB); and U.S. Department of Labor, Bureau of Labor Statistics, National Longitudinal Study of Youth (NLSY).

Table A.6 Comparison of black and white young adults in HSB observed 12 years after grade 10, by presence of all variables

Sample definition and descriptive variables	Not missing any variables		Missing one or more variables		Difference-of-means (missing - not missing)		Statistical significance
	Mean	Std. Error	Mean	Std. Error	Value	Std. Error	
Percentage of black and white children	(N=6317)		(N=4612)				
Percentage black	10.9	0.0	20.5	0.0	9.6	1.1	**
Population of white young adults	(N=5360)		(N=3422)				
Percentage female	51.6	0.8	49.0	0.0	-2.6	1.5	
Percentage attending high school in the Northeast	23.6	1.8	24.3	0.0	0.7	1.5	
Percentage attending high school in the Midwest	34.7	2.0	27.6	0.0	-7.1	1.7	**
Percentage attending high school in the South	27.8	1.9	29.2	0.0	1.4	1.8	
Percentage attending high school in the West	13.9	1.4	18.8	0.0	5.0	1.2	**
Socioeconomic status	0.102	0.016	-0.016	0.024	-0.118	0.023	**
Educational achievement	0.4	0.0	-0.1	0.0	-0.4	0.0	**
Percentage that completed high school	99.5	0.1	89.1	0.0	-10.5	0.9	**
Percentage that attended college	68.0	0.9	48.9	0.0	-19.1	1.8	**
Percentage that completed college	33.4	1.0	17.4	1.1	-16.0	1.2	**
Average years of work experience	7.9	0.0	7.2	0.1	-0.8	0.1	**
Percentage that are in the labor force	86.3	0.6	80.8	1.0	-5.5	1.2	**
Percentage that have earnings	92.0	0.5	83.6	0.9	-8.4	1.0	**
Earnings in thousands	22.0	0.3	18.7	0.4	-3.3	0.5	**
Population of black young adults	(N=957)		(N=1190)				
Percentage female	57.3	2.2	49.4	2.1	-7.9	3.1	**
Percentage attending high school in the Northeast	14.9	2.2	26.6	3.6	11.8	2.9	**
Percentage attending high school in the Midwest	14.3	2.3	18.8	2.8	4.4	2.3	**
Percentage attending high school in the South	64.8	3.4	47.5	3.8	-17.3	3.2	**
Percentage attending high school in the West	6.0	1.4	7.2	1.7	1.1	1.7	
Socioeconomic status	-0.405	0.036	-0.496	0.050	-0.090	0.056	
Educational achievement	-0.6	0.0	-0.9	0.0	-0.3	0.1	**
Percentage that completed high school	98.9	0.5	88.8	1.6	-10.1	1.6	**
Percentage that attended college	59.7	2.3	42.1	2.6	-17.5	3.4	**
Percentage that completed college	18.2	1.7	7.8	1.1	-10.4	1.9	**
Average years of work experience	7.1	0.1	6.1	0.2	-1.0	0.2	**
Percentage that are in the labor force	84.2	1.7	77.6	2.3	-6.6	2.9	**
Percentage that have earnings	88.1	1.7	72.7	2.6	-15.4	3.1	**
Earnings in thousands	18.2	0.7	14.6	0.9	-3.6	1.1	**
Difference between black and white children							
Percentage female	5.7	2.4	0.4	2.4	-5.3	3.5	
Percentage attending high school in the Northeast	-8.7	2.6	2.3	3.8	11.1	3.2	**
Percentage attending high school in the Midwest	-20.4	2.8	-8.8	3.2	11.5	2.7	**
Percentage attending high school in the South	36.9	3.4	18.2	3.9	-18.7	3.5	**
Percentage attending high school in the West	-7.8	1.8	-11.7	2.2	-3.8	2.1	
Socioeconomic status	-0.507	0.038	-0.480	0.052	0.027	0.059	
Educational achievement	-0.9	0.0	-0.8	0.0	0.1	0.1	**
Percentage that completed high school	-0.6	0.5	-0.2	1.8	0.4	1.8	
Percentage that attended college	-8.3	2.5	-6.8	2.9	1.6	3.7	
Percentage that completed college	-15.2	1.9	-9.6	1.5	5.6	2.2	**
Average years of work experience	-0.9	0.1	-1.1	0.2	-0.2	0.2	
Percentage that are in the labor force	-2.1	1.9	-3.2	2.4	-1.1	3.0	
Percentage that have earnings	-3.9	1.8	-10.9	2.7	-7.0	3.3	**
Earnings in thousands	-3.8	0.7	-4.1	1.0	-0.3	1.2	

** Difference-of-means between two groups is statistically significant at $p \leq .05$.

NOTE: Standard errors have been adjusted for complex survey designs. Numbers may not sum to totals due to rounding. Sample sizes are smaller than those reported previously because observations with a sampling weight of zero were excluded from the tabulations. Multiple imputation used to account for missing data.

SOURCE: U.S. Department of Education, National Center for Education Statistics, High School and Beyond (HSB); and U.S. Department of Labor, Bureau of Labor Statistics, National Longitudinal Study of Youth (NLSY).

Table A.7 Comparison of black and white young adults in NLSY observed 12 years after grade 10, by presence of all variables (in 1992 only)

Sample definition and descriptive variables	Not missing any variables		Missing one or more variables		Difference-of-means (missing - not missing)		Statistical significance
	Mean	Std. Error	Mean	Std. Error	Value	Std. Error	
Population of black and white children	(N=948)		(N=50)				
Percentage black	16.0	0.1	12.2	0.5	-3.9	3.6	
Population of white young adults	(N=632)		(N=36)				
Percentage female	48.3	2.3	50.8	1.5	2.5	9.6	
Percentage attending high school in the Northeast	20.6	4.2	17.7	1.1	-2.8	6.2	
Percentage attending high school in the Midwest	34.3	5.2	43.0	1.8	8.7	9.4	
Percentage attending high school in the South	27.8	4.8	19.5	1.2	-8.2	7.2	
Percentage attending high school in the West	17.4	3.9	19.8	1.4	2.4	6.6	
Socioeconomic status	0.078	0.050	-0.163	0.191	-0.241	0.188	
Educational achievement	0.2	0.0	-0.3	0.2	-0.5	0.2	**
Percentage that completed high school	81.6	1.9	84.9	1.0	3.3	6.1	
Percentage that attended college	53.6	2.5	31.1	1.5	-22.5	8.9	**
Percentage that completed college	26.6	2.1	14.1	6.3	-12.5	6.5	**
Average years of work experience	8.6	0.1	9.4	0.4	0.7	0.4	
Percentage that are in the labor force	87.4	1.5	88.2	5.4	0.8	5.5	
Percentage that have earnings	86.8	1.4	85.1	6.8	-1.7	6.8	
Earnings in thousands	19.3	0.6	20.4	4.1	1.1	4.1	
Population of black young adults	(N=316)		(N=14)				
Percentage female	49.9	2.8	53.7	15.3	3.8	14.8	
Percentage attending high school in the Northeast	15.4	4.0	23.5	12.1	8.2	11.6	
Percentage attending high school in the Midwest	18.6	4.5	34.6	14.5	16.0	12.8	
Percentage attending high school in the South	60.0	5.9	36.8	15.2	-23.2	14.6	
Percentage attending high school in the West	6.0	2.3	5.0	5.2	-1.0	5.4	
Socioeconomic status	-0.548	0.053	-0.290	0.285	0.258	0.286	
Educational achievement	-0.9	0.1	-0.8	0.3	0.1	0.3	
Percentage that completed high school	71.7	2.9	75.5	13.5	3.8	13.1	
Percentage that attended college	37.6	3.3	46.1	16.1	8.5	15.3	
Percentage that completed college	10.7	2.2	13.4	9.6	2.7	9.7	
Average years of work experience	6.8	0.2	5.9	1.1	-0.9	1.1	
Percentage that are in the labor force	80.3	2.6	83.5	11.2	3.2	10.4	
Percentage that have earnings	78.7	2.3	69.7	14.3	-9.1	13.4	
Earnings in thousands	13.2	0.7	10.5	4.1	-2.8	4.0	
Difference between black and white children							
Percentage female	1.6	3.6	2.8	17.9	1.3	18.0	
Percentage attending high school in the Northeast	-5.2	5.4	5.8	13.5	11.0	13.2	
Percentage attending high school in the Midwest	-15.7	6.2	-8.3	16.9	7.3	15.3	
Percentage attending high school in the South	32.3	6.7	17.3	16.2	-15.0	16.1	
Percentage attending high school in the West	-11.4	4.1	-14.8	9.4	-3.4	8.2	
Socioeconomic status	-0.626	0.073	-0.127	0.336	0.500	0.340	
Educational achievement	-1.1	0.1	-0.5	0.4	0.6	0.4	
Percentage that completed high school	-9.9	3.4	-9.4	14.5	0.5	14.4	
Percentage that attended college	-16.0	4.0	15.0	18.7	31.0	18.3	
Percentage that completed college	-15.9	3.1	-0.7	11.2	15.2	11.6	
Average years of work experience	-1.8	0.2	-3.5	1.2	-1.7	1.2	
Percentage that are in the labor force	-7.1	3.0	-4.7	12.2	2.4	11.9	
Percentage that have earnings	-8.1	2.7	-15.4	15.6	-7.4	15.2	
Earnings in thousands	-6.0	1.0	-9.9	5.5	-3.9	5.5	

** Difference-of-means between two groups is statistically significant at $p \leq .05$.

NOTE: Standard errors have been adjusted for complex survey designs. Numbers may not sum to totals due to rounding. Sample sizes are smaller than those reported previously because observations with a sampling weight of zero were excluded from the tabulations. Multiple imputation used to account for missing data.

SOURCE: U.S. Department of Education, National Center for Education Statistics, High School and Beyond (HSB); and U.S. Department of Labor, Bureau of Labor Statistics, National Longitudinal Study of Youth (NLSY).

- Black-white differences for complete cases in the NLSY sophomore sample are similar to black-white differences for cases with missing values. The only significant discrepancy in this sample is that among complete cases, whites were more likely to complete college than blacks, while the opposite is true for cases with missing data.
- In HSB, black-white differences tended to be smaller for cases with missing data than for complete cases for the region, educational achievement, and educational attainment variables, and larger for the labor market variables.

When comparisons were restricted to only cases from 1992 in the NLSY senior sample (table A.7), differences between complete cases and cases with missing values showed no particular pattern. This is probably due to the fact that only 50 cases in this subset had missing data, so the differences could not be measured precisely. In this subsample, there were no statistically significant differences in black-white gaps between complete cases and cases with missing values.

These results suggest that eliminating cases with missing data from the analysis could bias estimates of black-white differences. Careful imputation of missing values allows use of all available observations and should minimize this problem.

III. DATA USED IN THE ANALYSES OF EDUCATIONAL ATTAINMENT OUTCOMES

A. Samples

Six samples were used to compare the educational attainment of black and white young adults. Two of these samples were constructed for the analyses described in appendix D. First, samples of 2,695 black and white (non-Hispanic) youth (16- to 24-year-olds) and 5,421 black and white (non-Hispanic) young adults (25- to 34-year-olds) obtained from the 1992 National Adult Literacy Survey (NALS) were used to analyze black-white differences for both the civilian, noninstitutional population and the population of incarcerated persons. Next, a sample of 8,306 black and white (non-Hispanic) 28-year-olds was defined using 1985 to 1992 data from the NLSY to analyze trends in the completion of high school diplomas and equivalents over time, and to include some institutionalized persons in the sample.

The four remaining samples, the same as those used in the analysis of labor market outcomes, were used in the analysis of black-white differences in educational attainment as a function of black-white differences in educational achievement.

B. Variable Definitions

The following educational attainment outcome variables were defined in each sample:

1. *High school completion status*—indicator for young adults who had obtained a high school diploma or equivalent (GED, or General Educational Development certificate)
2. *College attendance status*—indicator for young adults who had attended at least some college (defined for young adults with high school diplomas or GEDs only)
3. *College completion status*—indicator for young adults who had completed at least four years of college or an equivalent bachelor's degree (defined for college attendees only)

Of the various surveys, only the NLSY inquired about the GED completion status of all the individuals in the sample, including those who subsequently attended college.

The analyses of educational attainment outcomes used the following variables to describe the different backgrounds of young adults:

1. *Race*—indicator for black children (defined for non-Hispanic blacks and non-Hispanic whites only; Hispanic, Asian, or Pacific Islander individuals were excluded from the sample, as were persons of unknown race)
2. *Sex*—indicator for girls
3. *Census region of high school (Northeast, North Central, South, or West)*—defined as of grade 12 for the NLS:72 sample, as of age 14 (and on the basis of the child’s residence) for the NLSY samples, and as of grade 10 for the HSB sample
4. *Parental socioeconomic status (SES) and socioeconomic status squared*—a composite variable constructed using any non-missing measures of mother’s education, father’s education, mother’s occupational status (in the NLSY samples only), father’s occupational status, family income, and (in the NLS:72 and HSB samples only) the presence of certain household items. For the NLS:72 and HSB samples, the baseline SES measures described by NCES (1997b) were used. For each of the class years represented in the NLSY samples (1976 through 1982 for the NLSY senior sample, and 1974 through 1980 for the NLSY sophomore sample), SES was measured as follows. Each non-missing component was normalized to have mean zero and standard deviation one for youth regardless of race. These components were then averaged for each youth, and the averages were re-normalized the average to have mean zero and standard deviation one across the entire sample. Education levels were expressed by years of schooling, occupational status by the Duncan index, and family income by the natural log of annual income. SES was considered missing only if *all* of the components that could be used to construct the index were missing. These components are assumed to be highly correlated with each other, making it possible to compare SES values across observations and datasets even when specific components may be missing.
5. *Educational achievement and educational achievement squared*—an average of mathematics and reading achievement scores, with each component and the overall average normalized to have mean zero and standard deviation one for all individuals (regardless of race) in the same grade in the same year in each sample

C. Sample Sizes and Multiple Imputation of Missing Data

The analysis of black-white differences in educational attainment used the same four datasets that were used in the analysis of labor market outcomes. Sample sizes for these datasets are reported in tables A.1 and A.2 above.

Between 30 and 57 percent of cases have missing values for one or more key analysis variables. We addressed this problem through multiple imputation. Section I.D above describes the basic approach, and appendix C gives details on the calculation of estimates and standard errors using imputed data.

IV. DATA USED IN THE ANALYSES OF EDUCATIONAL ACHIEVEMENT OUTCOMES

A. Samples

Four sources of data were used in the analysis of black-white differences in educational achievement outcomes. These sources included Cohorts 1, 3, and 7 of the Chapter One Prospects Study, and the National Educational Longitudinal Study of 1988 (NELS:88).

The Prospects Study is a panel study of elementary and junior high school education that includes three cohorts distinguished by the grade of students during the base year (grades 1, 3, and 7). Data collection, including mathematics and reading achievement testing, began in 1991 and 1992 and continued through 1994. When the samples were defined, 1994 data were unavailable for each cohort, so the Cohort 1 sample only included data through Grade 2, the Cohort 3 sample only included data through Grade 5, and the Cohort 7 sample only included data through Grade 9. For the NELS:88 sample, data were available for grades 8, 10, and 12, in 1988, 1990, and 1992, respectively.

B. Variable Definitions

The following educational achievement outcome variables were defined in each sample:

1. *Mathematics achievement*—mathematics achievement test score, measured in grade 8 standard deviation units
2. *Reading achievement*—reading achievement test score, measured in grade 8 standard deviation units

Test scores were converted to grade 8 standard deviation units by subtracting the grade 8 mean for all children (regardless of race) from the raw test score, and dividing that difference by the grade 8 standard deviation for all children (regardless of race). Since no grade 8 scores were available for children in the Prospects Cohort 1 and Prospects Cohort 3 samples, means and standard deviations from the Prospects Cohort 7 sample were used.

The analyses of educational achievement outcomes used the following variables to describe the different backgrounds of young adults:

1. *Race*—indicator for black children (defined for non-Hispanic blacks and non-Hispanic whites only; Hispanic, Asian, or Pacific Islander individuals were excluded from the sample, as were persons of unknown race)
2. *Sex*—indicator for girls
3. *Census region of school (Northeast, North Central, South, or West)*—defined as of grades 1, 3, and 7 for the respective Prospects cohorts, and as of grade 8 for the NELS:88 sample
4. *Parental socioeconomic status (SES) and socioeconomic status squared*—a composite variable constructed using any non-missing measures of mother's education, father's education, mother's occupational status, father's occupational status, family income, and (in the NELS:88 sample only) the presence of certain household items. For the NELS:88, the baseline SES measure described by NCES (1995b) was used. For each of the Prospects samples, SES was measured by normalizing each non-missing component to have mean zero and standard deviation one across the entire sample, averaging these components for each child, and then re-

normalizing the average to have mean zero and standard deviation one across the entire sample. Education levels were expressed by years of schooling, occupational status by the Duncan index, and family income by the natural log of annual income. SES was considered missing only if *all* of the components that could be used to construct the index were missing. These components are assumed to be highly correlated with each other, making it possible to compare SES values across observations and datasets even when specific components may be missing.

5. *Prior educational achievement* and *educational achievement squared*—the corresponding math or reading achievement score, measured in grade 8 standard deviation units

C. Sample Sizes

Table A.8 presents sample sizes for the four analyses samples. Cohort 1 of the Prospects Study includes 13,757 observations, 60 percent (8,236) of which were included in the analysis of achievement outcomes between grades 1 and 2.¹³ About 40 percent of the sample (5,513) were not identified as either non-Hispanic black or non-Hispanic white (having either another or a missing indicator for race/ethnicity); eight additional cases had missing background information (sex or Census region). Of cases used in the analysis, 41 percent were missing key analysis variables (test scores or parental SES).¹⁴

Cohort 3 of the Prospects Study includes 19,311 observations, 61 percent (11,847) of which were included in the analysis of achievement outcomes between grades 3 and 5. About 35 percent of the sample (4,842) were not identified as either non-Hispanic black or non-Hispanic white (having either another or a missing indicator for race/ethnicity); an additional 4 percent (314) had missing sex or Census region. Of cases used in the analysis, 62 percent were missing one or more analysis variables.¹⁵

Cohort 7 of the Prospects Study includes 9,986 observations, 67 percent (6,655) of which were included in the analysis of achievement outcomes between grades 7 and 9. About 31 percent of the sample (3,069) were not identified as either non-Hispanic black or non-Hispanic white (having either another or a missing indicator for race/ethnicity); an additional 2 percent (262) had missing sex or Census region. Two-thirds of cases used in the analysis were missing one or more analysis variables.¹⁶

The NELS:88 sample consists of 27,588 observations, 55 percent (15,251) of which were included in the analysis of achievement outcomes between grades 8 and 12. About 40 percent of the sample (11,151) were not identified as either non-Hispanic black or non-Hispanic white (having either another or a missing indicator for race/ethnicity); around 4 percent (1,186) of the observations had missing sex or Census region. Of cases used in the analysis, 37 percent had missing values for one or more key variables.¹⁷

¹³ Both fall and spring scores were available for grade 1. The spring scores were used in the analyses, because the fall math scores covered conceptual skills only, and not computational skills.

¹⁴ Item nonresponse for the sample of blacks and whites in Prospects Cohort 1 ranged from 8.67 percent for the SES variable to 28.91 percent for the grade 2 math score.

¹⁴ Item nonresponse for the sample of blacks and whites in Prospects Cohort 3 ranged from 19.68 percent for the SES variable to 46.26 percent for the grade 5 math score.

¹⁵ Item nonresponse for the sample of blacks and whites in Prospects Cohort 7 ranged from 18.81 percent for the SES variable to 48.69 percent for the grade 9 math score.

¹⁶ Item nonresponse for the sample of blacks and whites in NELS:88 ranged from 4.93 percent for the SES variable to 26.90 percent for the grade 12 reading score.

Table A.8 Total sample sizes and excluded cases (achievement datasets)

	Prospects 1	Prospects 3	Prospects 7	NELS: 88
Total cases	13,757	19,311	9,986	27,588
Black	2,559	3,765	1,790	2,300
White	5,685	8,750	5,127	14,137
Other/not reported	5,513	6,796	3,069	11,151
Percent missing sex or region	1.10	6.25	6.36	29.34
Black	0.16	6.61	4.86	13.35
White	0.07	4.79	3.41	6.22
Other/not reported	2.59	7.92	12.15	61.95
Total included in analysis	8,236	11,847	6,655	15,251
Black	2,555	3,516	1,703	1,993
White	5,681	8,331	4,952	13,258
Percent with missing values	41.2	61.7	65.0	37.3
Black	43.8	64.2	74.3	47.0
White	40.0	60.7	61.8	35.8

D. Multiple Imputation of Missing Data

The restriction of each analysis sample to students with valid mathematics and reading scores and parental SES would reduce the number of observations included in the analyses of educational achievement outcomes by 40 to 65 percent. Preliminary comparisons revealed that children with missing test score data tended to have lower parental SES than children with valid test score data. While black-white differences in sex, Census region, and parental SES were similar for children with test score data present and children with test score data absent in the four samples, black-white differences in Census region were dissimilar between the two groups of children in the NELS:88 sample.

These differences suggest that simply excluding cases with missing test scores from the analysis may bias the estimation of equations predicting black-white differences in outcomes, yielding estimates that are only applicable to the population of students with valid SES and test score data. We address this problem through the use of multiple imputation, which allows the use of all available data for each case, even if some key variables such as test scores are missing. Multiple imputation is described in section I.D above, and in Schafer (1997).

The imputations used all observations from non-Hispanic blacks and non-Hispanic whites with known sex and Census region, and missing values were imputed for the remaining variables of interest. These variables were mathematics and reading achievement (initial, final, and grade 8, where available), and parental SES. To improve the efficiency of the imputation and capture school and regional effects, the sample weight, average SES by school and by race, and average initial math and reading achievement by school and by race were also included. Since all of these variables are continuous and approximately normally distributed, it was reasonable to assume a multivariate normal distribution for the variables in the imputation model. This assumption yields a simpler model than the general location model described above and used to impute missing data for mixed categorical/continuous variables. The software package NORM

Table A.9 Sample sizes by race/sex category (achievement datasets)

		Cohort 1	Cohort 3	Cohort 7	NELS:88
Total	all regions	8,236	11,847	6,655	15,251
White males	Northeast	448	814	475	1,407
	North Central	620	773	596	2,104
	South	1,179	1,650	1,007	2,179
	West	655	1,008	505	1,013
White females	Northeast	417	780	454	1,397
	North Central	626	806	582	2,088
	South	1,160	1,610	916	2,130
	West	576	890	417	940
Black males	Northeast	290	455	250	157
	North Central	104	145	86	149
	South	809	958	437	620
	West	84	185	63	60
Black females	Northeast	319	473	251	154
	North Central	103	155	111	145
	South	749	977	442	645
	West	97	168	63	63

Version 2.0, created by Joe Schafer, was used to create the imputed datasets under the multivariate normal assumption.

Since the relationships among test scores may differ among race, sex, and region categories, missing values were imputed separately for each of the 16 subsets defined by the cross-classification of these variables (2 sexes x 2 races x 4 regions). The sample sizes for these 16 groups are indicated in table A.9. For each of the 16 subsets of NELS:88 and Cohorts 1 and 3 of the Prospects Study, we ran a single series of 2,500 iterations of data augmentation, using the default starting values and standard noninformative priors. We checked this series for convergence using the diagnostic plots suggested by Schaefer (1997)¹⁸. We examined time series and autocorrelation function (ACF) plots for each parameter (means and covariances for all variables in the imputation model). For all parameters and all subsets of the data, the plots indicated convergence of the data augmentation algorithm after a few hundred iterations. Time series plots for each parameter resembled a horizontal band, and autocorrelations diminished to near zero after a few lags.

Once we were satisfied that the data augmentation series had converged in fewer than 500 steps, we imputed missing values after every 500th step, for a total of 5 imputations. The data augmentation routine appeared to take longer to converge on the Prospects Cohort 7 data files. For each of these data files we ran a series of 5,000 steps of data augmentation and imputed after every 1,000th step.

For the first of the five multiple imputations, we merged the 16 subsets together, creating a single file that included observations from all sex, race, and region categories. We repeated this step for each imputation, creating a total of five files for use in the analysis. In the case of the Prospects Cohort 7 and NELS:88 datasets, we also imputed miss-

¹⁸ Note that the method for assessing convergence is different than that described in section I.D. The software used for the general location model does not allow saving and plotting of series of parameters.

ing math and reading test scores for children of other races than black/non-Hispanic and white/non-Hispanic. These additional imputations allowed us to normalize, for the corresponding imputed datasets, all Prospects and NELS:88 test scores by the mean and standard deviation for the entire sample of 8th graders.

E. Characteristics of Cases with Missing Data

Tables A.10 through A.13 show summary statistics on the key variables in the analysis of educational achievement. Some key differences exist between the distributions of the data for complete cases and cases with missing values. Test scores for reading and math were lower among cases with missing values than among complete cases in each of the four datasets. Among white children, these differences are statistically significant. Differences among black children are similar in magnitude but are not statistically significant, probably due to the smaller sample sizes for blacks. Cases with missing data had lower average SES and were more likely to come from the bottom SES quartile. This difference holds for both blacks and whites in each dataset.

Despite the differences in characteristics of complete cases and cases with missing data, the estimated black-white differences were similar for the two groups. There were no significant differences in the Prospects samples. In the NELS:88 sample, black-white differences in the percentage attending high school in the Midwest and South Census regions were smaller for cases with missing data. These results suggest that restricting analysis to complete cases might bias estimates of black-white differences only slightly. Strictly speaking, however, results of an analysis restricted to complete cases could be generalized only to the population of students with valid data on test scores and SES. Multiple imputation of missing data avoids this problem.

Table A.10 Comparison of black and white children in Prospects Cohort 1, by presence of SES and test score data

Sample definition and descriptive variables	SES and test score data present		SES or test score data absent		Difference-of-means (absent - present)		Statistical significance
	Mean	Std. Error	Mean	Std. Error	Value	Std. Error	
Population of black and white children	(N=4474)		(N=3068)				
Percentage black	18.4	0.0	20.0	0.1	1.6	2.7	
Population of white children	(N=3205)		(N=2074)				
Percentage female	48.2	1.3	50.1	0.0	1.9	2.4	
Percentage attending high school in the Northeast	7.6	2.8	6.7	0.0	-1.0	2.3	
Percentage attending high school in the Midwest	29.2	7.0	26.7	0.2	-2.5	9.1	
Percentage attending high school in the South	45.8	7.6	41.0	0.2	-4.9	6.8	
Percentage attending high school in the West	17.3	5.5	25.6	0.2	8.3	6.6	
Socioeconomic status	0.222	0.060	0.076	0.065	-0.146	0.068	**
Percentage in bottom SES quartile	13.0	1.6	21.5	0.1	8.5	2.5	**
Percentage in lower-middle SES quartile	27.7	2.2	23.2	0.0	-4.5	2.2	**
Percentage in upper-middle SES quartile	30.6	1.7	25.7	0.0	-4.9	2.7	
Percentage in top SES quartile	28.7	3.0	29.6	0.1	0.9	3.4	
Mean grade 2 math score	-2.6	0.1	-2.8	0.1	-0.2	0.1	**
Mean grade 2 reading score	-2.0	0.1	-2.2	0.1	-0.2	0.1	**
Mean initial (grade 1) math score	-4.0	0.1	-4.3	0.1	-0.3	0.1	**
Mean initial (grade 1) reading score	-3.4	0.1	-3.6	0.1	-0.2	0.1	**
Population of black children	(N=1269)		(N=994)				
Percentage female	51.1	3.2	46.7	2.1	-4.4	2.7	
Percentage attending high school in the Northeast	16.9	7.5	14.5	5.2	-2.4	4.5	
Percentage attending high school in the Midwest	9.8	4.8	9.1	3.8	-0.7	2.7	
Percentage attending high school in the South	68.1	8.8	64.0	7.2	-4.1	5.5	
Percentage attending high school in the West	5.2	2.4	12.4	4.2	7.2	3.0	**
Socioeconomic status	-0.479	0.119	-0.681	0.084	-0.202	0.095	**
Percentage in bottom SES quartile	43.4	4.9	51.3	3.2	7.9	4.4	
Percentage in lower-middle SES quartile	21.9	2.1	21.4	2.0	-0.5	3.0	
Percentage in upper-middle SES quartile	22.7	3.1	15.0	1.7	-7.6	3.3	**
Percentage in top SES quartile	12.0	3.4	12.3	2.5	0.2	3.0	
Mean grade 2 math score	-3.4	0.1	-3.7	0.2	-0.3	0.1	**
Mean grade 2 reading score	-3.2	0.2	-3.5	0.2	-0.3	0.1	**
Mean initial (grade 1) math score	-5.2	0.3	-5.4	0.2	-0.2	0.2	
Mean initial (grade 1) reading score	-4.4	0.2	-4.6	0.2	-0.2	0.2	
Difference between black and white children							
Percentage female	2.9	3.6	-3.4	2.8	-6.4	3.9	
Percentage attending high school in the Northeast	9.3	7.6	7.8	5.4	-1.4	4.9	
Percentage attending high school in the Midwest	-19.4	7.1	-17.7	8.9	1.8	8.5	
Percentage attending high school in the South	22.3	9.8	23.0	8.8	0.7	7.8	
Percentage attending high school in the West	-12.1	5.0	-13.2	6.8	-1.1	7.0	
Socioeconomic status	-0.702	0.111	-0.758	0.096	-0.056	0.108	
Percentage in bottom SES quartile	30.4	4.7	29.8	3.8	-0.6	4.9	
Percentage in lower-middle SES quartile	-5.8	2.8	-1.8	2.5	4.0	3.5	
Percentage in upper-middle SES quartile	-7.9	3.2	-10.6	2.5	-2.7	4.2	
Percentage in top SES quartile	-16.7	3.3	-17.4	3.8	-0.7	4.2	
Mean grade 2 math score	-0.8	0.1	-0.9	0.1	-0.1	0.2	
Mean grade 2 reading score	-1.2	0.2	-1.3	0.2	-0.1	0.2	
Mean initial (grade 1) math score	-1.1	0.2	-1.0	0.1	0.1	0.2	
Mean initial (grade 1) reading score	-1.0	0.2	-0.9	0.2	0.0	0.2	

** Difference-of-means between two groups is statistically significant at $p \leq .05$.

NOTE: Standard errors have been adjusted for complex survey designs. Numbers may not sum to totals due to rounding. Sample sizes are smaller than those reported in table A.2 because observations with a sampling weight of zero were excluded from the tabulations. Multiple imputation used to account for missing data.

SOURCE: U.S. Department of Education, Chapter 1 Prospects Study.

Table A.11 Comparison of black and white children in Prospects Cohort 3, by presence of SES and test score data

Sample definition and descriptive variables	SES and test score data present		SES or test score data absent		Difference-of-means (absent - present)		Statistical significance
	Mean	Std. Error	Mean	Std. Error	Value	Std. Error	
Population of black and white children	(N=4224)		(N=6793)				
Percentage black	15.7	2.5	16.0	2.2	0.3	2.1	
Population of white children	(N=3099)		(N=4773)				
Percentage female	51.1	1.9	48.7	1.0	-2.4	2.3	
Percentage attending high school in the Northeast	17.8	6.2	15.0	3.7	-2.9	5.6	
Percentage attending high school in the Midwest	27.5	6.6	25.7	5.9	-1.7	6.5	
Percentage attending high school in the South	38.0	7.2	39.9	6.5	1.9	6.9	
Percentage attending high school in the West	16.7	5.2	19.4	5.0	2.7	5.3	
Socioeconomic status	0.205	0.066	0.108	0.060	-0.097	0.073	
Percentage in bottom SES quartile	14.7	1.7	21.1	1.7	6.4	2.0	**
Percentage in lower-middle SES quartile	27.4	2.4	24.4	1.6	-3.0	2.5	
Percentage in upper-middle SES quartile	30.6	1.6	24.9	1.2	-5.7	2.0	**
Percentage in top SES quartile	27.4	3.5	29.7	3.0	2.3	3.9	
Mean grade 5 math score	-0.5	0.0	-0.7	0.0	-0.2	0.0	**
Mean grade 5 reading score	-0.4	0.0	-0.6	0.1	-0.2	0.1	**
Mean initial (grade 3) math score	-1.3	0.0	-1.4	0.0	-0.1	0.0	**
Mean initial (grade 3) reading score	-1.0	0.0	-1.2	0.0	-0.2	0.1	**
Population of black children	(N=1125)		(N=2020)				
Percentage female	52.8	2.0	48.6	1.9	-4.2	3.0	
Percentage attending high school in the Northeast	24.6	8.6	20.9	5.9	-3.7	4.3	
Percentage attending high school in the Midwest	6.9	3.7	12.5	4.1	5.5	2.9	**
Percentage attending high school in the South	61.6	8.4	57.9	6.6	-3.7	5.4	
Percentage attending high school in the West	6.9	2.3	8.8	2.6	1.9	2.1	
Socioeconomic status	-0.419	0.109	-0.507	0.069	-0.088	0.071	
Percentage in bottom SES quartile	37.1	4.5	44.4	2.7	7.3	3.4	**
Percentage in lower-middle SES quartile	28.5	2.5	21.7	1.8	-6.9	3.0	**
Percentage in upper-middle SES quartile	22.2	2.6	20.5	2.3	-1.7	2.1	
Percentage in top SES quartile	12.1	3.4	13.4	1.7	1.3	3.0	
Mean grade 5 math score	-1.0	0.1	-1.2	0.1	-0.3	0.1	**
Mean grade 5 reading score	-1.3	0.1	-1.5	0.1	-0.2	0.1	
Mean initial (grade 3) math score	-1.8	0.1	-2.0	0.1	-0.2	0.1	**
Mean initial (grade 3) reading score	-1.7	0.1	-2.0	0.1	-0.2	0.1	**
Difference between black and white children							
Percentage female	1.7	3.2	0.0	2.3	-1.7	4.4	
Percentage attending high school in the Northeast	6.8	9.4	5.9	6.5	-0.9	6.1	
Percentage attending high school in the Midwest	-20.5	7.2	-13.3	6.5	7.3	7.1	
Percentage attending high school in the South	23.6	9.3	18.0	8.0	-5.6	7.4	
Percentage attending high school in the West	-9.8	4.7	-10.6	5.2	-0.8	5.1	
Socioeconomic status	-0.624	0.112	-0.615	0.082	0.010	0.087	
Percentage in bottom SES quartile	22.5	4.5	23.3	3.0	0.9	3.5	
Percentage in lower-middle SES quartile	1.2	3.3	-2.7	2.5	-3.9	3.9	
Percentage in upper-middle SES quartile	-8.4	3.2	-4.4	2.5	4.0	3.0	
Percentage in top SES quartile	-15.3	4.3	-16.3	3.2	-1.0	4.4	
Mean grade 5 math score	-0.4	0.1	-0.5	0.1	-0.1	0.1	
Mean grade 5 reading score	-0.9	0.1	-0.9	0.1	0.0	0.1	
Mean initial (grade 3) math score	-0.5	0.1	-0.6	0.1	0.0	0.1	
Mean initial (grade 3) reading score	-0.7	0.1	-0.8	0.1	0.0	0.1	

** Difference-of-means between two groups is statistically significant at $p \leq .05$.

NOTE: Standard errors have been adjusted for complex survey designs. Numbers may not sum to totals due to rounding. Sample sizes are smaller than those reported in table A.2 because observations with a sampling weight of zero were excluded from the tabulations. Multiple imputation used to account for missing data.

SOURCE: U.S. Department of Education, Chapter 1 Prospects Study.

Table A.12 Comparison of black and white children in Prospects Cohort 7, by presence of SES and test score data

Sample definition and descriptive variables	SES and test score data present		SES or test score data absent		Difference-of-means (absent - present)		Statistical-significance
	Mean	Std. Error	Mean	Std. Error	Value	Std. Error	
Population of black and white children	(N=2332)		(N=4322)				
Percentage black	12.5	2.5	19.3	2.9	6.8	2.1	**
Population of white children	(N=1894)		(N=3057)				
Percentage female	50.3	1.5	46.6	1.4	-3.8	2.0	
Percentage attending high school in the Northeast	16.6	5.9	12.6	3.9	-3.9	4.2	
Percentage attending high school in the Midwest	29.7	8.2	26.2	6.1	-3.5	6.5	
Percentage attending high school in the South	37.7	7.8	42.0	6.5	4.3	6.1	
Percentage attending high school in the West	16.0	5.4	19.1	4.9	3.1	4.2	
Socioeconomic status	0.233	0.056	0.108	0.059	-0.124	0.064	**
Percentage in bottom SES quartile	13.3	1.5	21.1	1.8	7.8	1.9	**
Percentage in lower-middle SES quartile	28.2	1.7	24.9	1.4	-3.3	2.0	
Percentage in upper-middle SES quartile	30.5	1.7	25.4	1.7	-5.1	2.4	**
Percentage in top SES quartile	28.0	2.9	28.6	2.9	0.6	3.5	
Mean grade 9 math score	0.4	0.1	0.2	0.0	-0.3	0.1	**
Mean grade 9 reading score	0.4	0.1	0.2	0.0	-0.3	0.1	**
Mean initial (grade 7) math score	0.1	0.0	-0.1	0.0	-0.2	0.0	**
Mean initial (grade 7) reading score	0.2	0.0	0.0	0.0	-0.2	0.0	**
Population of black children	(N=438)		(N=1265)				
Percentage female	49.1	4.2	52.7	1.9	3.6	4.8	
Percentage attending high school in the Northeast	13.3	5.9	14.3	4.8	1.0	3.6	
Percentage attending high school in the Midwest	18.6	9.8	16.6	6.8	-2.0	6.0	
Percentage attending high school in the South	61.0	10.3	56.1	8.3	-4.9	7.2	
Percentage attending high school in the West	7.1	3.1	13.1	4.6	6.0	2.8	**
Socioeconomic status	-0.351	0.102	-0.524	0.101	-0.173	0.115	
Percentage in bottom SES quartile	35.1	4.6	45.7	3.8	10.5	5.0	**
Percentage in lower-middle SES quartile	25.9	3.1	19.2	1.8	-6.7	3.0	**
Percentage in upper-middle SES quartile	24.8	3.4	19.1	2.7	-5.6	4.3	
Percentage in top SES quartile	14.2	2.9	16.0	2.6	1.8	3.4	
Mean grade 9 math score	-0.3	0.1	-0.5	0.1	-0.2	0.1	**
Mean grade 9 reading score	-0.4	0.1	-0.6	0.1	-0.2	0.1	
Mean initial (grade 7) math score	-0.4	0.1	-0.6	0.1	-0.2	0.1	
Mean initial (grade 7) reading score	-0.4	0.1	-0.7	0.1	-0.3	0.1	**
Difference between black and white children							
Percentage female	-1.2	4.3	6.1	2.4	7.4	5.2	
Percentage attending high school in the Northeast	-3.3	7.4	1.6	5.8	4.9	4.8	
Percentage attending high school in the Midwest	-11.1	9.4	-9.6	8.1	1.4	6.4	
Percentage attending high school in the South	23.3	10.7	14.1	8.9	-9.2	7.7	
Percentage attending high school in the West	-8.9	5.5	-6.0	5.6	2.9	4.9	
Socioeconomic status	-0.584	0.108	-0.633	0.110	-0.048	0.122	
Percentage in bottom SES quartile	21.8	4.6	24.5	4.1	2.7	5.0	
Percentage in lower-middle SES quartile	-2.3	3.5	-5.7	2.2	-3.4	3.6	
Percentage in upper-middle SES quartile	-5.7	3.9	-6.2	3.2	-0.5	4.9	
Percentage in top SES quartile	-13.8	3.7	-12.6	3.7	1.2	4.5	
Mean grade 9 math score	-0.7	0.1	-0.7	0.1	0.0	0.1	
Mean grade 9 reading score	-0.8	0.1	-0.8	0.1	0.1	0.1	
Mean initial (grade 7) math score	-0.5	0.1	-0.5	0.1	0.0	0.1	
Mean initial (grade 7) reading score	-0.6	0.1	-0.7	0.1	-0.1	0.1	

** Difference-of-means between two groups is statistically significant at $p < .05$.

NOTE: Standard errors have been adjusted for complex survey designs. Numbers may not sum to totals due to rounding. Sample sizes are smaller than those reported in table A.2 because observations with a sampling weight of zero were excluded from the tabulations. Multiple imputation used to account for missing data.

SOURCE: U.S. Department of Education, Chapter 1 Prospects Study.

Table A.13 Comparison of black and white children in NELS:88, by presence of SES and test score data

Sample definition and descriptive variables	SES and test score data present		SES or test score data absent		Difference-of-means (absent - present)		Statistical significance
	Mean	Std. Error	Mean	Std. Error	Value	Std. Error	
Population of black and white children	(N=9564)		(N=3390)				
Percentage black	12.5	0.9	20.7	2.0	8.2	1.8	**
Population of white children	(N=8508)		(N=2878)				
Percentage female	49.6	0.9	50.8	1.5	1.1	1.8	
Percentage attending high school in the Northeast	21.4	2.0	20.1	2.0	-1.3	1.6	
Percentage attending high school in the Midwest	33.2	2.1	26.7	2.1	-6.5	1.9	**
Percentage attending high school in the South	30.5	1.9	32.9	2.1	2.4	1.9	
Percentage attending high school in the West	14.9	1.4	20.3	2.1	5.4	1.7	**
Socioeconomic status	0.12	0.02	0.02	0.03	-0.10	0.03	**
Percentage in bottom SES quartile	15.7	0.8	20.4	1.5	4.7	1.5	**
Percentage in lower-middle SES quartile	24.6	0.8	26.1	1.3	1.5	1.4	
Percentage in upper-middle SES quartile	28.3	0.8	25.9	1.3	-2.3	1.5	
Percentage in top SES quartile	31.5	1.2	27.6	1.5	-3.9	1.6	**
Mean grade 12 math score	1.32	0.03	0.90	0.04	-0.42	0.04	**
Mean grade 12 reading score	0.95	0.02	0.62	0.05	-0.33	0.05	**
Mean initial (grade 8) math score	0.31	0.02	-0.01	0.03	-0.32	0.03	**
Mean initial (grade 8) reading score	0.28	0.02	-0.02	0.03	-0.30	0.03	**
Population of black children	(N=1056)		(N=512)				
Percentage female	53.3	2.2	49.5	3.9	-3.8	4.4	
Percentage attending high school in the Northeast	13.1	2.3	19.3	4.3	6.1	3.7	
Percentage attending high school in the Midwest	12.8	2.3	13.3	2.7	0.5	2.7	
Percentage attending high school in the South	68.4	3.3	59.6	5.2	-8.8	5.0	
Percentage attending high school in the West	5.6	1.3	7.7	2.3	2.1	2.3	
Socioeconomic status	-0.37	0.04	-0.41	0.07	-0.04	0.07	
Percentage in bottom SES quartile	37.8	2.3	41.2	4.2	3.5	4.6	
Percentage in lower-middle SES quartile	28.3	2.5	24.0	3.2	-4.4	4.0	
Percentage in upper-middle SES quartile	20.6	1.8	22.4	3.6	1.9	4.1	
Percentage in top SES quartile	13.3	1.8	12.4	3.9	-1.0	4.1	
Mean grade 12 math score	0.42	0.06	0.18	0.12	-0.24	0.13	
Mean grade 12 reading score	0.21	0.05	-0.09	0.14	-0.30	0.14	**
Mean initial (grade 8) math score	-0.46	0.04	-0.64	0.08	-0.19	0.08	**
Mean initial (grade 8) reading score	-0.35	0.04	-0.61	0.09	-0.26	0.09	**
Difference between black and white children							
Percentage female	3.7	2.3	-1.3	4.1	-4.9	4.8	
Percentage attending high school in the Northeast	-8.3	2.9	-0.8	4.7	7.4	4.0	
Percentage attending high school in the Midwest	-20.4	2.8	-13.3	3.3	7.1	3.2	**
Percentage attending high school in the South	37.9	3.5	26.7	5.5	-11.1	5.3	**
Percentage attending high school in the West	-9.3	1.8	-12.6	3.0	-3.3	2.8	
Socioeconomic status	-0.49	0.04	-0.43	0.08	0.06	0.08	
Percentage in bottom SES quartile	22.1	2.4	20.9	4.5	-1.2	4.8	
Percentage in lower-middle SES quartile	3.8	2.6	-2.2	3.4	-5.9	4.2	
Percentage in upper-middle SES quartile	-7.7	1.9	-3.5	3.8	4.2	4.4	
Percentage in top SES quartile	-18.2	2.1	-15.2	4.1	2.9	4.4	
Mean grade 12 math score	-0.90	0.06	-0.72	0.12	0.18	0.13	
Mean grade 12 reading score	-0.74	0.06	-0.71	0.15	0.03	0.16	
Mean initial (grade 8) math score	-0.76	0.05	-0.64	0.08	0.13	0.09	
Mean initial (grade 8) reading score	-0.63	0.04	-0.59	0.10	0.04	0.10	

** Difference-of-means between two groups is statistically significant at $p \leq .05$.

NOTE: Standard errors have been adjusted for complex survey designs. Numbers may not sum to totals due to rounding. Sample sizes are smaller than those reported in table A.2 because observations with a sampling weight of zero were excluded from the tabulations. Multiple imputation used to account for missing data.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Educational Longitudinal Study of 1988 (NELS:88).

APPENDIX B: SAMPLE COMPARISONS

Several samples were constructed in the study of black-white differences in educational and economic outcomes. While efforts were made to construct comparable samples from different data sources, there are systematic differences among the samples. This appendix compares the samples used in analyses of educational achievement, labor market and attainment outcomes, and describes differences among them. In all instances, the comparisons rely on both reported and (multiple) imputed data to infer the characteristics of each sample of individuals.

I. COMPARISONS OF SAMPLES USED IN ANALYSES OF LABOR MARKET AND ATTAINMENT OUTCOMES

Comparison of the core background characteristics and levels of educational attainment and work experience in the NLS:72 and NLSY senior samples indicates some differences between the two samples of young adults (table B.1):

- Blacks represented a larger (weighted) proportion of the NLSY sample than the NLS:72 sample.
- Compared with white young adults in the NLS:72 sample, white young adults in the NLSY sample were less likely to have attended or completed college, had more work experience, and had lower wages and lower earnings.
- Compared with black young adults in the NLS:72 sample, black young adults in the NLSY sample were more likely to be male, were less likely to have attended or completed college, had less work experience, were less likely to participate in the labor force, and had lower wages and earnings.

The two samples had similar black-white differences in census region, educational achievement, SES, and college attendance and completion. The two samples had dissimilar black-white differences in sex, work experience, labor force participation, wages, and earnings.

While black young adults represented about the same proportion of the NLSY and HSB sophomore samples, several differences were apparent between these two samples (table B.2):

- Compared with white young adults in the NLSY sample, white young adults in the HSB sample were more likely to have completed high school, attended college and completed college, and had higher earnings.
- Compared with black young adults in the NLSY sample, black young adults in the HSB sample had higher parental SES, had higher levels of educational achievement, were more likely to have completed high school, had more work experience, and had higher earnings.

Table B.1 Comparison of black and white young adults in two datasets observed 7 years after grade 12

Sample definition and descriptive variables	Observed in 1979 (from NLS:72)		Observed in 1983-89 (from NLSY)		Difference-of-means (NLSY - NLS:72)		Statistical significance
	Mean	Std. Error	Mean	Std. Error	Value	Std. Error	
Population of black and white children	(N=12,577)		(N=7424)				
Percentage black	9.8	0.0	13.1	0.0	3.3	1.7	**
Population of white young adults	(N=10,978)		(N=5426)				
Percentage female	49.2	0.7	49.4	0.0	0.2	1.0	
Percentage attending high school in the Northeast	25.9	1.7	23.7	0.1	-2.2	4.6	
Percentage attending high school in the Midwest	33.4	1.8	35.1	0.1	1.7	5.1	
Percentage attending high school in the South	24.2	1.5	25.7	0.1	1.5	4.2	
Percentage attending high school in the West	16.5	1.4	15.5	0.0	-1.0	3.4	
Socioeconomic status	0.092	0.014	0.095	0.031	0.003	0.034	
Educational achievement	0.2	0.0	0.2	0.0	0.0	0.0	
Percentage that attended college	62.9	0.7	55.7	0.0	-7.2	1.5	**
Percentage that completed college	27.1	0.7	23.5	0.0	-3.6	1.3	**
Average years of work experience	4.8	0.0	5.0	0.1	0.2	0.1	**
Percentage that are in the labor force	84.2	0.4	84.8	0.7	0.6	0.8	
Natural log of wage	2.4	0.0	2.2	0.1	-0.2	0.1	**
Percentage that have earnings	89.4	0.4	87.2	0.6	-2.2	0.7	**
Earnings	19,052	178	16,261	356	-2,791	395	**
Population of black young adults	(N=1599)		(N=1998)				
Percentage female	57.6	1.5	52.1	1.2	-5.5	1.9	**
Percentage attending high school in the Northeast	15.0	2.3	17.4	3.7	2.5	4.4	
Percentage attending high school in the Midwest	17.6	3.0	18.2	3.9	0.6	4.9	
Percentage attending high school in the South	61.0	3.4	58.8	5.2	-2.2	6.2	
Percentage attending high school in the West	6.5	1.3	5.5	1.3	-0.9	1.8	
Socioeconomic status	-0.585	0.023	-0.557	0.035	0.028	0.042	
Educational achievement	-1.0	0.0	-1.0	0.0	0.0	0.0	
Percentage that attended college	58.7	1.6	49.5	1.6	-9.2	2.3	**
Percentage that completed college	17.6	1.1	11.6	0.9	-6.0	1.4	**
Average years of work experience	4.3	0.1	4.0	0.1	-0.3	0.1	**
Percentage that are in the labor force	86.1	1.0	79.2	1.3	-6.8	1.6	**
Natural log of wage	2.3	0.0	2.0	0.1	-0.3	0.1	**
Percentage that have earnings	87.1	1.0	78.5	1.3	-8.6	1.6	**
Earnings	16,307	439	10,822	367	-5,485	580	**
Difference between black and white children							
Percentage female	8.4	1.6	2.7	1.4	-5.7	2.1	**
Percentage attending high school in the Northeast	-11.0	2.6	-6.3	5.1	4.7	5.7	
Percentage attending high school in the Midwest	-15.8	3.3	-16.9	5.5	-1.1	6.5	
Percentage attending high school in the South	36.8	3.4	33.1	5.7	-3.7	6.6	
Percentage attending high school in the West	-10.0	1.8	-9.9	3.1	0.1	3.6	
Socioeconomic status	-0.677	0.026	-0.652	0.047	0.025	0.054	
Educational achievement	-115.0	3.4	-118.4	4.2	-3.4	5.4	
Percentage that attended college	-4.2	1.8	-6.2	2.0	-2.0	2.7	
Percentage that completed college	-9.6	1.3	-11.9	1.4	-2.3	1.9	
Average years of work experience	-0.5	0.1	-1.0	0.1	-0.5	0.1	**
Percentage that are in the labor force	1.8	1.1	-5.6	1.4	-7.4	1.8	**
Natural log of wage	-0.1	0.0	-0.2	0.0	-0.1	0.0	**
Percentage that have earnings	-2.3	1.1	-8.7	1.4	-6.4	1.7	**
Earnings	-2,745	475	-5,439	511	-2,694	714	**

** Difference-of-means between two groups is statistically significant at $p \leq .05$.

NOTE: Standard errors have been adjusted for complex survey designs. Numbers may not sum to totals due to rounding. Sample sizes are smaller than those reported previously because observations with a sampling weight of zero were excluded from the tabulations. Multiple imputation used to account for missing data.

SOURCE: U.S. Department of Education, National Center for Education Statistics, High School and Beyond (HSB); and U.S. Department of Labor, Bureau of Labor Statistics, National Longitudinal Survey of Youth (NLSY).

Table B.2 Comparison of black and white young adults in two datasets observed 12 years after grade 10

Sample definition and descriptive variables	Observed in 1986-92 (from NLSY)		Observed in 1992 (from HSB)		Difference-of-means (HSB - NLSY)		Statistical significance
	Mean	Std. Error	Mean	Std. Error	Value	Std. Error	
Population of black and white children	(N=8998)		(N=10,929)				
Percentage black	13.8	0.0	15.0	0.0	1.2	1.8	
Population of white young adults	(N=6554)		(N=8782)				
Percentage female	49.0	0.7	50.5	0.0	1.5	0.9	
Percentage attending high school in the Northeast	23.0	4.0	23.9	0.0	0.9	4.4	
Percentage attending high school in the Midwest	34.6	4.7	31.8	0.0	-2.8	5.0	
Percentage attending high school in the South	26.1	3.9	28.4	0.0	2.3	4.2	
Percentage attending high school in the West	16.2	3.2	15.9	0.0	-0.4	3.5	
Socioeconomic status	0.100	0.029	0.054	0.016	-0.046	0.033	
Educational achievement	0.2	0.0	0.2	0.0	0.0	0.0	
Percentage that completed high school	85.8	0.7	95.3	0.0	9.5	0.8	**
Percentage that attended college	53.9	1.2	60.2	0.0	6.3	1.5	**
Percentage that completed college	24.0	1.1	26.9	0.8	2.9	1.4	**
Average years of work experience	7.5	0.1	7.6	0.0	0.1	0.1	
Percentage that are in the labor force	84.2	0.7	84.0	0.5	-0.1	0.9	
Percentage that have earnings	82.5	0.7	88.6	0.5	6.1	0.9	**
Earnings	18,690	395	20,668	255	1,978	474	**
Population of black young adults	(N=2444)		(N=2147)				
Percentage female	50.1	1.1	52.6	1.6	2.5	1.9	
Percentage attending high school in the Northeast	18.1	3.8	21.8	2.8	3.8	4.8	
Percentage attending high school in the Midwest	19.3	4.1	17.0	2.4	-2.3	4.8	
Percentage attending high school in the South	56.7	5.2	54.5	3.4	-2.2	6.2	
Percentage attending high school in the West	5.9	1.4	6.7	1.3	0.8	1.9	
Socioeconomic status	-0.552	0.031	-0.459	0.035	0.093	0.047	**
Educational achievement	-1.0	0.0	-0.7	0.0	0.2	0.0	**
Percentage that completed high school	77.4	1.0	92.9	1.0	15.6	1.4	**
Percentage that attended college	45.8	1.5	49.3	1.9	3.5	2.4	
Percentage that completed college	11.9	0.9	12.0	1.0	0.1	1.3	
Average years of work experience	6.0	0.1	6.5	0.1	0.5	0.2	**
Percentage that are in the labor force	81.7	0.9	80.3	1.5	-1.4	1.7	
Percentage that have earnings	75.1	1.2	79.0	1.7	3.9	2.1	
Earnings	13,188	385	16,071	609	2,883	742	**
Difference between black and white children							
Percentage female	1.1	1.3	2.1	1.7	1.0	2.1	
Percentage attending high school in the Northeast	-4.9	5.0	-2.1	3.0	2.9	5.8	
Percentage attending high school in the Midwest	-15.3	5.5	-14.9	2.7	0.5	6.2	
Percentage attending high school in the South	30.6	5.7	26.1	3.3	-4.5	6.6	
Percentage attending high school in the West	-10.3	3.2	-9.2	1.7	1.1	3.6	
Socioeconomic status	-0.651	0.042	-0.513	0.036	0.138	0.056	**
Educational achievement	-1.1	0.0	-0.9	0.0	0.2	0.1	**
Percentage that completed high school	-8.5	1.2	-2.3	1.0	6.1	1.6	**
Percentage that attended college	-8.1	1.9	-10.9	2.0	-2.8	2.8	
Percentage that completed college	-12.2	1.4	-14.9	1.3	-2.7	1.9	
Average years of work experience	-1.5	0.1	-1.1	0.1	0.4	0.2	**
Percentage that are in the labor force	-2.5	1.1	-3.8	1.6	-1.3	2.0	
Percentage that have earnings	-7.4	1.3	-9.6	1.8	-2.2	2.3	
Earnings	-5,501	563	-4,597	661	905	909	

** Difference-of-means between two groups is statistically significant at $p \leq .05$.

NOTE: Standard errors have been adjusted for complex survey designs. Numbers may not sum to totals due to rounding. Sample sizes are smaller than those reported previously because observations with a sampling weight of zero were excluded from the tabulations. Multiple imputation used to account for missing data.

SOURCE: U.S. Department of Education, National Center for Education Statistics, High School and Beyond (HSB); and U.S. Department of Labor, Bureau of Labor Statistics, National Longitudinal Study of Youth (NLSY).

Black-white differences in background characteristics were similar between the two samples in the categories of sex, census region, college attendance and completion, and earnings. Black-white differences in background characteristics were dissimilar between the two samples in the categories of parental SES, educational achievement, high school completion, and work experience.

Since both the NLSY sophomore sample and the HSB sophomore sample contain observations from the year 1992, it is possible to compare the 1992 portion of the NLSY sample with the HSB sample. In general, the differences between the NLSY subsample and the HSB sample resemble the differences between the entire NLSY sample and the HSB sample (table B.3):

- Compared with white young adults in the NLSY 1992 sample, white young adults in the HSB sample were more likely to have completed high school and attended college, had less work experience, and had higher earnings.
- Compared with black young adults in the NLSY 1992 samples, black young adults in the HSB sample had higher levels of educational achievement, were more likely to have completed high school and attended college, and had higher earnings.

Black-white differences in background characteristics were similar between the two samples in the categories of sex, census region, parental SES, college attendance and completion, and earnings. Black-white differences in background characteristics were dissimilar between the two samples in the categories of educational achievement, high school completion, and work experience.

These findings suggest that there were systematic differences between the NLSY and HSB sophomore samples, even for young adults sampled during the same year. These differences could be due to differences in sample design or in the wording of survey questions. Consequently, differences between the NLSY and HSB sophomore samples cannot be attributed to time trends alone. In particular, since blacks appear to have been less disadvantaged in the HSB sample than in the NLSY sample in 1992, comparisons of the entire NLSY sample with the HSB sample may show relative gains for blacks between the samples, even though such gains may not have occurred between 1986-1992 and 1992.

Comparisons between NLS:72 and the NLSY senior sample, and HSB and the NLSY sophomore sample both used multiple imputation to account for missing labor market and attainment outcomes.

II. COMPARISONS OF SAMPLES USED IN ANALYSES OF EDUCATIONAL ACHIEVEMENT OUTCOMES

The across-sample comparisons of black-white differences in educational achievement compared children across the Prospects and NELS:88 samples. The expression of test scores in grade 8 standard deviation units, and the “vertical equation” of the Prospects scores across cohorts, enabled these comparisons. With few exceptions, the background characteristics of children were similar in all of the Prospects samples and in the Prospects and NELS:88 samples.

The Prospects Cohort 1 sample and the Prospects Cohort 3 sample were similar in terms of the proportion of the sample that was black. Black and white children in the

Table B.3 Comparison of black and white young adults in two datasets observed 12 years after grade 10 (in 1992 only)

Sample definition and descriptive variables	Observed in 1992 (from NLSY)		Observed in 1992 (from HSB)		Difference-of-means (HSB - NLSY)		Statistical significance
	Mean	Std. Error	Mean	Std. Error	Value	Std. Error	
Population of black and white children	(N=1228)		(N=10,929)				
Percentage black	15.7	0.1	15.0	0.0	-0.7	2.2	
Population of white young adults	(N=857)		(N=8782)				
Percentage female	49.2	2.2	50.5	0.0	1.3	2.3	
Percentage attending high school in the Northeast	21.3	4.1	23.9	0.0	2.6	4.5	
Percentage attending high school in the Midwest	34.8	5.1	31.8	0.0	-3.0	5.4	
Percentage attending high school in the South	26.8	4.4	28.4	0.0	1.6	4.7	
Percentage attending high school in the West	17.1	3.8	15.9	0.0	-1.2	4.0	
Socioeconomic status	0.105	0.048	0.054	0.016	-0.051	0.051	
Educational achievement	0.2	0.0	0.2	0.0	0.0	0.0	
Percentage that completed high school	80.1	1.7	95.3	0.0	15.2	1.7	**
Percentage that attended college	52.0	2.2	60.2	0.0	8.3	2.3	**
Percentage that completed college	25.3	1.9	26.9	0.8	1.6	2.0	
Average years of work experience	8.3	0.1	7.6	0.0	-0.7	0.1	**
Percentage that are in the labor force	86.1	1.5	84.0	0.5	-2.0	1.5	
Percentage that have earnings	82.7	1.5	88.6	0.5	5.9	1.6	**
Earnings	18,407	661	20,668	255	2,261	718	**
Population of black young adults	(N=371)		(N=2147)				
Percentage female	48.9	2.7	52.6	1.6	3.7	3.2	
Percentage attending high school in the Northeast	18.4	4.5	21.8	2.8	3.4	5.3	
Percentage attending high school in the Midwest	18.9	4.4	17.0	2.4	-1.9	5.0	
Percentage attending high school in the South	57.6	5.9	54.5	3.4	-3.1	6.7	
Percentage attending high school in the West	5.1	2.1	6.7	1.3	1.6	2.5	
Socioeconomic status	-0.524	0.049	-0.459	0.035	0.065	0.060	
Educational achievement	-0.9	0.0	-0.7	0.0	0.2	0.1	**
Percentage that completed high school	69.7	2.9	92.9	1.0	23.3	3.0	**
Percentage that attended college	36.0	3.1	49.3	1.9	13.2	3.7	**
Percentage that completed college	10.3	1.9	12.0	1.0	1.7	2.2	
Average years of work experience	6.4	0.2	6.5	0.1	0.1	0.2	
Percentage that are in the labor force	80.8	2.5	80.3	1.5	-0.5	2.9	
Percentage that have earnings	72.5	2.7	79.0	1.7	6.5	3.3	**
Earnings	11,970	697	16,071	609	4,101	933	**
Difference between black and white children							
Percentage female	-0.3	3.5	2.1	1.7	2.5	3.9	
Percentage attending high school in the Northeast	-2.9	5.6	-2.1	3.0	0.8	6.3	
Percentage attending high school in the Midwest	-15.9	5.9	-14.9	2.7	1.1	6.5	
Percentage attending high school in the South	30.8	6.5	26.1	3.3	-4.7	7.3	
Percentage attending high school in the West	-12.0	4.1	-9.2	1.7	2.8	4.4	
Socioeconomic status	-0.629	0.070	-0.513	0.036	0.116	0.078	
Educational achievement	-1.1	0.1	-0.9	0.0	0.2	0.1	**
Percentage that completed high school	-10.4	3.3	-2.3	1.0	8.1	3.5	**
Percentage that attended college	-15.9	3.7	-10.9	2.0	5.0	4.3	
Percentage that completed college	-15.0	2.8	-14.9	1.3	0.1	3.1	
Average years of work experience	-1.9	0.2	-1.1	0.1	0.8	0.3	**
Percentage that are in the labor force	-5.2	2.8	-3.8	1.6	1.5	3.2	
Percentage that have earnings	-10.3	3.1	-9.6	1.8	0.6	3.7	
Earnings	-6,436	969	-4,597	661	1,840	1,234	

** Difference-of-means between two groups is statistically significant at $p < .05$.

NOTE: Standard errors have been adjusted for complex survey designs. Numbers may not sum to totals due to rounding. Sample sizes are smaller than those reported previously because observations with a sampling weight of zero were excluded from the tabulations. Multiple imputation used to account for missing data.

SOURCE: U.S. Department of Education, National Center for Education Statistics, High School and Beyond (HSB); and U.S. Department of Labor, Bureau of Labor Statistics, National Longitudinal Study of Youth (NLSY).

Prospects Cohort 3 sample were more likely to live in the Northeast, but otherwise, the characteristics of children in both samples were similar (table B.4). The black-white difference in the proportion of the sample in the bottom SES quartile was dissimilar between the two samples; black-white differences in background characteristics were otherwise similar for Cohort 1 and Cohort 3.

The Prospects Cohort 1 sample and the Prospects Cohort 7 sample were similar in terms of the proportion of the sample that was black, and the background characteristics of children in both samples were similar by race (table B.5). Black-white differences in background characteristics were similar for Cohort 1 and Cohort 7.

The Prospects Cohort 1 sample and the NELS:88 sample were similar in terms of the proportion of the sample that was black (table B.6). Compared with white children in the Cohort 1 sample, white children in the NELS:88 sample were more likely to attend school in the Northeast. Compared with black children in the Cohort 1 sample, black children in the NELS:88 sample were more likely to be in the lower-middle SES quartile. The background characteristics of children in the two samples were otherwise similar by race. The black-white differences in the proportion of the sample living in the Northeast, proportion in the lower-middle SES quartile, and average parental SES were dissimilar between the two samples; otherwise, Cohort 1 and NELS:88 samples were similar in black-white differences in background characteristics.

The Prospects Cohort 3 sample and the Prospects Cohort 7 sample were similar in terms of the proportion of the sample that was black, and the background characteristics of children in both samples were similar by race (table B.7). Black-white differences in background characteristics were similar for Cohort 3 and Cohort 7.

The Prospects Cohort 3 sample and the NELS:88 sample were similar in terms of the proportion of the sample that was black, the background characteristics of children in both samples were similar by race (table B.8), and black-white differences in background characteristics were similar for the two samples.

The Prospects Cohort 7 sample and the NELS:88 sample were similar in terms of the proportion of the sample that was black (table B.9). Background characteristics of white children were similar in the two samples. Compared with black children in the Cohort 7 sample, black children in the NELS:88 sample were more likely to be in the lower-middle SES quartile. The black-white difference in the proportion of the sample in the lower-middle SES quartile was dissimilar between the two samples; otherwise, black-white differences in background characteristics were similar for Cohort 7 and NELS:88.

With few exceptions, the background characteristics of children were very similar across the four samples used in the analyses of educational achievement. These similarities suggest that comparisons of black-white differences in educational achievement across the samples are not biased by systematic differences in the observed characteristics of each sample.

Table B.4 Comparison of black and white children in Prospects Cohort 1 with black and white children in Prospects Cohort 3

Sample definition and descriptive variables	Prospects Cohort 1 sample		Prospects Cohort 3 sample		Difference-of-means (absent - present)		Statistical significance
	Mean	Std. Error	Mean	Std. Error	Value	Std. Error	
Population of black and white children	(N=7542)		(N=11,017)				
Percentage black	19.0	2.9	15.9	2.0	-3.2	2.1	
Population of white children	(N=5279)		(N=7872)				
Percentage female	48.9	0.9	49.7	0.9	0.7	1.2	
Percentage attending high school in the Northeast	7.3	2.3	16.1	4.1	8.9	3.6	**
Percentage attending high school in the Midwest	28.3	6.4	26.4	5.3	-1.8	3.0	
Percentage attending high school in the South	43.9	6.9	39.1	5.9	-4.8	4.1	
Percentage attending high school in the West	20.5	5.2	18.3	4.4	-2.2	3.2	
Socioeconomic status	0.165	0.052	0.148	0.051	-0.018	0.049	
Percentage in bottom SES quartile	16.3	1.5	18.4	1.4	2.2	1.6	
Percentage in lower-middle SES quartile	26.0	1.6	25.6	1.5	-0.3	1.5	
Percentage in upper-middle SES quartile	28.7	1.2	27.2	1.0	-1.5	1.4	
Percentage in top SES quartile	29.1	2.5	28.7	2.6	-0.3	2.5	
Population of black children	(N=2263)		(N=3145)				
Percentage female	49.3	2.4	50.3	1.3	1.1	2.7	
Percentage attending high school in the Northeast	15.9	6.3	22.4	6.8	6.5	2.8	**
Percentage attending high school in the Midwest	9.5	4.2	10.2	3.7	0.7	2.2	
Percentage attending high school in the South	66.4	7.7	59.4	6.9	-7.0	5.1	
Percentage attending high school in the West	8.2	3.0	8.0	2.3	-0.2	2.5	
Socioeconomic status	-0.563	0.094	-0.471	0.081	0.092	0.067	
Percentage in bottom SES quartile	46.6	3.7	41.4	3.2	-5.2	3.1	
Percentage in lower-middle SES quartile	21.7	1.5	24.5	1.5	2.8	1.9	
Percentage in upper-middle SES quartile	19.5	2.1	21.2	2.2	1.7	2.7	
Percentage in top SES quartile	12.1	2.7	12.9	2.0	0.8	2.6	
Difference between black and white children							
Percentage female	0.3	2.7	0.7	1.6	0.3	3.0	
Percentage attending high school in the Northeast	8.7	6.4	6.2	7.2	-2.4	3.9	
Percentage attending high school in the Midwest	-18.8	6.6	-16.2	5.8	2.5	3.1	
Percentage attending high school in the South	22.5	8.6	20.3	7.7	-2.2	5.2	
Percentage attending high school in the West	-12.3	4.7	-10.3	4.3	2.1	2.7	
Socioeconomic status	-0.729	0.090	-0.619	0.085	0.110	0.066	
Percentage in bottom SES quartile	30.4	3.5	23.0	3.2	-7.4	3.2	**
Percentage in lower-middle SES quartile	-4.2	2.1	-1.1	2.1	3.1	2.4	
Percentage in upper-middle SES quartile	-9.2	2.2	-6.0	2.4	3.1	2.7	
Percentage in top SES quartile	-16.9	2.8	-15.9	2.9	1.1	3.0	

** Difference-of-means between two groups is statistically significant at $p \leq .05$.

NOTE: Standard errors have been adjusted for complex survey designs. Numbers may not sum to totals due to rounding. Sample sizes are smaller than those reported in table A.2 because observations with a sampling weight of zero were excluded from the tabulations. Multiple imputation used to account for missing data.

SOURCE: U.S. Department of Education, Chapter 1 Prospects Study.

Table B.5 Comparison of black and white children in Prospects Cohort 1 with black and white children in Prospects Cohort 7

Sample definition and descriptive variables	Prospects Cohort 1		Prospects Cohort 7		Difference-of-means (absent - present)		Statistical significance
	Mean	Std. Error	Mean	Std. Error	Value	Std. Error	
Population of black and white children	(N=7542)		(N=6654)				
Percentage black	19.0	0.0	16.7	0.0	-2.4	3.8	
Population of white children	(N=5279)		(N=4951)				
Percentage female	48.9	0.9	48.1	1.1	-0.8	1.4	
Percentage attending high school in the Northeast	7.3	2.3	14.2	4.3	7.0	4.9	
Percentage attending high school in the Midwest	28.3	6.4	27.7	6.3	-0.6	8.5	
Percentage attending high school in the South	43.9	6.9	40.2	6.4	-3.7	9.2	
Percentage attending high school in the West	20.5	5.2	17.9	4.7	-2.7	6.9	
Socioeconomic status	0.165	0.052	0.159	0.048	-0.006	0.069	
Percentage in bottom SES quartile	16.3	1.5	17.9	1.4	1.6	2.1	
Percentage in lower-middle SES quartile	26.0	1.6	26.3	1.2	0.3	2.0	
Percentage in upper-middle SES quartile	28.7	1.2	27.5	1.2	-1.2	1.7	
Percentage in top SES quartile	29.1	2.5	28.3	2.4	-0.7	3.4	
Population of black children	(N=2263)		(N=1703)				
Percentage female	49.3	2.4	51.6	1.7	2.4	2.9	
Percentage attending high school in the Northeast	15.9	6.3	14.0	4.9	-2.0	7.9	
Percentage attending high school in the Midwest	9.5	4.2	17.2	7.3	7.7	8.4	
Percentage attending high school in the South	66.4	7.7	57.5	8.3	-8.9	11.3	
Percentage attending high school in the West	8.2	3.0	11.3	4.0	3.2	5.0	
Socioeconomic status	-0.563	0.094	-0.474	0.085	0.089	0.128	
Percentage in bottom SES quartile	46.6	3.7	42.6	3.3	-4.1	5.0	
Percentage in lower-middle SES quartile	21.7	1.5	21.2	1.8	-0.5	2.3	
Percentage in upper-middle SES quartile	19.5	2.1	20.8	2.2	1.3	3.1	
Percentage in top SES quartile	12.1	2.7	15.5	2.2	3.3	3.5	
Difference between black and white children							
Percentage female	0.3	2.7	3.5	2.0	3.2	3.3	
Percentage attending high school in the Northeast	8.7	6.4	-0.3	6.1	-8.9	8.7	
Percentage attending high school in the Midwest	-18.8	6.6	-10.5	8.0	8.3	9.9	
Percentage attending high school in the South	22.5	8.6	17.3	8.9	-5.2	12.2	
Percentage attending high school in the West	-12.3	4.7	-6.5	5.2	5.8	6.9	
Socioeconomic status	-0.729	0.090	-0.633	0.093	0.096	0.129	
Percentage in bottom SES quartile	30.4	3.5	24.7	3.5	-5.7	5.0	
Percentage in lower-middle SES quartile	-4.2	2.1	-5.1	2.0	-0.8	2.9	
Percentage in upper-middle SES quartile	-9.2	2.2	-6.7	2.5	2.5	3.3	
Percentage in top SES quartile	-16.9	2.8	-12.9	3.0	4.1	4.1	

** Difference-of-means between two groups is statistically significant at $p \leq .05$.

NOTE: Standard errors have been adjusted for complex survey designs. Numbers may not sum to totals due to rounding. Sample sizes are smaller than those reported in table A.2 because observations with a sampling weight of zero were excluded from the tabulations. Multiple imputation used to account for missing data.

SOURCE: U.S. Department of Education, Chapter 1 Prospects Study.

Table B.6 Comparison of black and white children in Prospects Cohort 1 with black and white children in NELS:88

Sample definition and descriptive variables	Prospects Cohort 1 sample		NELS:88 sample		Difference-of-means (absent - present)		Statistical significance
	Mean	Std. Error	Mean	Std. Error	Value	Std. Error	
Population of black and white children	(N=7542)		(N=12,954)				
Percentage black	19.0	2.9	15.0	1.1	-4.1	3.1	
Population of white children	(N=5279)		(N=11,386)				
Percentage female	48.9	0.9	49.9	0.7	1.0	1.1	
Percentage attending high school in the Northeast	7.3	2.3	21.0	1.9	13.8	3.0	**
Percentage attending high school in the Midwest	28.3	6.4	31.4	1.9	3.1	6.7	
Percentage attending high school in the South	43.9	6.9	31.2	1.8	-12.7	7.1	
Percentage attending high school in the West	20.5	5.2	16.4	1.5	-4.1	5.4	
Socioeconomic status	0.165	0.052	0.095	0.020	-0.070	0.056	
Percentage in bottom SES quartile	16.3	1.5	17.0	0.8	0.7	1.7	
Percentage in lower-middle SES quartile	26.0	1.6	25.0	0.7	-0.9	1.8	
Percentage in upper-middle SES quartile	28.7	1.2	27.6	0.7	-1.1	1.3	
Percentage in top SES quartile	29.1	2.5	30.4	1.1	1.3	2.7	
Population of black children	(N=2263)		(N=1568)				
Percentage female	49.3	2.4	51.7	2.1	2.5	3.2	
Percentage attending high school in the Northeast	15.9	6.3	15.7	2.8	-0.2	6.8	
Percentage attending high school in the Midwest	9.5	4.2	13.0	2.1	3.5	4.7	
Percentage attending high school in the South	66.4	7.7	64.8	3.5	-1.6	8.5	
Percentage attending high school in the West	8.2	3.0	6.5	1.3	-1.7	3.3	
Socioeconomic status	-0.563	0.094	-0.387	0.041	0.176	0.102	
Percentage in bottom SES quartile	46.6	3.7	39.2	2.3	-7.4	4.4	
Percentage in lower-middle SES quartile	21.7	1.5	26.5	2.0	4.8	2.5	**
Percentage in upper-middle SES quartile	19.5	2.1	21.4	1.8	1.9	2.8	
Percentage in top SES quartile	12.1	2.7	12.9	2.0	0.8	3.3	
Difference between black and white children							
Percentage female	0.3	2.7	1.8	2.2	1.5	3.4	
Percentage attending high school in the Northeast	8.7	6.4	-5.3	3.2	-14.0	7.1	**
Percentage attending high school in the Midwest	-18.8	6.6	-18.3	2.6	0.4	7.1	
Percentage attending high school in the South	22.5	8.6	33.6	3.6	11.1	9.3	
Percentage attending high school in the West	-12.3	4.7	-9.9	1.9	2.4	5.1	
Socioeconomic status	-0.729	0.090	-0.483	0.044	0.246	0.100	**
Percentage in bottom SES quartile	30.4	3.5	22.2	2.4	-8.1	4.3	
Percentage in lower-middle SES quartile	-4.2	2.1	1.5	2.1	5.7	3.0	**
Percentage in upper-middle SES quartile	-9.2	2.2	-6.2	1.9	2.9	2.9	
Percentage in top SES quartile	-16.9	2.8	-17.5	2.2	-0.5	3.6	

** Difference-of-means between two groups is statistically significant at $p \leq .05$.

NOTE: Standard errors have been adjusted for complex survey designs. Numbers may not sum to totals due to rounding. Sample sizes are smaller than those reported in table A.2 because observations with a sampling weight of zero were excluded from the tabulations. Multiple imputation used to account for missing data.

SOURCE: U.S. Department of Education, Chapter 1 Prospects Study and National Educational Longitudinal Study of 1988.

Table B.7 Comparison of black and white children in Prospects Cohort 3 with black and white children in Prospects Cohort 7

Sample definition and descriptive variables	Prospects Cohort 3 sample		Prospects Cohort 7 sample		Difference-of-means (absent - present)		Statistical significance
	Mean	Std. Error	Mean	Std. Error	Value	Std. Error	
Population of black and white children	(N=11,017)		(N=6654)				
Percentage black	15.9	2.0	16.7	2.5	0.8	3.2	
Population of white children	(N=7872)		(N=4951)				
Percentage female	49.7	0.9	48.1	1.1	-1.6	1.4	
Percentage attending high school in the Northeast	16.1	4.1	14.2	4.3	-1.9	5.9	
Percentage attending high school in the Midwest	26.4	5.3	27.7	6.3	1.2	7.7	
Percentage attending high school in the South	39.1	5.9	40.2	6.4	1.1	8.6	
Percentage attending high school in the West	18.3	4.4	17.9	4.7	-0.4	6.3	
Socioeconomic status	0.148	0.051	0.159	0.048	0.011	0.069	
Percentage in bottom SES quartile	18.4	1.4	17.9	1.4	-0.5	1.9	
Percentage in lower-middle SES quartile	25.6	1.5	26.3	1.2	0.6	1.9	
Percentage in upper-middle SES quartile	27.2	1.0	27.5	1.2	0.3	1.6	
Percentage in top SES quartile	28.7	2.6	28.3	2.4	-0.4	3.5	
Population of black children	(N=3145)		(N=1703)				
Percentage female	50.3	1.3	51.6	1.7	1.3	2.1	
Percentage attending high school in the Northeast	22.4	6.8	14.0	4.9	-8.4	8.3	
Percentage attending high school in the Midwest	10.2	3.7	17.2	7.3	7.0	8.2	
Percentage attending high school in the South	59.4	6.9	57.5	8.3	-1.9	10.7	
Percentage attending high school in the West	8.0	2.3	11.3	4.0	3.3	4.6	
Socioeconomic status	-0.471	0.081	-0.474	0.085	-0.003	0.117	
Percentage in bottom SES quartile	41.4	3.2	42.6	3.3	1.1	4.7	
Percentage in lower-middle SES quartile	24.5	1.5	21.2	1.8	-3.3	2.4	
Percentage in upper-middle SES quartile	21.2	2.2	20.8	2.2	-0.4	3.1	
Percentage in top SES quartile	12.9	2.0	15.5	2.2	2.6	2.9	
Difference between black and white children							
Percentage female	0.7	1.6	3.5	2.0	2.9	2.5	
Percentage attending high school in the Northeast	6.2	7.2	-0.3	6.1	-6.5	9.4	
Percentage attending high school in the Midwest	-16.2	5.8	-10.5	8.0	5.8	9.5	
Percentage attending high school in the South	20.3	7.7	17.3	8.9	-3.0	11.6	
Percentage attending high school in the West	-10.3	4.3	-6.5	5.2	3.8	6.7	
Socioeconomic status	-0.619	0.085	-0.633	0.093	-0.014	0.125	
Percentage in bottom SES quartile	23.0	3.2	24.7	3.5	1.6	4.8	
Percentage in lower-middle SES quartile	-1.1	2.1	-5.1	2.0	-3.9	3.0	
Percentage in upper-middle SES quartile	-6.0	2.4	-6.7	2.5	-0.7	3.5	
Percentage in top SES quartile	-15.9	2.9	-12.9	3.0	3.0	4.0	

** Difference-of-means between two groups is statistically significant at $p \leq .05$.

NOTE: Standard errors have been adjusted for complex survey designs. Numbers may not sum to totals due to rounding. Sample sizes are smaller than those reported in table A.2 because observations with a sampling weight of zero were excluded from the tabulations. Multiple imputation used to account for missing data.

SOURCE: U.S. Department of Education, Chapter 1 Prospects Study.

Table B.8 Comparison of black and white children in Prospects Cohort 3 with black and white children in NELS:88

Sample definition and descriptive variables	Prospects Cohort 3 sample		NELS:88 sample		Difference-of-means (absent - present)		Statistical significance
	Mean	Std. Error	Mean	Std. Error	Value	Std. Error	
Population of black and white children	(N=11,017)		(N=12,954)				
Percentage black	15.9	2.0	15.0	1.1	-0.9	2.3	
Population of white children	(N=7872)		(N=11,386)				
Percentage female	49.7	0.9	49.9	0.7	0.3	1.2	
Percentage attending high school in the Northeast	16.1	4.1	21.0	1.9	4.9	4.5	
Percentage attending high school in the Midwest	26.4	5.3	31.4	1.9	4.9	5.6	
Percentage attending high school in the South	39.1	5.9	31.2	1.8	-7.9	6.2	
Percentage attending high school in the West	18.3	4.4	16.4	1.5	-1.9	4.6	
Socioeconomic status	0.148	0.051	0.095	0.020	-0.052	0.055	
Percentage in bottom SES quartile	18.4	1.4	17.0	0.8	-1.5	1.6	
Percentage in lower-middle SES quartile	25.6	1.5	25.0	0.7	-0.6	1.7	
Percentage in upper-middle SES quartile	27.2	1.0	27.6	0.7	0.4	1.2	
Percentage in top SES quartile	28.7	2.6	30.4	1.1	1.7	2.8	
Population of black children	(N=3145)		(N=1568)				
Percentage female	50.3	1.3	51.7	2.1	1.4	2.4	
Percentage attending high school in the Northeast	22.4	6.8	15.7	2.8	-6.7	7.4	
Percentage attending high school in the Midwest	10.2	3.7	13.0	2.1	2.8	4.2	
Percentage attending high school in the South	59.4	6.9	64.8	3.5	5.4	7.7	
Percentage attending high school in the West	8.0	2.3	6.5	1.3	-1.5	2.7	
Socioeconomic status	-0.471	0.081	-0.387	0.041	0.084	0.090	
Percentage in bottom SES quartile	41.4	3.2	39.2	2.3	-2.2	3.9	
Percentage in lower-middle SES quartile	24.5	1.5	26.5	2.0	2.0	2.5	
Percentage in upper-middle SES quartile	21.2	2.2	21.4	1.8	0.2	2.8	
Percentage in top SES quartile	12.9	2.0	12.9	2.0	0.0	2.8	
Difference between black and white children							
Percentage female	0.7	1.6	1.8	2.2	1.1	2.7	
Percentage attending high school in the Northeast	6.2	7.2	-5.3	3.2	-11.6	7.9	
Percentage attending high school in the Midwest	-16.2	5.8	-18.3	2.6	-2.1	6.3	
Percentage attending high school in the South	20.3	7.7	33.6	3.6	13.3	8.5	
Percentage attending high school in the West	-10.3	4.3	-9.9	1.9	0.4	4.7	
Socioeconomic status	-0.619	0.085	-0.483	0.044	0.136	0.096	
Percentage in bottom SES quartile	23.0	3.2	22.2	2.4	-0.8	4.0	
Percentage in lower-middle SES quartile	-1.1	2.1	1.5	2.1	2.6	3.0	
Percentage in upper-middle SES quartile	-6.0	2.4	-6.2	1.9	-0.2	3.0	
Percentage in top SES quartile	-15.9	2.9	-17.5	2.2	-1.6	3.7	

** Difference-of-means between two groups is statistically significant at $p \leq .05$.

NOTE: Standard errors have been adjusted for complex survey designs. Numbers may not sum to totals due to rounding. Sample sizes are smaller than those reported in table A.2 because observations with a sampling weight of zero were excluded from the tabulations. Multiple imputation used to account for missing data.

SOURCE: U.S. Department of Education, Chapter 1 Prospects Study and National Educational Longitudinal Study of 1988.

Table B.9 Comparison of black and white children in Prospects Cohort 7 with black and white children in NELS:88

Sample definition and descriptive variables	Prospects Cohort 7 sample		NELS:88 sample		Difference-of-means (absent - present)		Statistical significance
	Mean	Std. Error	Mean	Std. Error	Value	Std. Error	
Population of black and white children	(N=6654)		(N=12,954)				
Percentage black	16.7	2.5	15.0	1.1	-1.7	2.7	
Population of white children	(N=4951)		(N=11,386)				
Percentage female	48.1	1.1	49.9	0.7	1.8	1.3	
Percentage attending high school in the Northeast	14.2	4.3	21.0	1.9	6.8	4.7	
Percentage attending high school in the Midwest	27.7	6.3	31.4	1.9	3.7	6.6	
Percentage attending high school in the South	40.2	6.4	31.2	1.8	-9.0	6.6	
Percentage attending high school in the West	17.9	4.7	16.4	1.5	-1.4	4.9	
Socioeconomic status	0.159	0.048	0.095	0.020	-0.064	0.052	
Percentage in bottom SES quartile	17.9	1.4	17.0	0.8	-1.0	1.6	
Percentage in lower-middle SES quartile	26.3	1.2	25.0	0.7	-1.2	1.4	
Percentage in upper-middle SES quartile	27.5	1.2	27.6	0.7	0.1	1.4	
Percentage in top SES quartile	28.3	2.4	30.4	1.1	2.1	2.6	
Grade 8 mathematics score	0.2	0.0	0.2	0.0	0.1	0.0	
Grade 8 reading score	0.2	0.0	0.2	0.0	0.0	0.0	
Population of black children	(N=1703)		(N=1568)				
Percentage female	51.6	1.7	51.7	2.1	0.1	2.7	
Percentage attending high school in the Northeast	14.0	4.9	15.7	2.8	1.7	5.6	
Percentage attending high school in the Midwest	17.2	7.3	13.0	2.1	-4.2	7.6	
Percentage attending high school in the South	57.5	8.3	64.8	3.5	7.3	9.0	
Percentage attending high school in the West	11.3	4.0	6.5	1.3	-4.8	4.3	
Socioeconomic status	-0.474	0.085	-0.387	0.041	0.086	0.094	
Percentage in bottom SES quartile	42.6	3.3	39.2	2.3	-3.4	4.0	
Percentage in lower-middle SES quartile	21.2	1.8	26.5	2.0	5.3	2.7	**
Percentage in upper-middle SES quartile	20.8	2.2	21.4	1.8	0.6	2.8	
Percentage in top SES quartile	15.5	2.2	12.9	2.0	-2.5	3.0	
Grade 8 mathematics score	-0.3	0.1	-0.5	0.0	-0.2	0.1	
Grade 8 reading score	-0.4	0.1	-0.5	0.0	0.0	0.1	
Difference between black and white children							
Percentage female	3.5	2.0	1.8	2.2	-1.7	2.9	
Percentage attending high school in the Northeast	-0.3	6.1	-5.3	3.2	-5.1	6.8	
Percentage attending high school in the Midwest	-10.5	8.0	-18.3	2.6	-7.8	8.4	
Percentage attending high school in the South	17.3	8.9	33.6	3.6	16.3	9.5	
Percentage attending high school in the West	-6.5	5.2	-9.9	1.9	-3.4	5.5	
Socioeconomic status	-0.633	0.093	-0.483	0.044	0.150	0.102	
Percentage in bottom SES quartile	24.7	3.5	22.2	2.4	-2.4	4.2	
Percentage in lower-middle SES quartile	-5.1	2.0	1.5	2.1	6.6	2.9	**
Percentage in upper-middle SES quartile	-6.7	2.5	-6.2	1.9	0.4	3.2	
Percentage in top SES quartile	-12.9	3.0	-17.5	2.2	-4.6	3.7	
Grade 8 mathematics score	-0.5	0.1	-0.8	0.0	-0.3	0.1	
Grade 8 reading score	-0.6	0.1	-0.7	0.0	0.0	0.1	

** Difference-of-means between two groups is statistically significant at $p \leq .05$.

NOTE: Standard errors have been adjusted for complex survey designs. Numbers may not sum to totals due to rounding. Sample sizes are smaller than those reported in table A.2 because observations with a sampling weight of zero were excluded from the tabulations. Multiple imputation used to account for missing data.

SOURCE: U.S. Department of Education, Chapter 1 Prospects Study and National Educational Longitudinal Study of 1988.

APPENDIX C. ANALYSIS METHODOLOGY

I. APPROACH TO ESTIMATION OF BLACK-WHITE DIFFERENCES IN OUTCOMES

Chapters 2, 3, and 4 involve the analysis of black-white differences in outcomes for individuals similar in terms of prior educational achievement and/or other factors. In general, this approach involves estimating a series of linear regressions for the same set of observations, of the form:

$$OUTCOME_i = b_{11} + b_{12} * BLACK_i + e_{1i} \quad (1)$$

where $OUTCOME_i$ is the outcome value for student i (for example, growth in educational achievement between two grades); $BLACK_i$ is a variable set equal to one if student i is black and zero if student i is white; the b 's are individual parameters to be estimated; and e_{1i} is an error term for student i .

For each of the five datasets containing imputed values, we estimated black-white differences in outcomes. Equation (1) estimates the difference of the outcome for blacks and whites (b_{12}). This difference is compared with the black-white difference in the outcome for individuals with similar levels of prior educational achievement.

The propensity score approach described by Rosenbaum and Rubin (1984) and Rubin (1997) was used to account for prior educational achievement. Essentially, this method accounts for differences between two groups of individuals by classifying them into five groups, based on their propensity of being in the same group. Propensity scores were used to classify each sample into five groups based on the likelihood of individuals being black. Separate logistic regressions of race on prior achievement and prior achievement squared were estimated for males and females. The estimated propensity score is defined as the predicted probability of being black from this logistic regression. Propensity score quintiles were identified from the weighted distribution of the propensity scores for whites, and blacks were placed in one of the five groups depending on their propensity scores. A linear regression of the outcome variable on race was fit separately to data in each of the five groups:

$$OUTCOME_{ij} = b_{21ij} + b_{22ij} * BLACK_{ij} + e_{2ij} \quad (2)$$

where $j=1$ to 5 indexes the 5 subgroups defined by propensity score quintiles. The black-white difference for the outcome in question, after adjusting for achievement, (b_{22}) is estimated by the simple average of the five race coefficients b_{221} through b_{225} . The variance of this black-white difference was estimated as 1/25 of the sum of the variances of

the five race coefficients (since the variance of the average of five independent coefficients equals 1/25 of the sum of the variances of each coefficient).

Estimates of black-white differences for individuals with similar levels of prior achievement plus similar levels of other factors were obtained in a similar fashion. Propensity scores were estimated from a logistic regression of race on prior achievement, prior achievement squared, and other factors. In the analysis of labor market outcomes, the other factors include educational attainment and work experience; in the analyses of educational attainment and achievement outcomes, the other factors include parental socioeconomic status (in both linear and quadratic terms) and census region of residence. After classifying the sample into five subgroups on the basis of the propensity score quintiles, black-white differences were estimated for each subgroup using equation (3):

$$OUTCOME_i = b_{31k} + b_{32k} * BLACK_i + e_{3ik} \quad (3)$$

where $k=1$ to 5 indexes the 5 subgroups defined by propensity score quintiles based on achievement and other factors. The overall estimate of black-white difference after adjusting for achievement and other factors, b_{32} , is the simple average of the five race coefficients b_{321} through b_{325} . The variance of this black-white difference was estimated as 1/25 of the sum of the variances of the five race coefficients.

Two sorts of statistical tests were performed to analyze the black-white differences estimated in equations (1) through (3). First, t-tests were performed to see whether each estimated black-white difference (b_{12} , b_{22} , and b_{32})—averaged across quintiles—was statistically different from zero. Next, tests were performed to see whether the addition of a particular set of variables to the outcome equation changed the estimate of black-white differences in outcomes. In particular, to test whether the addition of prior educational achievement changes the estimated difference in outcomes, a test was performed to determine whether the difference ($b_{22}-b_{12}$) was significantly different from zero. Following the work of Clogg, Petkova, and Cheng (1995), the unconditional variance of ($b_{22}-b_{12}$), which we denote as $s_u^2(b_{22}-b_{12})$, was estimated as follows:

$$s_u^2(b_{22}-b_{12}) = s^2(b_{22}) + s^2(b_{12}) - 2s^2(b_{12}) * (MSE_2/MSE_1) \quad (4)$$

where $s^2(b_{22})$ and $s^2(b_{12})$ are the variances of b_{22} and b_{12} , respectively, and MSE_1 and MSE_2 are the mean-squared errors from equations (1) and (2), respectively. MSE_2 is estimated by calculating the variance of the residuals in each of the five propensity score subgroups and taking the simple average of the results. The variance estimate is unconditional since it allows the predictor variables (for example, *BLACK* and *ACHIEVE*) to include random components (that is, it does not treat the predictor variables as fixed).

The use of the Clogg/Petkova/Cheng unconditional variance estimator placed certain constraints on the estimation of equations predicting educational and economic outcomes. First, since the estimator was only provided for linear regression models, binary outcome equations were estimated as linear probability models rather than using probit or logit maximum likelihood procedures. Second, because the estimator tests the difference of individual regression coefficients only, outcome equations were estimated pooling blacks and whites rather than separately by race. The test of black-white equality in outcomes was therefore the test of black-white equality in the constant term in outcome equations; other coefficients in these equations (such as those capturing the

relationship between achievement and the outcome variable) were assumed to be the same for blacks and whites.¹

The Clogg/Petkova/Cheng test for estimating the unconditional variance of regression coefficients was used to perform a t-test of whether $(b_{22}-b_{12})$ was significantly different from zero. A similar test was also performed to determine whether the addition of other factors to prior educational achievement changed the black-white difference estimated including prior achievement only; this test examined the significance of the difference $(b_{32}-b_{22})$.

In all instances, black-white differences were calculated jointly for males and females as well as separately by sex. When differences were calculated jointly, equations (1) through (3) were estimated for the full sample of men and women. In the joint model, black-white differences were constrained to be the same for males and females (that is, b_{12} , b_{22} , and b_{32} were not allowed to vary by sex), but the contributions of achievement and other explanatory variables to the propensity score were allowed to vary by sex. When analyses were performed separately by sex, equations (1), (2), and (3) were estimated separately for men and for women, and *all* of the parameters in the propensity score equations were allowed to vary by sex. Testing the equality of black-white differences for men and women was equivalent to testing whether the parameters b_{12} , b_{22} , or b_{32} differed for men and for women.

These results were estimated separately for each of the five datasets containing imputed values. Equations developed by Rubin (1987) were used to combine the estimates and standard errors of the regression coefficients. For each estimate of a value Q with variance V using $i = 1$ to m iterations of imputed data, the following quantities were calculated, using notation based on that employed by Schafer and Olsen (1998):

the mean estimate:

$$E(Q) = \sum_m Qi / m \quad (5)$$

the within-imputation variance:

$$W = \sum_m Vi / m \quad (6)$$

the between-imputation variance:

$$\begin{aligned} B &= \sum_m (Qi - E(Q))^2 / (m - 1) \\ &= (\sum_m Qi^2 - m E(Q)^2) / (m - 1) \end{aligned} \quad (7)$$

the total variance:

$$T = W + (1 + 1/m) B \quad (8)$$

The small-sample degrees of freedom for t-tests were calculated from the observed-data degrees of freedom as specified by Rubin (1987) and repeated by Schafer and Olsen (1998):

¹ Despite this restriction, black-white differences estimated under this strategy are similar in magnitude to those estimated using separate outcome equations for blacks and whites.

$$dfO = (m - 1) [1 + mW / (m + 1) / B]^2 \quad (9)$$

where m , W , and B are defined above.

Defining gamma as $1 / (1 + mW / (m + 1) / B)$ we can see that

$$\text{observed-data degrees of freedom: } dfO = (m - 1)\text{gamma}^{-2} \quad (10)$$

Following Barnard and Rubin (1999), the small-sample degrees of freedom (dfS) were defined as:

$$dfS = dfC / (1 / \text{lambda} / (1 - \text{gamma}) + dfC / dfO) \quad (11)$$

where dfC = the complete-data degrees of freedom and $\text{lambda} = (dfC + 1) / (dfC + 3)$. Significance tests were performed by comparing the ratio $E(Q)/\sqrt{T}$ to a Student's t -distribution with degrees of freedom dfS .

II. EFFECTS OF SAMPLE SELECTION ON ESTIMATES OF BLACK-WHITE DIFFERENCES

The estimation of outcome equations did not include corrections for sample selection bias. While selection bias arising from missing data was corrected for through multiple imputation, other selection bias arises from the definition of the outcome itself (for example, college completion for the sample of college attendees, or earnings for individuals with positive earnings). Sample selection models require the identification of factors influencing selection into the sample but not the outcome itself, and these factors are typically difficult to isolate. As long as selection into a sample is random, estimated black-white differences should be similar to those differences for the population as a whole. To the extent that selection into the sample is nonrandom, estimated black-white differences are valid in describing the population over which the estimation occurs, and reflect black-white differences in selection into the sample.

III. EFFECTS OF MEASUREMENT ERROR ON ESTIMATES OF BLACK-WHITE DIFFERENCES

Because test scores are an imperfect measure of actual educational achievement, and SES values are an imperfect measure of actual socioeconomic status, it is likely that measurement error affects the estimates of the propensity scores described above. However, it is unlikely that this measurement error has a large effect on estimated black-white differences. The propensity scoring method divides cases into subgroups on the basis of their estimated propensity scores. Even though measurement error in educational achievement and SES result in less precise estimates of the propensity scores, this only affects the estimation of black-white differences to the extent that it distorts the rank order of the cases by the estimated propensity score. A few cases with propensity scores close to the subgroup cutoffs may be placed in the wrong subgroup, but otherwise the measurement error does not affect the estimation of black-white differences.

APPENDIX D: SAMPLE SIZES AND STANDARD ERRORS FOR TABLES IN CHAPTERS 2, 3, AND 4

This appendix presents tables of sample sizes and standard errors corresponding to the results presented in chapters 2, 3, and 4. The standard errors were calculated accounting for the sample designs of the datasets studied. In particular, adjustments were made for the clustering of sample observations in the Primary Sampling Units (PSUs) included in each survey. The “linearization” approach was used, employing a first-order Taylor expansion to estimate the variance of the parameters of each linear regression equation.¹ Because of limited information on the stratification of PSUs in the NLSY and Prospects samples, no adjustments were made for stratification of PSUs (or of observations within PSUs). However, comparison of standard errors in the NELS:88 sample with adjustments for clustering only, and with adjustments for both clustering and stratification, indicated that adding adjustments for stratification increased sample design effects by no more than 0.2 percent, making it unlikely that the results of any hypothesis tests would differ if stratification were accounted for.

¹ Details of the estimation procedure used are discussed in the *Stata Reference Manual*, Release 5, Volume 3, P–Z, pp. 427–429.

Table D.1 Standard errors for Table 2.1: Black-white differences in labor force participation rates for young adults in 4 samples, 1979-1992

Description of sample of young adults, source of data, and sex of subgroup	Sample size	S.E. of black-white difference adjusting for			S.E. of change in black-white difference		
		(1) no variables	(2) achievement only	(3) multiple factors+	difference between (2) and (1)	difference between (3) and (1)	difference between (3) and (2)
High school seniors, 7 years later							
In 1979 (from the NLS-72)							
Men and women combined	12243	0.97	1.78	1.49	1.50	1.12	1.15
Men only	5791	1.16	3.01	1.43	2.78	0.84	2.67
Women only	6452	1.43	1.89	2.20	1.23	1.61	1.24
Difference between men and women	12243	1.84	3.53	2.62	3.02	1.81	2.69
In 1983-1989 (from the NLSY)							
Men and women combined	6893	1.21	1.60	1.67	1.07	1.18	0.58
Men only	3303	1.29	1.94	1.89	1.48	1.46	0.63
Women only	3590	1.86	2.57	2.76	1.80	2.06	0.98
Difference between men and women	6893	2.28	3.24	3.36	2.33	2.51	0.98
Difference between 1983-1989 and 1979 samples							
Men and women combined	19136	1.55	2.40	2.24	1.84	1.63	1.29
Men only	9094	1.74	3.59	2.38	3.15	1.68	2.75
Women only	10042	2.35	3.19	3.53	2.18	2.61	1.58
Difference between men and women	19136	2.93	4.79	4.26	3.82	3.10	2.86
High school sophomores, 12 years later							
In 1986-1992 (from the NLSY)							
Men and women combined	8838	1.14	1.17	1.66	0.33	1.20	1.17
Men only	4368	1.30	1.31	1.20	0.61	0.82	0.66
Women only	4470	1.69	1.84	3.15	0.80	2.70	2.59
Difference between men and women	8838	2.15	2.24	3.33	0.89	2.63	2.53
In 1992 (from HSB)							
Men and women combined	10857	1.61	1.34	1.33	0.88	0.91	0.62
Men only	5302	2.48	2.07	1.76	1.51	1.74	1.10
Women only	5555	1.89	1.82	1.87	0.69	0.62	0.61
Difference between men and women	10857	3.09	2.90	2.53	1.51	1.64	1.24
Difference between 1992 and 1986-1992 samples							
Men and women combined	19695	1.98	1.78	2.13	0.94	1.51	1.33
Men only	9670	2.81	2.45	2.13	1.62	1.93	1.29
Women only	10025	2.53	2.59	3.67	1.06	2.78	2.66
Difference between men and women	19695	3.77	3.66	4.18	1.75	3.10	2.82

+ = Multiple factors include educational achievement, educational attainment, and length of work experience

NOTE: Difference between group X and group Y refers to the mean for group X minus the mean for group Y. S.E. = standard error. Standard errors have been adjusted for complex survey designs. Samples restricted to civilians. Multiple imputation used to account for missing data.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Longitudinal Study of the High School Class of 1972

(NLS-72), High School and Beyond (HSB); and U.S. Department of Labor, Bureau of Labor Statistics, National Longitudinal Survey of Youth (NLSY).

Table D.2 Standard errors for Table 2.2: Black-white differences in unemployment rates for young adults in 4 samples, 1979-1992

Description of sample of young adults, source of data, and sex of subgroup	Sample size	S.E. of black-white difference adjusting for			S.E. of change in black-white difference		
		(1) no variables	(2) achievement only	(3) multiple factors+	difference between (2) and (1)	difference between (3) and (1)	difference between (3) and (2)
High school seniors, 7 years later							
In 1979 (from the NLS-72)							
Men and women combined	10581	0.75	0.82	0.93	0.34	0.57	0.45
Men only	5501	1.06	1.32	1.52	0.80	1.10	0.75
Women only	5080	1.08	0.84	0.66	0.66	0.85	0.54
Difference between men and women	10581	1.51	1.57	1.66	0.48	0.72	0.53
In 1983-1989 (from the NLSY)							
Men and women combined	5966	1.10	1.30	1.08	0.71	0.34	0.71
Men only	3062	1.68	2.00	1.24	1.08	1.05	1.50
Women only	2904	1.52	1.60	1.80	0.55	1.01	0.85
Difference between men and women	5966	2.30	2.62	2.22	1.24	0.57	1.24
Difference between 1983-1989 and 1979 samples							
Men and women combined	16547	1.33	1.54	1.43	0.79	0.67	0.84
Men only	8563	1.98	2.40	1.96	1.35	1.52	1.68
Women only	7984	1.86	1.81	1.92	0.86	1.33	1.00
Difference between men and women	16547	2.75	3.06	2.77	1.33	0.92	1.35
High school sophomores, 12 years later							
In 1986-1992 (from the NLSY)							
Men and women combined	7425	1.37	1.64	1.46	1.07	0.97	0.77
Men only	3890	2.02	2.43	2.13	1.71	1.58	0.94
Women only	3535	1.57	1.60	1.56	0.38	0.46	0.50
Difference between men and women	7425	2.44	2.89	2.63	1.73	1.58	1.07
In 1992 (from HSB)							
Men and women combined	9189	1.21	1.22	1.12	0.69	0.78	0.44
Men only	4884	1.45	1.47	1.21	0.78	0.86	0.78
Women only	4305	1.81	1.86	1.21	1.02	1.34	1.24
Difference between men and women	9189	2.23	2.22	1.47	1.17	1.61	1.57
Difference between 1992 and 1986-1992 samples							
Men and women combined	16614	1.83	2.05	1.84	1.27	1.24	0.88
Men only	8774	2.48	2.84	2.45	1.88	1.80	1.22
Women only	7840	2.40	2.45	1.97	1.08	1.41	1.34
Difference between men and women	16614	3.31	3.64	3.01	2.09	2.25	1.90

+ = Multiple factors include educational achievement, educational attainment, and length of work experience

NOTE: Difference between group X and group Y refers to the mean for group X minus the mean for group Y. S.E. = standard error. Standard errors have been adjusted for complex survey designs. Samples restricted to civilian labor force participants. Multiple imputation used to account for missing data.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Longitudinal Study of the High School Class of 1972 (NLS-72), High School and Beyond (HSB); and U.S. Department of Labor, Bureau of Labor Statistics, National Longitudinal Survey of Youth (NLSY).

Table D.3 Standard errors for Table 2.3: Black-white differences in average annual earnings for young adults in 4 samples, 1979-1992

Description of sample of young adults, source of data, and sex of subgroup	Sample size	S.E. of black-white difference adjusting for			S.E. of change in black-white difference		
		(1) no variables	(2) achievement only	(3) multiple factors+	difference between (2) and (1)	difference between (3) and (1)	difference between (3) and (2)
High school seniors, 7 years later							
In 1979 (from the NLS-72)							
Men and women combined	11,162	0.034	0.043	0.040	0.028	0.023	0.018
Men only	5,834	0.044	0.064	0.058	0.050	0.041	0.033
Women only	5,328	0.047	0.053	0.047	0.026	0.016	0.024
Difference between men and women	11,162	0.064	0.083	0.073	0.057	0.042	0.042
In 1983-1989 (from the NLSY)							
Men and women combined	6,095	0.044	0.044	0.054	0.015	0.038	0.034
Men only	3,102	0.055	0.063	0.072	0.034	0.054	0.040
Women only	2,993	0.053	0.065	0.082	0.041	0.066	0.052
Difference between men and women	6,095	0.076	0.090	0.110	0.053	0.085	0.065
Difference between 1983-1989 and 1979 samples							
Men and women combined	17,257	0.055	0.062	0.067	0.031	0.044	0.039
Men only	8,936	0.071	0.090	0.093	0.061	0.067	0.052
Women only	8,321	0.071	0.084	0.095	0.049	0.068	0.057
Difference between men and women	17,257	0.100	0.123	0.133	0.077	0.095	0.077
High school sophomores, 12 years later							
In 1986-1992 (from the NLSY)							
Men and women combined	6,666	0.041	0.048	0.059	0.024	0.045	0.041
Men only	3,459	0.050	0.056	0.059	0.029	0.037	0.022
Women only	3,207	0.059	0.063	0.100	0.029	0.084	0.080
Difference between men and women	6,666	0.077	0.084	0.115	0.042	0.090	0.082
In 1992 (from HSB)							
Men and women combined	9,592	0.036	0.039	0.045	0.015	0.022	0.018
Men only	5,002	0.045	0.046	0.046	0.021	0.021	0.022
Women only	4,590	0.054	0.058	0.075	0.026	0.044	0.031
Difference between men and women	9,592	0.073	0.071	0.081	0.034	0.042	0.035
Difference between 1992 and 1986-1992 samples							
Men and women combined	16,258	0.055	0.062	0.075	0.028	0.050	0.045
Men only	8,461	0.067	0.072	0.075	0.036	0.042	0.031
Women only	7,797	0.080	0.085	0.125	0.039	0.094	0.085
Difference between men and women	16,258	0.106	0.110	0.141	0.054	0.100	0.090

+ = Multiple factors include educational achievement, educational attainment, and length of work experience

NOTE: Difference between group X and group Y refers to the mean for group X minus the mean for group Y. S.E. = standard error of the log of annual earnings. Standard errors have been adjusted for complex survey designs. Samples restricted to employed civilians. Multiple imputation used to account for missing data.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Longitudinal Study of the High School Class of 1972 (NLS-72), High School and Beyond (HSB); and U.S. Department of Labor, Bureau of Labor Statistics, National Longitudinal Survey of Youth (NLSY).

Table D.4 Standard errors for Table 2.4: Black-white differences in hourly wages for young adults in 3 samples, 1979-1992

Description of sample of young adults, source of data, and sex of subgroup	Sample size	S.E. of black-white difference adjusting for			S.E. of change in black-white difference		
		(1) no variables	(2) achievement only	(3) multiple factors+	difference between (2) and (1)	difference between (3) and (1)	difference between (3) and (2)
High school seniors, 7 years later							
In 1979 (from the NLS-72)							
Men and women combined	10106	0.018	0.037	0.032	0.032	0.025	0.016
Men only	5333	0.030	0.064	0.046	0.055	0.035	0.037
Women only	4773	0.020	0.028	0.027	0.021	0.020	0.011
Difference between men and women	10106	0.036	0.069	0.054	0.060	0.042	0.038
In 1983-1989 (from the NLSY)							
Men and women combined	5441	0.020	0.028	0.028	0.021	0.021	0.005
Men only	2794	0.025	0.046	0.046	0.039	0.040	0.011
Women only	2647	0.025	0.037	0.036	0.029	0.028	0.013
Difference between men and women	5441	0.036	0.059	0.061	0.049	0.049	0.017
Difference between 1983-1989 and 1979 samples							
Men and women combined	15547	0.027	0.046	0.042	0.038	0.033	0.017
Men only	8127	0.039	0.078	0.066	0.068	0.053	0.038
Women only	7420	0.032	0.046	0.045	0.036	0.034	0.017
Difference between men and women	15547	0.051	0.091	0.081	0.077	0.064	0.042
High school sophomores, 12 years later							
In 1986-1992 (from the NLSY)							
Men and women combined	6170	0.023	0.026	0.033	0.015	0.025	0.019
Men only	3291	0.027	0.041	0.047	0.033	0.039	0.021
Women only	2879	0.030	0.031	0.041	0.015	0.031	0.027
Difference between men and women	6170	0.040	0.051	0.062	0.036	0.050	0.034

+ = Multiple factors include educational achievement, educational attainment, and length of work experience

NOTE: Difference between group X and group Y refers to the mean for group X minus the mean for group Y. S.E. = standard error of the log of hourly wage. Standard errors have been adjusted for complex survey designs. Samples restricted to employed civilians. Multiple imputation used to account for missing data.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Longitudinal Study of the High School Class of 1972 (NLS-72); and U.S. Department of Labor, Bureau of Labor Statistics, National Longitudinal Survey of Youth (NLSY).

Table D.5 Standard errors for Table 3.1: Black-white differences in high school/GED completion rates for young adults in 3 samples, 1983-1992

Description of sample of young adults, source of data, and sex of subgroup	Sample size	S.E. of black-white difference adjusting for			S.E. of change in black-white difference		
		(1) no variables	(2) achievement only	(3) multiple factors+	difference between (2) and (1)	difference between (3) and (1)	difference between (3) and (2)
High school seniors, 7 years later							
In 1983-1989 (from the NLSY)							
Men and women combined	7424	0.97	1.29	1.46	0.89	1.12	0.69
Men only	3636	1.43	1.26	1.43	0.63	0.36	0.72
Women only	3788	1.18	2.12	2.53	1.78	2.26	1.43
Difference between men and women	7424	1.85	2.46	2.92	1.67	2.29	1.57
High school sophomores, 12 years later							
In 1986-1992 (from the NLSY)							
Men and women combined	8998	1.22	0.93	1.07	0.66	0.40	0.53
Men only	4506	1.70	1.33	1.70	0.81	0.76	1.04
Women only	4492	1.54	1.09	1.26	0.93	0.75	0.61
Difference between men and women	8998	2.29	1.72	2.11	1.23	0.51	1.20
In 1992 (from HSB)							
Men and women combined	10929	1.02	0.66	0.70	0.80	0.70	0.41
Men only	5367	1.66	1.05	1.31	1.15	0.78	0.82
Women only	5562	1.55	1.01	0.69	1.15	1.24	0.61
Difference between men and women	10929	2.55	1.46	1.59	1.68	1.58	0.72
Difference between 1992 and 1986-1992 samples							
Men and women combined	19927	1.59	1.14	1.28	1.04	0.80	0.68
Men only	9873	2.38	1.70	2.14	1.41	1.09	1.32
Women only	10054	2.18	1.48	1.44	1.48	1.45	0.87
Difference between men and women	19927	3.43	2.26	2.64	2.09	1.66	1.40

+ = Multiple factors include educational achievement, parental socioeconomic status, and Census region

NOTE: Difference between group X and group Y refers to the mean for group X minus the mean for group Y. S.E. = standard error. Standard errors have been adjusted for complex survey designs. Multiple imputation used to account for missing data.

SOURCE: U.S. Department of Education, National Center for Education Statistics, High School and Beyond (HSB); and U.S. Department of Labor, Bureau of Labor Statistics, National Longitudinal Survey of Youth (NLSY).

Table D.6 Standard errors for Table 3.2: Black-white differences in college attendance rates for young adults in 4 samples, 1979-1992

Description of sample of young adults, source of data, and sex of subgroup	Sample size	S.E. of black-white difference adjusting for			S.E. of change in black-white difference		
		(1) no variables	(2) achievement only	(3) multiple factors+	difference between (2) and (1)	difference between (3) and (1)	difference between (3) and (2)
High school seniors, 7 years later							
In 1979 (from the NLS-72)							
Men and women combined	12577	1.77	1.25	1.19	0.71	0.82	0.50
Men only	6078	2.47	1.85	1.84	0.64	0.95	0.83
Women only	6499	2.16	1.87	2.84	0.61	2.28	2.20
Difference between men and women	12577	3.29	2.64	3.37	0.46	2.17	2.16
In 1983-1989 (from the NLSY)							
Men and women combined	7424	2.00	1.85	2.34	0.96	1.64	1.36
Men only	3636	2.44	2.13	2.68	1.08	1.87	1.54
Women only	3788	2.54	2.97	3.47	2.20	2.74	1.54
Difference between men and women	7424	3.53	3.66	4.39	2.51	3.36	2.20
Difference between 1983-1989 and 1979 samples							
Men and women combined	20001	2.67	2.23	2.62	1.19	1.84	1.45
Men only	9714	3.48	2.83	3.25	1.26	2.10	1.75
Women only	10287	3.33	3.51	4.48	2.29	3.56	2.68
Difference between men and women	20001	4.82	4.51	5.54	2.55	4.00	3.09
High school sophomores, 12 years later							
In 1986-1992 (from the NLSY)							
Men and women combined	7344	2.22	1.56	1.89	1.02	0.65	1.09
Men only	3534	2.67	2.12	2.17	0.45	0.64	0.61
Women only	3810	2.65	1.97	2.14	0.85	0.77	0.55
Difference between men and women	7344	3.76	2.89	3.05	0.85	0.96	0.60
In 1992 (from HSB)							
Men and women combined	10278	1.99	2.33	2.96	1.81	2.47	1.72
Men only	5015	2.76	3.48	4.43	2.61	3.64	2.55
Women only	5263	2.71	2.54	2.67	1.47	1.68	1.24
Difference between men and women	10278	3.85	4.21	5.24	2.87	4.03	3.20
Difference between 1992 and 1986-1992 samples							
Men and women combined	17622	2.98	2.81	3.51	2.08	2.56	2.04
Men only	8549	3.85	4.08	4.93	2.65	3.69	2.63
Women only	9073	3.79	3.22	3.42	1.70	1.85	1.36
Difference between men and women	17622	5.38	5.11	6.06	3.00	4.15	3.25

+ = Multiple factors include educational achievement, parental socioeconomic status, and Census region

NOTE: Difference between group X and group Y refers to the mean for group X minus the mean for group Y. S.E. = standard error. Standard errors have been adjusted for complex survey designs. Sophomore samples restricted to high school graduates. Multiple imputation used to account for missing data.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Longitudinal Study of the High School Class of 1972 (NLS-72), High School and Beyond (HSB); and U.S. Department of Labor, Bureau of Labor Statistics, National Longitudinal Survey of Youth (NLSY).

Table D.7 Standard errors for Table 3.3: Black-white differences in college completion rates for young adults in 4 samples, 1979-1992

Description of sample of young adults, source of data, and sex of subgroup	Sample size	S.E. of black-white difference adjusting for			S.E. of change in black-white difference		
		(1) no variables	(2) achievement only	(3) multiple factors+	difference between (2) and (1)	difference between (3) and (1)	difference between (3) and (2)
High school seniors, 7 years later							
In 1979 (from the NLS-72)							
Men and women combined	7899	1.92	3.77	4.69	3.42	4.395	2.82
Men only	3965	2.71	5.11	7.54	4.55	7.17	5.61
Women only	3934	2.50	5.26	5.98	4.80	5.58	2.89
Difference between men and women	7899	3.68	7.28	9.47	6.52	8.92	6.30
In 1983-1989 (from the NLSY)							
Men and women combined	3807	2.31	4.40	4.92	3.95	4.48	1.75
Men only	1776	3.03	5.65	6.46	5.01	5.84	2.64
Women only	2031	3.05	5.76	6.68	5.25	6.22	3.15
Difference between men and women	3807	4.30	8.09	9.29	7.28	8.55	3.56
Difference between 1983-1989 and 1979 samples							
Men and women combined	11706	3.01	5.79	6.80	5.22	6.28	3.32
Men only	5741	4.06	7.62	9.93	6.77	9.25	6.20
Women only	5965	3.94	7.80	8.97	7.12	8.36	4.27
Difference between men and women	11706	5.66	10.88	13.27	9.78	12.35	7.24
High school sophomores, 12 years later							
In 1986-1992 (from the NLSY)							
Men and women combined	4300	2.28	3.98	4.54	3.54	4.11	1.90
Men only	2011	2.92	5.34	6.32	4.80	5.77	2.67
Women only	2289	2.85	4.36	5.52	3.77	4.98	3.03
Difference between men and women	4300	4.08	6.82	8.38	6.04	7.63	4.08
In 1992 (from HSB)							
Men and women combined	6835	2.09	2.60	2.83	1.85	2.17	1.31
Men only	3263	3.01	3.59	4.06	2.42	2.98	2.32
Women only	3572	2.66	3.97	3.72	3.02	3.19	2.19
Difference between men and women	6835	4.04	5.56	5.60	3.92	4.52	3.66
Difference between 1992 and 1986-1992 samples							
Men and women combined	11135	3.10	4.76	5.35	4.00	4.64	2.31
Men only	5274	4.19	6.44	7.51	5.38	6.50	3.54
Women only	5861	3.90	5.89	6.66	4.83	5.92	3.74
Difference between men and women	11135	5.74	8.80	10.08	7.20	8.86	5.48

+ = Multiple factors include educational achievement, parental socioeconomic status, and Census region

NOTE: Difference between group X and group Y refers to the mean for group X minus the mean for group Y. S.E. = standard error. Standard errors have been adjusted for complex survey designs. Sophomore samples are restricted to persons who have attended some college. Multiple imputation used to account for missing data.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Longitudinal Study of the High School Class of 1972 (NLS-72), High School and Beyond (HSB); and U.S. Department of Labor, Bureau of Labor Statistics, National Longitudinal Survey of Youth (NLSY).

Table D.8 Standard errors for Table 4.1: Black-white differences in average mathematics achievement in grades 1 to 2 and 3 to 5 for 2 samples of children, 1991-1993

Description of source of data, grade and outcome, and sex of subgroup	Sample size	S.E. of black-white difference adjusting for			S.E. of change in black-white difference		
		(1) no variables	(2) achievement only	(3) multiple factors+	difference between (2) and (1)	difference between (3) and (1)	difference between (3) and (2)
Prospects Cohort 1 (1992-1993)							
Grade 1 mathematics achievement							
Boys and girls combined	7542	0.176	—	—	—	—	—
Boys only	3829	0.202	—	—	—	—	—
Girls only	3713	0.158	—	—	—	—	—
Difference between boys and girls	7542	0.254	—	—	—	—	—
Grade 2 mathematics achievement							
Boys and girls combined	7542	0.117	0.072	0.116	0.033	0.065	0.057
Boys only	3829	0.112	0.096	0.201	0.070	0.156	0.135
Girls only	3713	0.145	0.078	0.132	0.043	0.091	0.069
Difference between boys and girls	7542	0.180	0.120	0.237	0.096	0.177	0.151
Growth between grades 1 and 2							
Boys and girls combined	7542	0.130	0.068	0.078	0.081	0.099	0.027
Boys only	3829	0.165	0.088	0.119	0.116	0.141	0.068
Girls only	3713	0.116	0.080	0.117	0.055	0.080	0.062
Difference between boys and girls	7542	0.187	0.113	0.166	0.131	0.173	0.106
Prospects Cohort 3 (1991-1993)							
Grade 3 mathematics achievement							
Boys and girls combined	11017	0.067	—	—	—	—	—
Boys only	5553	0.099	—	—	—	—	—
Girls only	5464	0.065	—	—	—	—	—
Difference between boys and girls	11017	0.123	—	—	—	—	—
Grade 5 mathematics achievement							
Boys and girls combined	11017	0.060	0.034	0.082	0.022	0.064	0.070
Boys only	5553	0.066	0.045	0.154	0.041	0.140	0.154
Girls only	5464	0.071	0.044	0.084	0.028	0.053	0.055
Difference between boys and girls	11017	0.096	0.067	0.182	0.046	0.156	0.163
Growth between grades 3 and 5							
Boys and girls combined	11017	0.107	0.073	0.118	0.062	0.110	0.103
Boys only	5553	0.180	0.105	0.153	0.131	0.196	0.112
Girls only	5464	0.101	0.089	0.199	0.048	0.156	0.144
Difference between boys and girls	11017	0.216	0.149	0.286	0.151	0.261	0.190

+ = Multiple factors include achievement (average of math and reading scores from initial grade), parental socioeconomic status, and Census region

— = Not applicable because of absence of information on math or reading achievement prior to the initial grade for each sample

NOTE: Difference between group X and group Y refers to the mean for group X minus the mean for group Y. Standard errors have been adjusted for complex survey designs. 8th Grade SDUs = eighth-grade standard deviation units; 8th grade score is normalized to have mean of zero and standard deviation of 1.00. Multiple imputation used to account for missing data.

SOURCE: U.S. Department of Education, Chapter 1 Prospects Study.

Table D.9 Standard errors for Table 4.2: Black-white differences in average mathematics achievement in grades 7 to 9 and 10 to 12 for 2 samples of children, 1990-1993

Description of source of data, grade and outcome, and sex of subgroup	Sample size	S.E. of black-white difference adjusting for			S.E. of change in black-white difference		
		(1) no variables	(2) achievement only	(3) multiple factors+	difference between (2) and (1)	difference between (3) and (1)	difference between (3) and (2)
Prospects Cohort 7 (1991-1993)							
Grade 7 mathematics achievement							
Boys and girls combined	6654	0.056	—	—	—	—	—
Boys only	3419	0.066	—	—	—	—	—
Girls only	3235	0.087	—	—	—	—	—
Difference between boys and girls	6654	0.114	—	—	—	—	—
Grade 9 mathematics achievement							
Boys and girls combined	6654	0.076	0.060	0.081	0.044	0.056	0.058
Boys only	3419	0.089	0.069	0.111	0.050	0.086	0.082
Girls only	3235	0.095	0.103	0.117	0.072	0.086	0.049
Difference between boys and girls	6654	0.134	0.128	0.167	0.085	0.121	0.099
Growth between grades 7 and 9							
Boys and girls combined	6654	0.111	0.118	0.152	0.081	0.094	0.109
Boys only	3419	0.146	0.138	0.205	0.072	0.106	0.135
Girls only	3235	0.167	0.197	0.240	0.125	0.171	0.132
Difference between boys and girls	6654	0.237	0.243	0.323	0.121	0.197	0.187
NELS:88 (1990-1992)							
Grade 8 mathematics achievement							
Boys and girls combined	12954	0.046	—	—	—	—	—
Boys only	6409	0.058	—	—	—	—	—
Girls only	6545	0.060	—	—	—	—	—
Difference between boys and girls	12954	0.083	—	—	—	—	—
Grade 12 mathematics achievement							
Boys and girls combined	12954	0.065	0.039	0.060	0.061	0.059	0.054
Boys only	6409	0.077	0.048	0.083	0.074	0.083	0.055
Girls only	6545	0.089	0.063	0.076	0.087	0.074	0.105
Difference between boys and girls	12954	0.117	0.083	0.106	0.112	0.094	0.106
Growth between grades 8 and 12							
Boys and girls combined	12954	0.155	0.153	0.163	0.083	0.157	0.100
Boys only	6409	0.154	0.163	0.275	0.112	0.253	0.180
Girls only	6545	0.247	0.244	0.203	0.102	0.216	0.189
Difference between boys and girls	12954	0.274	0.304	0.315	0.104	0.168	0.165

+ = Multiple factors include achievement (average of math and reading scores from initial grade), parental socioeconomic status, and Census region

— = Not applicable because of absence of information on math or reading achievement prior to the initial grade for each sample

NOTE: Difference between group X and group Y refers to the mean for group X minus the mean for group Y. Standard errors have been adjusted for complex survey designs. 8th Grade SDUs = eighth-grade standard deviation units; 8th grade score is normalized to have mean of zero and standard deviation of 1.00. Multiple imputation used to account for missing data.

SOURCE: U.S. Department of Education, Chapter 1 Prospects Study, and National Educational Longitudinal Study of 1988 (NELS:88).

Table D.10 Standard errors for Table 4.3: Black-white differences in average mathematics achievement in grades 2, 5, 9, and 12 for 4 samples of children, 1992-1993

Description of sample(s), year(s) of data, and grades for which differences are calculated	Standard error of difference between blacks and whites (in 8th grade SDUs) for				
	Sample size	boys and girls combined	boys only	girls only	difference of boys and girls
Grade 2—Prospects Cohort 1 data (observed 1993)	4474	0.00	0.27	0.23	0.13
Grade 5—Prospects Cohort 3 data (observed 1993)	4224	0.00	0.11	0.09	0.09
Grade 5 math gap minus Grade 2 math gap	8698	0.26	0.29	0.23	0.16
Grade 5—Prospects Cohort 3 data (observed 1993)	4224	0.00	0.11	0.09	0.09
Grade 9—Prospects Cohort 7 data (observed 1993)	2332	0.00	0.13	0.11	0.14
Grade 9 math gap minus Grade 5 math gap	6556	0.13	0.17	0.14	0.16
Grade 9—Prospects Cohort 7 data (observed 1993)	2332	0.00	0.13	0.11	0.14
Grade 12—NELS:88 data (observed 1992)	9571	0.00	0.08	0.07	0.10
Grade 12 math gap minus Grade 9 math gap	11903	0.11	0.15	0.13	0.17
Grade 2—Prospects Cohort 1 data (observed 1993)	4474	0.00	0.27	0.23	0.13
Grade 9—Prospects Cohort 7 data (observed 1993)	2332	0.00	0.13	0.11	0.14
Grade 9 math gap minus Grade 2 math gap	6806	0.26	0.30	0.25	0.19
Grade 2—Prospects Cohort 1 data (observed 1993)	4474	0.00	0.27	0.23	0.13
Grade 12—NELS:88 data (observed 1992)	9571	0.00	0.08	0.07	0.10
Grade 12 math gap minus Grade 2 math gap	14045	0.25	0.28	0.24	0.16

NOTE: Difference between group X and group Y refers to the mean for group X minus the mean for group Y. Standard errors have been adjusted for complex survey designs. 8th Grade SDUs = eighth-grade standard deviation units; 8th grade score is normalized to have mean of zero and standard deviation of 1.00. Multiple imputation has been used to account for missing data.

SOURCE: U.S. Department of Education, Chapter 1 Prospects Study, and National Educational Longitudinal Study of 1988 (NELS:88).

Table D.11 Standard error for Table 4.4: Black-white differences in average reading achievement in grades 1 to 2 and 3 to 5 for 2 samples of children, 1991-1993

Description of source of data, grade and outcome, and sex of subgroup	Sample size	S.E. of black-white difference adjusting for			S.E. of change in black-white difference		
		(1) no variables	(2) achievement only	(3) multiple factors+	difference between (2) and (1)	difference between (3) and (1)	difference between (3) and (2)
Prospects Cohort 1 (1992-1993)							
Grade 1 reading achievement							
Boys and girls combined	7542	0.165	—	—	—	—	—
Boys only	3829	0.180	—	—	—	—	—
Girls only	3713	0.151	—	—	—	—	—
Difference between boys and girls	7542	0.234	—	—	—	—	—
Grade 2 reading achievement							
Boys and girls combined	7542	0.174	0.070	0.144	0.055	0.115	0.099
Boys only	3829	0.224	0.097	0.181	0.095	0.133	0.108
Girls only	3713	0.124	0.071	0.118	0.067	0.073	0.068
Difference between boys and girls	7542	0.255	0.092	0.180	0.107	0.089	0.119
Growth between grades 1 and 2							
Boys and girls combined	7542	0.068	0.065	0.086	0.035	0.055	0.053
Boys only	3829	0.099	0.084	0.118	0.067	0.073	0.073
Girls only	3713	0.086	0.069	0.081	0.042	0.051	0.052
Difference between boys and girls	7542	0.122	0.085	0.141	0.077	0.089	0.098
Prospects Cohort 3 (1991-1993)							
Grade 3 reading achievement							
Boys and girls combined	11017	0.071	—	—	—	—	—
Boys only	5553	0.105	—	—	—	—	—
Girls only	5464	0.072	—	—	—	—	—
Difference between boys and girls	11017	0.130	—	—	—	—	—
Grade 5 reading achievement							
Boys and girls combined	11017	0.071	0.055	0.101	0.065	0.106	0.068
Boys only	5553	0.083	0.087	0.171	0.101	0.186	0.127
Girls only	5464	0.085	0.059	0.087	0.051	0.062	0.043
Difference between boys and girls	11017	0.111	0.109	0.210	0.110	0.202	0.136
Growth between grades 3 and 5							
Boys and girls combined	22034	0.099	0.159	0.237	0.147	0.228	0.143
Boys only	11106	0.156	0.295	0.416	0.251	0.372	0.195
Girls only	10928	0.133	0.113	0.171	0.048	0.127	0.123
Difference between boys and girls	11017	0.218	0.341	0.491	0.242	0.393	0.251

+ = Multiple factors include achievement (average of math and reading scores from initial grade), parental socioeconomic status, and Census region

— = Not applicable because of absence of information on math or reading achievement prior to the initial grade for each sample

NOTE: Difference between group X and group Y refers to the mean for group X minus the mean for group Y. Standard errors have been adjusted for complex survey designs. 8th Grade SDUs = eighth-grade standard deviation units; 8th grade score is normalized to have mean of zero and standard deviation of 1.00. Multiple imputation used to account for missing data.

SOURCE: U.S. Department of Education, Chapter 1 Prospects Study.

Table D.12 Standard errors for Table 4.5: Black-white differences in average reading achievement in grades 7 to 9 and 10 to 12 for 2 samples of children, 1990-1993

Description of source of data, grade and outcome, and sex of subgroup	Sample size	S.E. of black-white difference adjusting for			S.E. of change in black-white difference		
		(1) no variables	(2) achievement only	(3) multiple factors+	difference between (2) and (1)	difference between (3) and (1)	difference between (3) and (2)
Prospects Cohort 7 (1991-1993)							
Grade 7 reading achievement							
Boys and girls combined	6654	0.066	—	—	—	—	—
Boys only	3419	0.079	—	—	—	—	—
Girls only	3235	0.086	—	—	—	—	—
Difference between boys and girls	6654	0.114	—	—	—	—	—
Grade 9 reading achievement							
Boys and girls combined	6654	0.066	0.060	0.104	0.047	0.078	0.067
Boys only	3419	0.099	0.096	0.187	0.064	0.138	0.126
Girls only	3235	0.084	0.085	0.101	0.072	0.070	0.056
Difference between boys and girls	6654	0.136	0.132	0.222	0.094	0.155	0.131
Growth between grades 7 and 9							
Boys and girls combined	6654	0.066	0.063	0.104	0.038	0.070	0.067
Boys only	3419	0.096	0.096	0.187	0.044	0.127	0.126
Girls only	3235	0.082	0.073	0.097	0.041	0.040	0.050
Difference between boys and girls	6654	0.132	0.119	0.219	0.041	0.128	0.138
NELS:88 (1990-1992)							
Grade 8 reading achievement							
Boys and girls combined	12954	0.048	—	—	—	—	—
Boys only	6409	0.073	—	—	—	—	—
Girls only	6545	0.057	—	—	—	—	—
Difference between boys and girls	12954	0.091	—	—	—	—	—
Grade 12 reading achievement							
Boys and girls combined	12954	0.070	0.058	0.066	0.049	0.055	0.051
Boys only	6409	0.100	0.061	0.103	0.076	0.085	0.084
Girls only	6545	0.094	0.079	0.077	0.070	0.060	0.064
Difference between boys and girls	12954	0.138	0.083	0.139	0.108	0.105	0.066
Growth between grades 8 and 12							
Boys and girls combined	12954	0.236	0.254	0.181	0.074	0.149	0.160
Boys only	6409	0.205	0.255	0.293	0.165	0.245	0.209
Girls only	6545	0.370	0.306	0.248	0.126	0.196	0.115
Difference between boys and girls	12954	0.376	0.311	0.384	0.177	0.223	0.162

+ = Multiple factors include achievement (average of math and reading scores from initial grade), parental socioeconomic status, and Census region

— = Not applicable because of absence of information on math or reading achievement prior to the initial grade for each sample

NOTE: Difference between group X and group Y refers to the mean for group X minus the mean for group Y. Standard errors have been adjusted for complex survey designs. 8th Grade SDUs = eighth-grade standard deviation units; 8th grade score is normalized to have mean of zero and standard deviation of 1.00. Multiple imputation used to account for missing data.

SOURCE: U.S. Department of Education, Chapter 1 Prospects Study, and National Educational Longitudinal Study of 1988 (NELS:88).

Table D.13 Standard errors for Table 4.6: Black-white differences in average reading achievement in grades 2, 5, 9, and 12 for 4 samples of children, 1992-1993

Description of sample(s), year(s) of data, and grades for which differences are calculated	Standard error of difference between blacks and whites (in 8th grade SDUs) for				
	Sample size	boys and girls combined	boys only	girls only	difference of boys and girls
Grade 2—Prospects Cohort 1 data (observed 1993)	4474	0.21	0.21	0.20	0.15
Grade 5—Prospects Cohort 3 data (observed 1993)	4224	0.10	0.13	0.12	0.15
Grade 5 reading gap minus Grade 2 reading gap	8698	0.21	0.25	0.20	0.21
Grade 5—Prospects Cohort 3 data (observed 1993)	4224	0.10	0.13	0.12	0.15
Grade 9—Prospects Cohort 7 data (observed 1993)	2332	0.09	0.13	0.12	0.17
Grade 9 reading gap minus Grade 5 reading gap	6556	0.13	0.18	0.17	0.22
Grade 9—Prospects Cohort 7 data (observed 1993)	2332	0.09	0.13	0.12	0.17
Grade 12—NELS:88 data (observed 1992)	9571	0.05	0.08	0.06	0.09
Grade 12 reading gap minus Grade 9 reading gap	11903	0.10	0.15	0.13	0.20
Grade 2—Prospects Cohort 1 data (observed 1993)	4474	0.21	0.21	0.20	0.15
Grade 9—Prospects Cohort 7 data (observed 1993)	9571	0.05	0.08	0.06	0.09
Grade 9 reading gap minus Grade 2 reading gap	6806	0.23	0.25	0.23	0.23
Grade 2—Prospects Cohort 1 data (observed 1993)	4474	0.21	0.21	0.20	0.15
Grade 12—NELS:88 data (observed 1992)	9571	0.05	0.08	0.06	0.09
Grade 12 reading gap minus Grade 2 reading gap	14045	0.22	0.23	0.21	0.17

NOTE: Difference between group X and group Y refers to the mean for group X minus the mean for group Y. Standard errors have been adjusted for complex survey designs. 8th Grade SDUs = eighth-grade standard deviation units; 8th grade score is normalized to have mean of zero and standard deviation of 1.00. Multiple imputation used to account for missing data.

SOURCE: U.S. Department of Education, Chapter 1 Prospects Study, and National Educational Longitudinal Study of 1988 (NELS:88).

APPENDIX E: THE SENSITIVITY OF BLACK-WHITE DIFFERENCES IN HIGH SCHOOL COMPLETION TO VARIABLE AND SAMPLE DEFINITIONS

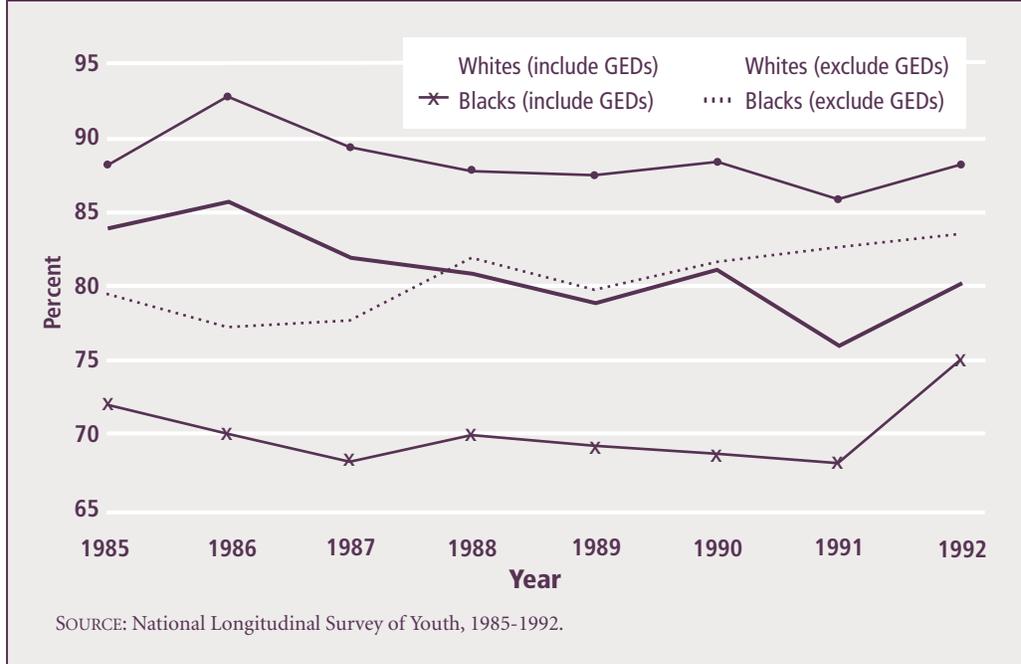
The analyses in this appendix, based on two samples of young adults,¹ investigate the sensitivity of black-white differences in high school completion to variable and sample definitions. First, evidence is presented on whether the black-white gaps in high school completion are different from black-white gaps in high school/GED completion. Second, evidence is presented on whether black-white gaps in high school/GED completion rates differ when prisoners and other institutionalized persons are included in the sample.

I. HIGH SCHOOL COMPLETION RATES WHEN GEDS ARE NOT COUNTED AS HIGH SCHOOL EQUIVALENTS

When comparing black-white gaps in high school completion rates, it is important to recognize the possible sensitivity of results to variable definitions, and in particular, to the treatment of General Educational Development certificates (GEDs) as equivalent to high school diplomas. As noted in chapter 1, research by Cameron and Heckman (1991) suggests that high school diplomas and GEDs differ in many important ways, such that GEDs may not represent true equivalents to high school diplomas.

¹ The first sample used for these analyses is a sample of young adults in the National Longitudinal Survey of Youth, observed between 1985 and 1992, during the calendar year when they turn 28. (Note that this sample was not "freshened" with recent immigrants to the United States, so it does not represent the full population of 28-year-olds in each year.) The second sample used for these analyses is a sample of 25- to 34-year-olds in the National Adult Literacy Survey, observed in 1992.

Figure E.1 High school completion rates for 28-year-olds (distinguishing high school diplomas from GEDs)



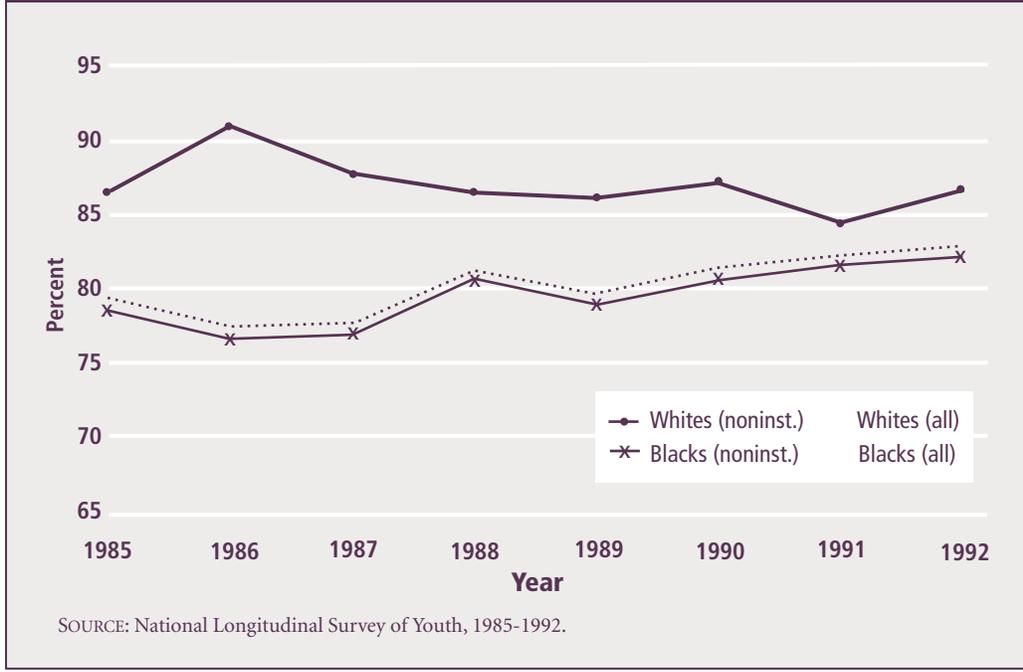
Distinguishing high school diplomas from GEDs alters the magnitude, but not the trend, in the black-white gap in high school completion between 1985 and 1992 (figure E.1). When GEDs are no longer counted as equivalent to a high school diploma, the average black-white gap in high school completion rates over this period increases by 36 percent, from 8 percentage points to 11 percentage points. However, black-white gaps in high school completion rates narrowed at about the same annual rate between 1985 and 1992, regardless of whether GEDs are treated as true high school equivalents. Over this period, the black-white gap in the completion of high school diplomas or GEDs by 28-year-olds narrowed at an average rate of 1.2 percentage points per year, while the black-white gap in the completion of high school diplomas narrowed at an average rate of 1.1 percentage points per year.

II. HIGH SCHOOL/GED COMPLETION RATES INCLUDING INSTITUTIONALIZED PERSONS

Another factor that may influence reported trends in high school completion is the exclusion of institutionalized persons from the sample over which high school completion is being calculated. Most statistics on the high school completion rates of blacks and whites are calculated using household or school-based surveys that exclude institutionalized persons, such as prisoners. If blacks represent a disproportionate share of the institutionalized population, and institutionalized persons are less likely to have high school diplomas (or GEDs), statistics on high school completion reported from civilian, noninstitutional surveys may understate the true magnitude of black-white gaps in the completion rates of high school diplomas and GEDs.

Analysis of high school/GED completion rates in two samples reveals that including prisoners or other institutionalized persons in the sample changes black-white gaps in

Figure E.2 High school/GED completion rates for 28-year-olds (with and without institutionalized persons)



high school/GED completion rates by very little. While black youth and young adults are more likely than white youth and young adults to be incarcerated, and prisoners are less likely than nonprisoners to have high school diplomas or GEDs, the proportion of the black population that is in prison is about 5 percent for 16- to 24-year-old blacks and about 6 percent for 25- to 34-year-old blacks. Because prisoners represent only a small fraction of the total population, the black-white gap in high school/GED completion rate is about the same (11 percentage points for 16- to 24-year-olds and 12 percentage points for 25- to 34-year olds), regardless of whether prisoners are included in the sample.

Including prisoners and other institutionalized persons in the sample does little to alter trends in black-white differences in high school/GED completion between 1985 and 1992 (figure E.2). Between these two years, the black-white gap in high school/GED completion rates of 28-year-olds narrowed at a rate of approximately 1.2 percentage points per year for the noninstitutionalized population, and at a rate of approximately 8.0 percentage points per year for the institutionalized population. While these two rates are noticeably different, the fact that institutionalized persons represent a small fraction of the overall population means that the rate of convergence in the high school/GED completion rates of all 28-year-old blacks and whites was about the same as for the non-institutionalized population only: 1.2 percentage points per year.

United States
Department of Education
ED Pubs
8242-B Sandy Court
Jessup, MD 20794-1398

Official Business
Penalty for Private Use, \$300

Postage and Fees Paid
U.S. Department of Education
Permit No. G-17

Standard Mail (A)

