

Appendix Table 1: Impact of Alternative Income Distribution Measures on Boys' Likelihood of Dropping Out of High School, by Socioeconomic Status (MSA Level)

	50/10 ratio (1)	90/50 ratio (2)	10 th Percentile of Income (in \$10,000s) (3)	50 th Percentile of Income (in \$10,000s) (4)	Income Share top 1% (5)
Correlation between 50/10 ratio and characteristic:		0.57	-0.65	-0.14	0.39
50/10 Ratio*	0.073	0.071	0.068	0.073	0.069
Mom HS Dropout	(0.018)	(0.024)	(0.027)	(0.019)	(0.021)
50/10 Ratio*	0.028	0.018	0.037	0.029	0.025
Mom HS Graduate	(0.016)	(0.018)	(0.021)	(0.015)	(0.018)
State Characteristic*	---	0.009	-0.002	-0.0002	0.002
Mom HS Dropout	---	(0.056)	(0.005)	(0.001)	(0.002)
State Characteristic*	---	0.045	0.002	0.0002	0.002
Mom HS Graduate	---	(0.039)	(0.004)	(0.001)	(0.002)

Notes: Reported standard errors (in parenthesis) are adjusted for clustering at the state level. Additional explanatory variables in each regression include maternal educational attainment, gender (where appropriate) race/ethnicity, and an indicator variable for living with a single parent at age 14, along with state and cohort fixed effects. The sample includes data from the ELS, NLSY79, and NLSY97. The total sample size is 11,013 boys.

Appendix Table 2: Impact of Measures of Segregation on Boys' Likelihood of Dropping Out of High School, by Socioeconomic Status (MSA Level)

	50/10 ratio (1)	Racial Segregation Index (2)	Income Segregation Index (3)	Poverty Segregation Index (4)
Correlation between 50/10 ratio and characteristic:		0.52	0.31	0.36
50/10 Ratio* Mom HS Dropout	0.073 (0.018)	0.045 (0.022)	0.066 (0.018)	0.065 (0.018)
50/10 Ratio* Mom HS Graduate	0.028 (0.016)	0.026 (0.020)	0.024 (0.017)	0.024 (0.018)
State Characteristic* Mom HS Dropout	--- ---	0.267 (0.116)	0.570 (0.372)	0.640 (0.460)
State Characteristic* Mom HS Graduate	--- ---	0.027 (0.084)	0.352 (0.302)	0.341 (0.349)

Notes: see notes to Appendix Table 1. Interacted state characteristic is listed in column headers.

Appendix Table 3: Impact of Potential Mediating/Confounding Factors on Boys' Likelihood of Dropping Out of High School, by Socioeconomic Status (MSA Level)

	50/10 ratio (1)	Fraction Single Parent (5)	Fraction Employed in Manufacturing (5)
Correlation between 50/10 ratio and characteristic:		0.56	-0.09
50/10 Ratio* Mom HS Dropout	0.073 (0.018)	0.059 (0.027)	0.074 (0.018)
50/10 Ratio* Mom HS Graduate	0.028 (0.016)	0.008 (0.018)	0.029 (0.015)
State Characteristic* Mom HS Dropout	--- ---	0.418 (0.409)	-0.115 (0.177)
State Characteristic* Mom HS Graduate	--- ---	0.550 (0.268)	-0.112 (0.127)

Notes: see notes to Appendix Table 1. Interacted state characteristic is listed in column headers.

**DATA APPENDIX:
MEASURING EDUCATIONAL ATTAINMENT IN NLSY AND NCES DATA**

This data appendix provides further details regarding the specific samples used in our analysis. All calculations performed include the sample restrictions described in the text, where we indicate that respondents whose educational attainment by age 20 is unknown and those whose mother's educational attainment is unknown are not included in the sample.

I. NATIONAL LONGITUDINAL SURVEYS OF YOUTH

A. 1979 Cohort

This data source originally surveyed 12,686 respondents born between 1957 and 1964, who were between the ages of 14 and 22 on the first survey date in 1979. The sample is not nationally representative, but sample weights are available to provide national representative estimates. Retention rates have been very high in these data, reducing the likelihood of attrition bias, particularly over relatively short periods. Respondents were re-interviewed every year through 1994 and then every other year after that. Because the NLSY is not a school-based survey, the universe of respondents is not restricted to those currently enrolled in a certain grade, as in the NCES data sources described below. On the other hand, some respondents are older than mandatory schooling ages on the initial survey and report their ultimate educational attainment and the timing of its completion retrospectively, introducing the possibility of recall bias.

B. 1997 Cohort

These data include information on 8,984 respondents who were born between 1980 and 1984, making them 12 to 18 on the first survey date. The sample is not nationally representative, but weights are available to provide nationally representative estimates.

Retention rates have been very high in these data, reducing the likelihood of attrition bias, particularly over relatively short periods. Respondents have been re-interviewed every other year since 1997 with the most recent available survey having been completed in 2011. Relative to the NLSY79, these data have the advantage that virtually all students are still in school at the time of the initial survey, so we can more reliably track their high school degree status as they age.

II. NATIONAL CENTER FOR EDUCATIONAL STATISTICS DATA

A. High School and Beyond (HSB)

HSB initially surveyed high school sophomores and seniors in the spring of 1980; we restrict our attention to the sophomores, most of whom were around 16 years old in that year. Respondents were re-interviewed every two years through 1986 and then again in 1992. HSB is a school-based survey; specific schools were selected to participate and the survey was administered to several students within the school. Over 30,000 sophomores in 1,015 high schools were surveyed in 1980. Of the original sample, half were selected to participate in the follow-up surveys and 79 percent responded to the follow-up survey. We restricted our analysis to those students who also participated in the base year survey.

We measure respondents' educational attainment in the second follow up, conducted in the spring and summer of 1984, when the respondents would have been around 20 years old. The second follow-up survey asks a direct question about whether respondents had graduated from high school. Respondents could have reported in response that they had graduated, had left school, were still enrolled in school, or whether they had earned a GED. For those still enrolled in school in the second follow-up, we code them as not having completed their degree by age 20 (i.e. as a "drop out").

B. National Education Longitudinal Survey (NELS)

NELS initially surveyed 8th graders in the spring of 1988, when most of them were 14 years old. They were re-interviewed in 1990, 1992, 1994, and 2000. In total, 14,915 respondents were interviewed initially in 1988 and again in the 1994 round, which represents the point at which we measure educational outcomes. Survey responses regarding educational attainment were recorded in each of these survey years and a subsample of these responses were checked against transcript records indicating their accuracy. The survey excluded 5.4 percent of selected students in the base year “because of physical or mental disabilities, or because of limited English language proficiency” (Ingels and Quinn, 1996). This restriction introduces sample selection bias since these students are more likely to drop out of high school subsequently.

The sample was “freshened” in subsequent surveys so that representative estimates could be drawn for the sophomore class in 1990 and the senior class in 1992. We focus on those respondents surveyed in the base year because using respondents from the refreshed sample would introduce an upward bias in measures of educational attainment in these data. Those students who have made it to their sophomore or senior years are a positively selected group of students, as we discuss below.

C. Education Longitudinal Survey of 2002 (ELS)

This survey included students who were in 10th grade in the spring of 2002. Students were re-surveyed in 2004 and 2006, so that they are around 20 years old in the latest year of available data. There were 15,300 students who responded to both the base year survey and the 2006 survey, when educational outcomes were measured.

III. DESCRIPTIVE STATISTICS

As Heckman and LaFontaine (2010) make clear, comparing educational attainment statistics from various micro datasets is a difficult task because of the idiosyncrasies of each. It is not our goal to track trends using these datasets, as they did, so we do not try to reconcile these differences. We do report statistics on educational attainment from each of them, though, for the purposes of detailing those differences and comparing the calculated statistics with outside sources for verification. In our econometric analysis, we control for these differences by including “dataset fixed effects.”

Appendix Table 1 indicates the percentage of survey respondents who graduate from high school, receive a GED, or drop out of high school in each dataset. Sample weights are used to adjust for the various sampling techniques used in each dataset. Discrepancies across datasets are extensive, yet they are consistent with past estimates (allowing for modest variation attributable to the sample restrictions we impose). For instance, Hill and Holzer (2007) examine data from the two NLSY surveys. We focus on educational attainment by age 20 and they focus on educational outcomes between 20 and 22. We find that 16.3 and 12.2 percent dropped out of high school and 5.1 percent and 6.9 percent have a GED in the 1979 and 1997 surveys, respectively. Their results are comparable: 16.8 percent and 12.8 percent dropped out and 4.3 percent and 5 percent earned a GED in the respective surveys. For HSB, we are able to replicate reported results (National Center for Educational Statistics, 1984), although our sample restrictions change the final values reported in this table somewhat. We match previous estimates because we are coding high school completion status directly for a single survey question. In the NELS, we estimate that 9.3 percent of students drop out and 5.0 percent of students obtain a GED by around age 20. Our estimates are comparable to those in Hurst, et al.

(2004), who find that 12 percent of students drop out and 6 percent of students obtained a GED by 1994 (when most respondents are age 20). In the ELS, we estimate that 7.5 percent of students drop out of high school and 4.3 percent obtain a GED by around age 20. This compares to 7.8 percent and 3.9 percent, respectively, reported in Bozick, et al. (2007).

It is difficult to determine the extent to which the differences in estimates across datasets are attributable to changes in outcomes over time or the differences in the nature of the surveys. An important difference is the sampling strategies used by the different surveys. The three NCES surveys are school-based and require students to be still enrolled in school to participate. This is particularly troublesome with the HSB and ELS surveys, in which youth need to “survive” to 10th grade to participate. In the NELS, students only need to “survive” to 8th grade, which is less likely to introduce bias. Nevertheless, these sampling strategies indicate that we should expect higher dropout rates in the two NLSY surveys, which is exactly what we see. It would also be reasonable that NELS had the next highest dropout rate and that hypothesis is confirmed in these data as well.

To better document this problem, we use data from the two NLSY surveys to examine the degree status of students and their highest grade completed at age 20. The results are presented in Appendix Table 2. In the NLSY79 and NLSY97, 5.1 percent and 6.9 percent of respondents, respectively, never make it to 10th grade by age 20. Of those who fail to reach that grade, most drop out. Omitting those students from the sample, as occurs in the HSB and ELS, imposes an upward bias in educational attainment. Indeed, this is a problem, albeit considerably smaller, even when starting a sample in 8th grade, as occurs in the NELS. Around half a percent of students fail to reach that grade in the two NLSY surveys.

Data Appendix Table 1: Educational Attainment Measured
in Alternative Longitudinal Data Sources.

Educational Attainment by Age 20			
	GED	High School Dropout	High School Graduate
NLSY79	5.1	16.3	78.6
HSB (1980)	3.8	7.1	89.2
NELS (1988)	5.0	9.3	85.7
NLSY97	6.9	12.2	81.0
ELS (2002)	4.3	7.5	88.3

Source: Authors' calculations.

Data Appendix Table 2: Degree Status by Highest Grade Completed at Age 20

	Below 8 th Grade	8 th Grade	9 th Grade	10 th Grade	11 th Grade	12 th Grade and Higher
NLSY79						
Percent at Level	0.7	1.6	2.8	4.3	5.0	85.6
<u>Degree Status:</u>						
HS Dropout	99.4	97.4	95.7	95.5	89.5	3.3
GED	0.0	1.3	3.1	4.4	4.8	5.4
HS Graduate	0.6	1.3	1.3	0.1	5.8	91.4
NLSY97						
Percent at Level	0.4	2.6	3.9	5.1	6.0	82.0
<u>Degree Status:</u>						
HS Dropout	93.9	73.7	63.6	58.1	59.5	1.0
GED	4.6	23.2	33.5	38.8	33.5	1.2
HS Graduate	1.6	3.1	2.9	3.2	7.0	97.8

Source: Authors' calculations.

SOURCES:

Bozick, Robert, Erich Lauff and John Wirt (2007). *Education Longitudinal Study of 2002 (ELS:2002): A First Look at the Initial Postsecondary Experiences of the High School Sophomore Class of 2002*. Washington, DC: National Center for Education Statistics.

Haskins, Ron. 2008. "Education and Economic Mobility." In *Getting Ahead or Losing Ground: Economic Mobility in America*, edited by Julia B. Isaacs, Isabel V. Sawhill, and Ron Haskins. Washington, DC: The Brookings Institution and the Economic Mobility Project.

Heckman, James J. and Paul A. LaFontaine (2010). "The American High School Graduation Rate: Trends and Levels," *Review of Economics and Statistics* 92 (2): 244-262.

Hill, Carolyn J. and Harry J. Holzer (2007). "Labor Market Experiences and the Transition to Adulthood," in Sheldon Danziger and Cecilia Rouse (eds.) *The Price of Independence: The Economics of Early Adulthood*. New York: Russell Sage Foundation.

Hurst, David, Dana Kelly, and Daniel Princiotta (2004). *Educational Attainment of High School Dropouts 8 Years Later*. Washington, DC: U.S. Department of Education, National Center for Educational Statistics.

Ingels, Steven J., and Peggy Quinn (1996). *Sample Exclusion in NELS:88: Characteristics of Base Year Ineligible Students; Changes in Eligibility Status after Four Years [National Education Longitudinal Study of 1988 Second Follow-Up]*. Washington, DC: U.S. Department of Education. National Center for Education Statistics.

National Center for Education Statistics (1984). *High School and Beyond, 1980: Sophomore Cohort Second Follow-up (1984). Volume I*. Washington, DC: Government Printing Office.