ECONOMIC POLICY INSTITUTE
1333 H STREET, NW • SUITE 300, EAST TOWER • WASHINGTON, DC 20005
PHONE: 202.775.8810 • FAX: 202.775.0819

Questions and Answers on Measurement of High School Graduation Rates and Trends

by Lawrence Mishel and Joydeep Roy

April 27, 2006

What data did you analyze in your study?

Some commentators have been describing our report, *Rethinking High School Graduation Rates and Trends* (http://www.epi.org/content.cfm/book_grad_rates) as primarily relying on the Current Population Survey (CPS), which is the monthly labor force survey. In fact, we have assessed graduation rates in every source we could identify. In all cases we develop estimates of the graduation rate with regular diplomas, excluding GEDs. The various sources of data include:

National longitudinal surveys

The first type of data we examined was from national longitudinal surveys—which follow students over time. We identify the National Education Longitudinal Study (NELS) as the very best source of data because it tracks individual students over time and verifies students' graduation status with transcripts independently obtained from schools. The NELS began with 8th graders in the spring of 1988 and followed them over the next 12 years, with interviews at regular intervals in 1990, 1992, 1994, and 2000. The NELS is considered by us and most researchers as the "gold standard." The NELS shows that by 1994, two years after their normal or on-time graduation date, slightly over 82% of all students had completed high school with a regular diploma. This number was 85% for whites, 95% for Asians, 74% for blacks, and 74% for Hispanics. By 2000, when most of these people were of age 26, an additional 7% of whites, 14% of blacks, and 9% of Hispanics had acquired a GED diploma, which facilitates access to college and the military. These results alone are enough to question the new conventional wisdom that only two-thirds of all students, and only half of all minorities, graduate from high school with a diploma (the conventional wisdom estimates are based, as we explain below, on diploma counts that include fifth- and sixth-year diplomas notwithstanding the fact that the conventional estimates are often inaccurately described as "on-time" graduation rates).

The NELS results are confirmed by the NLSY97 and the NLSY79—large scale longitudinal surveys, called the National Longitudinal Survey of Youth, conducted by the Bureau of Labor Statistics. Chen and Holzer (2006) show that, for those ages 20-22 in 2002 (including those in prison), the overall graduation rate is the same as in NELS: 82% overall, 75% for blacks, and 76% for Hispanics. These data also show an improvement in graduation rates for every race and gender group since 1984, except for black men. The improvements are particularly large and significant for Hispanics, both males and females.

Household surveys

The second type of data we explore is household surveys, which allow us to identify the share of the population at a particular age that has graduated from high school. The Current Population Survey (CPS) results have been questioned because they exclude the institutional (including prison) population and because they undercount low-income blacks. Therefore, we

analyzed the decennial census data for 2000. These data include people in institutions (inmates of prisons, nursing homes, hospitals, etc.) and the military and miss very few people in the sample (recall that these are the data that determine the number of congressional seats for each state and therefore their collection was the focus of great effort and attention). We find that these Census data yield graduation rates very similar to those in the NELS. We also use the CPS and adjust for the increased incarceration rates of black men (the only demographic group where this omission creates a bias). Again, the results yield graduation rates far in excess of those claimed in the new conventional wisdom.

Enrollment and diploma data

The new conventional wisdom, based largely on analyses of Jay P. Greene of the Manhattan Institute and University of Arkansas, and of Christopher Swanson, formerly of the Urban Institute and now with *Education Week*, relies on a third type of data that consists of the enrollment and diploma counts presented in the Common Core of Data, or CCD for short. The CCD provides enrollment data for particular grades and diploma counts (all diplomas, on-time or not) for particular years. Graduation rates can be estimated by constructing artificial "cohorts," the simplest method being to examine diplomas in one year and divide by the 9th grade enrollment three years earlier. There are different answers one can get from the CCD depending upon how one constructs the cohort. These choices matter, particularly for estimating the graduation rates for minorities, as we explain below. Specifically, we show that Drs. Greene and Swanson make unexamined choices in constructing these cohorts that result in dramatic under-estimates of the true graduation rates.

What are the challenges with using the enrollments and diploma in the Common Core of Data (CCD) to estimate graduation rates?

The enrollment and diploma counts in the CCD are the only data available at the state and local areas, so it is not surprising that researchers have tried to compute graduation rates with these data. We believe there are important limitations to any computation of graduation rates using the CCD, setting aside any question of the quality and completeness of the data. *This is because the CCD does not measure high school graduation rates of entering ninth graders*. Consequently, researchers must estimate graduation rates by constructing "cohort graduation rates" based on enrollment and diploma data for particular years. There are several data limitations that frustrate this effort:

1. ***Diploma Counts****:* It is not generally understood that the diploma counts in the CCD include all diplomas, on-time or not, even though some people refer to the rates calculated using the CCD as "on-time" rates.

2. ***Entering Ninth Graders:*** It seems to be generally acknowledged that the graduation rate should reflect how many entering ninth graders complete high school with a diploma. Unfortunately, the CCD does not report *entering ninth graders*; rather, it reports *ninth grade enrollment*, including students who are repeating 9th grade and who entered 9th grade a year or more earlier This is an important distinction because there is substantial retention of students, particularly minorities, in 9th grade and sometimes in 10th grade. We find that for the nation as a whole there are 12-13% more students in 9th grade in public schools than in the 8th grade in the previous year; for blacks and Hispanics the rate is more like 25%. Since retention is larger for some demographic groups, in some states the method for accounting for retention—or not doing so—can greatly affect racial comparisons and state comparisons.

3. ***Transfers:*** The diploma counts include diplomas that are earned by students who transferred into a school/district/state. Consequently, graduation rates can be distorted in areas where there are substantial increases or decreases in the student population. Some computations using the CCD do not account for this while others do. The results of adjusting for transfers in and out (not present with longitudinal student data) are problematic and making no adjustment is problematic.

4. ***Diploma Definitions:*** Each state defines what it means to graduate with a diploma, and that definition can change over time. It is important to be careful when making comparisons across states to have as consistent a definition as possible. An added complication is the fact that several states have different categories of completion as reported in the CCD. For example, while in states like Georgia, Oregon, and Alabama, the share of the category called "other high school completers," which does not include GEDs— which includes those with certificates of attendance or certificates of completion—is more than 9% of all completers, in states like California, Illinois, and Massachusetts, there are no completers in this category.

The point is that the CCD does not provide the measure that we seek: the graduation rate of entering 9th graders, either on-time or eventual/final. These problems do not invalidate the use of the CCD, but acknowledging them is important. The CCD is certainly not a data set that can be described as a "census" of what we want measured. At best, the CCD is a 'census' of enrollment and diplomas but these are just ingredients in a graduation formula that reflects many choices to address the limitations of the CCD.

What are the problems with the Greene and Winters methodology?

There are a number of problems with the Greene and Winters method, all of which flow from choices they make to overcome inherent weaknesses in measuring graduation rates based on enrollment and diploma counts in particular years. The first issue involves "retention bulge" (see below) and arises because the CCD does not have data on "entering ninth graders." Second, the issue of "population adjustment" becomes problematic because the CCD provides no information on transfers in and out of a state or district, yet the diploma count includes transfer students. This problem doesn't arise with longitudinal data, which follow students over time in a national sample.

Retention bulge

Greene's initial methodology in a report released in November 2001 (revised in April 2002) compared diplomas to 8th-grade enrollment four years earlier. As such, the resulting graduation rates are not distorted by retention rates or the ninth-grade bulge (if a student is retained in some later grade and ultimately gets a diploma, his or her graduation will count in some year, but if there is no increase/decrease in retention, then there should not be a bias in the trend). Greene and Winter's current methodology acknowledges the retention problem in 9th grade and claims to adjust for it by using the average of 8th, 9th, and 10th grade enrollments of a cohort as their estimate for entering 9th graders. This change in methodology actually introduces a distortion, especially for minorities, even if the distortion is not as large as simply relying on 9th grade enrollment. Greene and Winter's estimate of entering 9th graders exceeds the number of 8th graders in the previous year by more than 10% for Hispanics and more than 8% for blacks—a gap unexplainable by transfers of students from private middle schools to public high schools between the 8th and 9th grades. We show below what the impact of the retention bulge is on Greene's estimates. By the way, Swanson's reliance on changes in enrollment between 9th grade and every ensuing year generates an even larger distortion. Miao and Haney's (2004) measure of high school completion is the simple diploma to 8th grade ratio—called the basic completion ratio—and is not affected by the retention bulge. Warren (2005) shows in his study, based on actual retention rates from several states (Texas, Massachusetts, North Carolina), that the best predictor of the number of this year's entering 9th graders is the number of last year's 8th graders.

The choice of denominator (8th grade enrollment, 9th grade enrollment; or Swanson's use of enrollment erosion each year starting from 9th grade) can greatly affect measured rates, with the Swanson and Greene measures yielding the lowest graduation rates. This is especially true for minorities and in some states, such as Texas, where retention rates are high. For example, in Texas, the retention rate for black students is roughly 25.0% in grade 9 and 12.5% in grade 10.

Population adjustments

Another important reason that Greene's estimates in particular diverge so much from most other results is his use of a population adjustment in scaling up his estimates of freshmen 9th graders. Thus, for example, Greene adjusts his denominator (reflecting enrollment in the 8th, 9th, and 10th grade) for Hispanics upwards by 21% and Asians upwards by 25% based on the increase in the populations of these respective groups as they advance from 14 years old to 17 years old. For estimates at a district or state level, one could argue that such an adjustment reflects net transfers. However, at the national level such an adjustment solely reflects immigration at ages 15, 16, and 17. The implicit assumption is

that these recent immigrants can be expected to enroll and graduate at the same pace as students who were present in the United States for their entire schooling. Any failure of these recent immigrants to graduate is labeled a school failure by those using Greene's method.

Whatever one thinks of this population adjustment at the national or local level, it is clear that the adjustment is hugely important, as seen in Table 1:

Table 1: Greene's population adjustments

		Graduation rate		
	Population change	**Unadjusted**	**Adjusted**	**Population adjustment**
Whites	2%	79%	78%	-2%
Blacks	8%	59%	55%	-4%
Hispanics	21%	64%	53%	-11%
Asians	25%	90%	72%	-18%
Total	7%	75%	70%	-5%

The population adjustment results in a lowering of the Hispanic and Asian rates by 11 and 18 percentage points, respectively. In contrast, the white and black rates are impacted by 2 and 4 percentage points. Thus, this adjustment greatly affects the rates for some groups and any comparisons with whites, who are largely unaffected. The population adjustment lowers the national graduation rate by 5 percentage points.

How to address the measurement problem posed by immigration is unclear, and how various methods bias the results is also unclear. If any immigrants actually obtain a diploma (counted in the CCD), then those measures that rely on 8th grade enrollment may be overstating graduation rates. In contrast, Greene's adjustment will greatly understate graduation if very few of these immigrants receive diplomas, at least if one seeks to measure the graduation success of entering 9th graders.

Without knowing how many of these immigrants even enroll in U.S. high schools and how many of them obtain high school diplomas from U.S. high schools, it is not possible to gauge the bias from Greene's adjustment. This reveals that Greene's calculations embed a value judgment that schools should be judged as harshly for the failure of a student immigrating to the United States at age 15 as it is for a student immigrating at age 5 or one even born in the United States (we want to be clear that we are not dismissing the needs of recent immigrant students and the substantial help many of them need and should be provided, but as a measurement issue and a policy issue, they should not be treated as comparable to students enrolled at much earlier ages.) How one addresses this issue is by nature an arbitrary decision. If anything, this underscores the need for longitudinal studies like the NELS, which track individual students over time and record their graduation status.

Diploma definition

Not only does each state have different numbers and types of exit options, every state has its own definition for a "regular" diploma (see the NCES report of the Task Force on Graduation, Completion, and Dropout Indicators). Moreover, states often change the requirements for diplomas, as New York did in the last decade—requirements for earning a local diploma went up from 20.5 credits to 22 credits. None of this will be reflected in the CCD data that Greene, Swanson, and others use for comparing graduation rates across states and over time. The bottom line is that, without additional adjustments, which neither Greene nor Swanson make, you cannot do either a state-by-state or a year-by-year comparison with the existing CCD data.

Looking beneath the hood

Table 2 shows the difference between Greene's estimates of graduation and an estimate, such as Haney's, based on a simple ratio of diplomas to the 8th grade enrollment four years earlier. The results are presented by race/ethnic group and also for Texas so that we can illustrate the impact of the adjustments on measured graduation rates in differing locations and for differing demographic groups.

The Haney calculations show a graduation rate lower than the NELS, but the gap is much smaller—just 3-4 percentage points (82% versus 78%), though the gap for blacks is larger (12-13 percentage points). In contrast, the gap between NELS and Greene's estimates are 12 percentage points

overall and about 20 percentage points for blacks. This underlines that the choices that Greene and others have made in using the CCD greatly explains the much lower graduation rates they report.

Table 2: Greene's estimates and adjustments*

	Nation	Texas
Total		
Diploma/8th grade	78%	79%
Population adjustment	-5%	-7%
Bulge impact	-3%	-4%
Greene	70%	69%
Whites		
Diploma/8th grade	81%	83%
Population adjustment	-2%	-3%
Bulge impact	-2%	-1%
Greene	78%	78%
Blacks		
Diploma/8th grade	64%	76%
Population adjustment	-4%	-3%
Bulge impact	-5%	-6%
Greene	55%	67%
Hispanics		
Diploma/8th grade	70%	74%
Population adjustment	-11%	-10%
Bulge impact	-7%	-6%
Greene	53%	59%
Asians		
Diploma/8th grade	98%	108%
Population adjustment	-18%	-18%
Bulge impact	-8%	-9%
Greene	72%	81%

* Totals may not add due to rounding

The Greene adjustment for population growth and the bias imparted by the retention bulge(s) have a small effect on whites (three percentage points) but a large impact on every other group, ranging from a 26 percentage point impact on Asian graduation rates to a 17 percentage point adjustment for Hispanics and a smaller 9 percentage point adjustment for blacks. What one believes about the gaps in graduation between whites and other groups depends greatly on whether one accepts the Greene adjustments as appropriate. Whatever one's judgment on these adjustments, it should be clear that Greene's calculations are much more than a simple representation of a census of graduation rates reported by school districts to the states and to NCES. It is noteworthy that the Asian graduation rate in Texas with the diploma-to–8th-grade-enrollment ratio is an implausible 108%. This motivates the rationale for why a population adjustment may have merit for some groups. Whether there is no adjustment or an 18 percentage point adjustment (as in Greene) is completely arbitrary because there are no data to guide us to the proper adjustment. But regardless, it still clearly matters.

What can we learn from student longitudinal data developed by school districts?

We also examine places where schools have developed student longitudinal data (i.e., tracking the graduation of actual students). We compared these graduation rates to those obtained by Greene and

Swanson based on the enrollment and diploma counts from these same students. We find that the Greene and Swanson estimates are dramatically lower than the longitudinal graduation rates presented by the state of Florida or New York City but are not so off for Chicago (except Greene and Swanson's estimates get the trends wrong). This analysis is another confirmation that the Greene and Swanson estimates, which are the basis for the conventional wisdom, are erroneous.

What about the ability of surveys to identify poor students and those at the margins?

One possible critique of national data is that they may not properly sample low-income students, especially minorities, thereby leading to an overstatement of graduation rates. This is a serious issue with which any agency compiling data has to wrestle. In fact, the national longitudinal data upon which we rely has addressed this issue and minimized the problem. The Census data we analyze is the result of the massive effort in 2000 to assure as complete a Census as possible, with the results being the best census ever and one that is characterized by lower levels of under-representation than in any other census or survey.

National Education Longitudinal Study, 1988

Since one of the main objectives of the *National Education Longitudinal Study, 1988* (NELS) was to provide a valuable framework for the analysis of educational outcomes and academic persistence, a great deal of attention was paid to the issue of gathering information on dropouts and their educational outcomes. Among the most important objectives for the survey design were: first, dropouts were to be retained in the sample with near certainty to ensure an accurate examination of the dropping out phenomenon; second, to retain the maximum number of Hispanics, Asians, and American Indians; third, non-respondents in earlier follow-ups were to be retained with certainty in the next follow-ups to minimize non-response bias. Also, to achieve disproportionate retention of minority students, most of the schools containing these students were retained in the sample, particularly for the transcript survey in 1992.

Table 3 shows that the issue of non-response from dropouts—who are likely to come from the more-marginalized sections of society—was sufficiently small in the NELS to assure us of the validity of its results regarding high school completion. (The category "out of scope" refers to students who were either deceased or were out of the country.)

Table 3: Distribution of status of NELS:88 students in the first and second follow-ups

Base Year (1988)	First follow-up status (1990)	Second follow-up status(1992)	
		Categories	**Number** (N)
	Dropouts (N = 1,029)	Dropouts	611
		Alternative Completer or Alt. Student	222
		Student	69
		Out of Scope	9
		Status Unknown	118
	Students (N = 18,270)	Dropouts	1,041
		Alternative Completer or Alt. Student	542
		Student	16,339
		Out of Scope	82
		Status Unknown	266
Students (N = 20,062)	**Out of Scope** (N = 129)	Dropouts	11
		Alternative Completer or Alt. Student	6
		Student	11
		Out of Scope	83
		Status Unknown	18
	Status Unknown (N = 634)	Dropouts	58
		Alternative Completer or Alt. Student	20
		Student	466
		Out of Scope	6
		Status Unknown	84

The table clearly highlights that non-response was sufficiently small—in the first follow-up survey in 1990, the status of only 634 out of over 20,000 students (about 3%) was unknown. In the second follow-up, the status of only 486 students (about 2.5% of the initial sample) had unknown statuses. Note that most of the students whose status was unknown in 1990 could be contacted successfully in 1992—only 84 students, less than half of 1% of the initial sample, could not be contacted in either year.

A final point to note: about 5.35% of the initially sampled students were excluded in the base year because of language impairments (difficulty in completing survey questionnaire) or mental or physical disabilities. In later follow-up surveys, about 74% of those initially excluded due to language barriers were included, as were a significant number of those excluded due to different forms of disabilities.

Therefore, it is not accurate to say that the NELS estimates are significantly biased by the inability to track down dropouts. The great amount of time and resources spent to ensure adequate representation of dropouts and minorities—not only in the initial sample but in follow-up samples as well—suggests that the NELS estimates of high school completion are quite accurate, with minimal bias from under coverage of the more marginalized and vulnerable sections of the society.

National Longitudinal Study of Youth: 97 (NLSY97)

The NLSY97 consists of a nationally representative sample of approximately 9,000 youths who were 12 to 16 years old as of December 31, 1996. In round 1 of the survey, which took place in 1997, both the eligible youth and one of that youth's parents received hour-long personal interviews. Since then, youths have been interviewed on an annual basis. Educational data collected in NLSY97 include schooling history, performance on standardized tests, course of study, the timing and types of degrees, and a detailed account of progression through post-secondary schooling. In winter 1999-2000 high school transcripts, containing information on high school graduation status among other things, were obtained for NLSY97 respondents who were no longer enrolled in high school—those who had either graduated from high school or were age 18 or older and no longer enrolled in high school. The agency plans to collect additional transcripts as more respondents become eligible.

The NLSY97 involved the selection of two independent samples, one of which was designed to produce a significant over-sampling of Hispanic and non-Hispanic black youths. In a technical report available on the BLS Web site, the authors compare the percentages of Hispanics and non-Hispanic blacks in this NLSY97 sample with those in the 1990 census, and find that the difference (undercount) is only 0.2 percent for blacks and 1.5 percent for Hispanics. This, along with the fact that persons in a jail, prison or similar detention facility were included in the NLSY97 sampling universe, implies that the estimates of high school graduation from NLSY97 provide a reasonably accurate picture of the educational attainment of these cohorts.

2000 Census

A study of the 2000 census conducted by a committee set up for the purpose finds that the net undercounts for Hispanics and non-Hispanic blacks were about 2.85% and 2.17%, respectively. These numbers compare favorably with the 1990 census, when the net undercounts were 4.99% for Hispanics and 4.57% for blacks. Within specific age groups, the net undercount in the 2000 census is estimated to be around 6.49% for black males and 0.12% for black females in the 18-29 age group. As such, it is unlikely that under-coverage, particularly of the minority populations who are likely to have higher rates of dropout, significantly impacts our estimates from the decennial census.

If the CCD is like a “census,” then doesn’t that make it the best source of data?

The CCD collects data on every public school and public school district in the country. However, this fact alone does not make it particularly suitable for calculating graduation rates. The problem is that the CCD does not track individual students over time, as is explained above. The best one could do to calculate graduation rates using the CCD data is to compare the number of diplomas in a particular year, say 2006, over the number of entering 9th graders in the fall of 2002. This is problematic because, first, we cannot track students who drop out or join between the 9th grade and graduation—that is, we cannot account for leavers and joiners. This is particularly important if we are to calculate graduation rates at the state and school district levels, as low-income minority youths—whose graduation rates are of the greatest concern—also have the highest rates of mobility. Second, since the CCD data only has grade-specific *enrollments*, a researcher does not know the number of *entering* 9th graders. Estimating the number of first-time 9th graders based on 9th grade enrollments—or even a

smoothed average of 8th, 9th, and 10th grade enrollments as Greene and Winters have done—is problematic because of significant grade retentions in the 9th grade. The major point here is that having a larger sample, or "census," on enrollment and diploma counts does not necessarily provide accurate graduation rates because the CCD is not designed to do so.

The longitudinal studies, on the other hand, do not suffer from either of these problems. For example, the NELS begins with students in their 8th grade, and then follows them over the next 12 years. This gives us the correct rates of high school graduation, including rates of on-time completion and completion via alternative methods like the GED. (It also allows us to relate the problem of non-completion to the respective families' socioeconomic status and other family and school indicators.) The issue of grade retention or the transferring of schools does not affect the NELS results. This is the same for other national longitudinal surveys such as the NLSY. Because these surveys have samples that allow them to minimize sampling error and since they measure what is desired, they are clearly preferable to the CCD, which does not provide a measure of graduation for a fixed cohort of students even if it does survey each school district in the nation.

A "census" in this respect would constitute a national student identifier system, so that we could track every student, from his or her entering of 9th grade until he or she graduates or drops out. The NELS:88 is the big, representative sample version of this idea—for its sample, it does exactly what we would do for the universe if it were possible. In statistical terms, saying the CCD is preferred over the NELS because it is a "census" is to overvalue reducing sampling error while ignoring much worse nonsampling errors.

Last, labeling the CCD as a census overlooks the fact that there is only slight quality control and checking of the data provided by school districts to their states and by the states to NCES. Whether the questionnaire is completed as NCES expects and is done so consistently across districts and states and over time is not known because there are no audits done of school district respondents. In contrast, the national longitudinal study data are very carefully compiled.

Don't people misreport their education status in surveys, and aren't they reluctant to admit they are dropouts?

Many people—including Greene and Swanson—have argued that this is the case in the Current Population Survey and imparts significant bias to estimates of graduation rates. However, there are no studies documenting the extent of self-reporting bias in high school completion. Note that the CPS surveys are conducted by highly trained individuals and that people self-report not only their education, but also their occupation, wages, incomes, etc. Census bureau surveyors check whether respondents provide different answers at different points in the respondents participation (households are in the CPS for four months then out for eight and then in for another four) and resolve inconsistencies.

The other important thing to note is that even if there is some untruthful reporting of high school completion, this will get cancelled out if we are comparing graduation rates across years. That is, unless people are becoming more or less truthful over years—an unlikely event—self-reporting bias in CPS surveys will not affect the trends in high school completion.

Given that the CPS estimates—with suitable corrections for exclusions of certain populations like the military and those in institutions (prisons and nursing homes)—are close to other reliable estimates from the longitudinal studies like the NELS (which verifies student reports against actual transcripts) and the NLSY, we believe that these estimates are important for tracking the performance of students, both overall and for particular racial groups, over time.

Self-reporting of educational attainment is also the norm in decennial censuses, but as we have shown, the results from the 2000 census are sufficiently close to the longitudinal estimates to suggest that the bias from self-reporting is minimal.