Comparing Graduation Rates Using One State's Data

November 16, 2006

Bobby J. Franklin, Ph.D.
Louisiana Department of Education

Mindy Crain-Dorough, Ph.D.
Louisiana Department of Education

Abstract

The National Governors' Association is urging states to develop a standardized cohort graduation rate definition and to develop data systems with quality data that can in turn be used to produce a reliable indicator. This paper is an examination of one state's graduation rate using the cohort method. The statistics generated are compared to a panel based graduation rate, the Average Freshmen Graduation Rate (AFGR) produced by the National Center for Education Statistics and the Cumulative Promotion Index (CPI) published in Education Week. The cohort, panel and AFGR methods generated similar results across all years while the CPI produced a lower rate for two of four years examined. The CPI appears to be more susceptible to population fluctuations than those that use four years of data.

Comparing Graduation Rates Using One State's Data

For over two decades the National Center for Educational Statistics (NCES) has been collecting and reporting education statistics. Involving states through the National Forum, the NCES has been an advocate of standard data definitions and collection processes. Their most noted endeavor is the development of a national definition for identifying high school dropouts. More recently they have attempted to tackle the issue of an on-time or longitudinal graduation rate. Policy analysts (Green & Winters, 2005; Warren, 2005) have criticized previously published dropout statistics and the data and methods used to generate these indicators. The federal legislation, No Child Left Behind, pushing for national accountability has demanded that schools improve student performance in all areas and track this performance across time. The need to produce longitudinal-based indicators has been the momentum behind the call for student-level or unit record systems that allow states to link student records (National Institute, 2004). The National Governors Association Task Force on State High School Graduation Data (National Governors Association, 2005) has urged states to develop a standardized cohort graduation rate definition and to develop data systems that produce quality data that can in turn be used to produce a reliable indicator.

The Common Core of Data compiled by NCES and the Current Population Survey collected by the U.S. Census Bureau have generally been the main sources of dropout and completion statistics (Green & Forrester, 2003; National Institute, 2004). The national-level statistics heretofore published are based on aggregate information and are at best simulations of cohort rates. The aggregate data make it difficult to adjust the population for those that should be excluded. The exclusion problem is centered mainly on the counting of students who are no longer enrolled. Graduate counts should be fairly accurate and can be collected at one point in

time. Even though this is a simplistic view, graduates are physically present and can be identified and counted. The other leavers, such as dropouts, are more difficult to categorize and therefore difficult to count, especially after they leave.

This report presents the findings of an analysis of 8th and 9th grade panels and cohorts of students using data collected through a state developed unit record system. The rates from three panels and three cohorts of students are examined and compared with each other as well as simulated rates produced by other researchers. Panels of students may be used instead of cohorts simply because panels are more easily tracked across time. Panels are adjusted only for students who leave the group while cohorts are adjusted for leavers as well as for those students that enter school at a later date at the appropriate grade level.

Data System

In 1993, the Louisiana Department of Education implemented its Student Information System (SIS), which is a student-level data collection system. In 1995-96, the Department began reporting dropout statistics extracted from SIS rather than depending on the aggregate reports provided by the 66 Local Education Agencies (LEAs). There was a stark difference in the statistic reported using the two different data collection systems. The dropout rate reported for 1994-95 was 3.8% for grades 9-12 but jumped to 11.8% the following year. This drastic increase was not due to a massive increase in dropouts but rather to a systemic change in the way dropout related data were collected and processed. While the definition had been standardized the application of that definition was not and varied among the districts. By collecting rudimentary exit information from each school system the state was positioned to process all the data and apply the NCES dropout definition in a consistent manner. The SIS currently maintains 29 different exit codes used in the tracking process. Student enrollments

can be tracked across time and district boundaries allowing the state to adjust dropout counts and increase accuracy. In addition, this tracking system allows LEAs access to a Student Inquiry System, which enables them to find students who relocate within the state (Author & Kochan, 2000).

Longitudinal data systems afford opportunities for states to see the impact of policy changes and other phenomena on the public education system. The data used in this study provided the authors the chance to examine the impact of a major policy change while developing a cohort graduation rate. Beginning with the 1999-00 school year, promotion from the eighth to the ninth grade was contingent on passing a criterion referenced test. This promotion criterion had implications for promotion, graduation and dropout rates. The initial impact, as seen in Figure 2, was to create a population shift between the eighth and ninth grades. This shift and its impact on the graduation rate will be discussed later.

Data Quality

Since the SIS is used to determine the Minimum Foundation Funding program, data quality is of utmost importance to the districts. There are numerous edit checks built into the data processing procedures to help ensure that the data are of the correct type. Tolerance procedures are run to identify aggregate data that may be out of normal ranges. An elementary student with 100% attendance is not unusual while an elementary school with 100% attendance is highly improbable. LEAs are provided reports of various kinds and degrees to help them identify data errors and additional time is provided to make corrections. At the state level alone there are easily over 10 man-years of effort annually generated to clean-up the data. There is no way to even estimate the time consumed within the LEAs. Even with the enormous effort put forth there is still error within the data but not enough to question the

integrity and quality of the information that can be generated. This system proved to be a valuable resource to many across the country in the days following the hurricanes of 2005.

Definitions

There has been much concern, discussion and focus on the development of a standardized graduation rate (National Institute, 2004; National Governors Association, 2005). Although graduation rates have been produced for years the quality of these rates has been severely criticized (Swanson & Chaplin, 2003; Swanson, 2004). While a graduation rate should report the activities of high school students over time, there are different opinions about the time periods and grade levels that should be included. A high school graduation rate provides some indication of the ability the school has in getting students to successfully complete the designed education program. Since high school is typically considered to begin with entrance into the 9th grade, most graduation rate formulas use this as a reference point for which to compare the number of graduates three years later at the end of the 12^{th} grade. Task forces set up to examine and propose a method for producing a graduation rate have recommended using a defined cohort model (National Institute, 2004; National Governors Association, 2005). The recommended graduation rate is a comparison of the number of students who enter high school for the first time in the 9^{th} grade in a given year (appointed time) to the number of students in this group who graduate in a prescribed amount of time. The formula would generally be represented as $\frac{\sum CohortGraduates}{\sum Cohort}$. The term cohort is not specific to individuals but is more universal. It must be constrained by institutional and time dimensions. There must be an appointed time for cohort entry, which could be a specific year or even a specific day within a given year and there must be some institutional boundary such as state, district or school. Thus

the first time (FT) 9th grade students in a given year for a given state are not bound by a set membership. As time progresses the composition of the cohort may change as individuals move in and out of the cohort. As 9th grade students in YEAR = Y move to YEAR = Y+1, some will leave the institutional boundary for legitimate reasons and can be removed from the cohort. Others may enter the cohort if they do so at the properly appointed time. For example, in YEAR = Y+1 FT 10th grade students new to the state public education system can be added to the cohort. In YEAR = Y+2 FT 11th grade students new to the state public education system can be added to the cohort and so on. Once the cohort is established then the process turns toward identifying the last record of each cohort member within the designated time of completion.

An alternative to the cohort procedure is that of the panel. A panel is different from a cohort in that it is set by a specific membership. This membership is not adjusted for new entrants as with the cohort. Once the panel is established the focus is on what happens to the individuals within the panel and not with individuals who enter the system after the initial year. In this case the comparison is between the numbers of students in this group who graduate in a prescribed amount of time to the numbers of students who enter high school for the first time in the 9th grade within an appointed time. The formula would generally be represented as $\frac{\sum PanelGraduates}{\sum Panel}$. The term panel is specific to individuals in this case and does not allow for positive membership changes. It also is constrained by institutional and time dimensions. There must be an appointed time, which could be a specific year or even a specific day within a given year and there must be some institutional boundary such as state, district or school just as is with the cohort. Once the panel is established then the process turns toward identifying the

last record of each panel member within the designated time of completion. This is the same procedure used within the cohort model. The primary difference is that only one year is used to establish panel membership whereas with the cohort multiple years or time periods must be used.

The NGA graduation rate can be represented by $\frac{\sum (C_G)}{\sum \left(C_{1\rightarrow 4} - C_{1\rightarrow 4}^{L}\right)} \times 100$

Where

C_G = Cohort graduates and is defined as those students who graduate on time based on their year of entry into the state public school system.

$C_{1\rightarrow 4}$ = Cohort membership and is defined as the count of students who are scheduled to graduate on time based on their year of entry into the state public school system. The symbol 1→4 represents the four years of time that designates their membership in high school

$C_{1\rightarrow 4}^{L}$ = Cohort leavers and is defined as the count of students who are legitimate leavers based on their year of exit from the state public school system.

If special consideration is given to students with disabilities or students who lack English speaking skills then the process as well as the formula becomes more complex.

The panel graduation rate can be represented with the following formula:

$\frac{\sum P_G}{\sum P_1 - \sum P_{1\rightarrow 4}^{L}}$. Where P_1 represents panel members who entered the 9th grade for the first time in the first year of high school, PG denotes those panel members who graduated by the fourth year of high school and $P_{1\rightarrow 4}^{L}$ stands for those panel members who transferred out for

legitimate reasons. Since there is no movement in and out of the panel the only adjustment is for those students who leave for legitimate reasons.

Policy analysts across the country have been calling for better statistics on exit patterns of high school students. This has been primarily directed toward graduation and dropout rates. Then National Center for Education Statistics has also been instrumental in developing standards to produce reliable education statistics. Education Week through its *Diplomas Count* (Edwards, 2006) report has recommended a procedure whereby aggregated data can be used to estimate or simulate a cohort graduation rate. This is the Cumulative Promotion Index (CPI) developed by Chris Swanson. This method uses enrollment data across two years for the four grade levels of high school plus the count of diploma recipients in the second year. This method generally produces a lower graduation rate than what many states had been reporting in the past. NCES has also created and published a similar statistic called the Average Freshman Graduation Rate (AFGR) (Seastrom, Hoffman, Chapman, and Stillwell, 2005). The AFGR also uses aggregate data but differs from the CPI in that the time covered is longer and the formula is more complex. Results obtained from both the CPI and AFGR methods of calculating graduation rates are compared to the panel and cohort statistics.

Method

The data collection system used by the Louisiana Department of Education is a transactional process that collects all student records within a school year. For this reason the data must be processed and placed into a logical format and pattern to enhance the proper analysis. Because of the complex nature of these data, care must be used to ensure the data are not misinterpreted. The raw data contain multiple records for individual students across time as well as institutions. The districts are required to submit every student record stored within the

districts' data systems. For example, John Doe can attend schools A, B, A and C during a single school year. School A would report two records for John Doe while School B and C would report one record. Each record would be unique because of differences in membership time and location. Therefore, for John Doe there are four records that must be placed in the proper sequential order for analysis. In many instances we are only interested in the first and last record so intermediate data may be ignored.

A second important issue that must be addressed is that of student identification. Since the system is designed to collect multiple records on an individual, we must be concerned that these multiple records are not due to data errors. We must also be concerned that multiple records are not overlooked due to data errors. If we also find Jon Doe attending School D, we must determine if this is the same student reported as the John Doe attending Schools A, B, and C. Likewise we need to be sure that Jonathan Doe attending School E is not the same as John Doe. Although the data system looks at multiple indicators to assign unique identifiers the system is not fail-safe and before it can be analyzed the data should be carefully processed to ensure an accurate matching of student records.

Having a student unit record system is a prerequisite to developing a cohort graduation rate. An important second step is developing processing rules for the longitudinal data that must be sorted, filtered and labeled in a manner that allows for easy extraction of the information. Louisiana's data files provide demographic information on the student, time of entry and exit, location data, and education status. Using this information students are assigned a cohort graduation year based on the date they entered the 9^{th} grade for the first time or when they first enter the Louisiana public school system as a 10^{th}, 11^{th}, or 12^{th} grader. This identifies those student records associated with a particular cohort. It does not; however, identify the point of

entry so a cohort join year flag was created to permit cohort members to be sorted into four groups if needed. For example, the cohort join year flag easily permits the user to identify a 9th grade panel for analysis if needed. A student who enters the cohort as a 9th grader is assigned a cohort join year flag of "1" , a 10th grader is assigned a "2", 11th a "3" and 12th a "4". Once a student is assigned a cohort graduation year and join year flag, these data should never change and all records for a particular student should show the same graduation cohort year and join year flag.

To meet the needs of multiple users, aggregation flags were created. These flags designate whether the aggregation occurs at the school, district, or state level. There are two sets—one for cohort aggregations and another for panel aggregations. Policy dictates that new students entering the cohort after a certain point in time would be treated at the school and district levels differently from the state. Two additional sets of flags were developed to handle transfers within the cohort and transfers within the panel.

Once the cohort membership was established the next step was to analyze student exit codes and movement patterns to determine the final status of the student. Five different status codes were developed for final aggregation purposes. Student records were sorted into these five categories using preexisting SIS exit codes plus dropout and graduation flags. Only the last record within the file was used to determine final status. The five codes are as follows:

4x for Graduate (The x is a place holder for codes that represent future variations in graduates/diplomas.),

10 for Dropout,

20 for Attendee,

30 for Completer, and

99 for Legitimate Leaver.

A graduate is any student who receives a regular high school diploma. (Currently in Louisiana a regular high school diploma is the only type diploma issued.) Dropouts are students who can be identified as such using the NCES definition of a dropout based on their last record. An attendee is any student who does not exit at the end of the cohort period for any reason. These students are expected to return to school the following year. Completers are students who received some completion credential other than a regular high school diploma. Legitimate leavers are those students who transfer to another acceptable education program. This includes transfers to other states, nonpublic schools, home schools, and correctional facilities. Legitimate leavers do not include those categories associated with dropout classifications such as unknown reasons but do include deceased students.

Sample

This study reports the comparison of three student panels and three student cohorts for the school years 1999-00, 2000-01, and 2001-02 corresponding to the graduating classes of 2003, 2004, and 2005 respectively. A fourth graduating class, 2002, was included when the cohort and panel rates were compared to the AFGR and the CPI. The data were collected by the Louisiana Department of Education from the Local Education Agencies within the state at the end of the school year and represent all records generated throughout the year. The longitudinal data extends from
1999-00 through 2004-05. Under normal conditions the 2005-06 beginning year collection would have been included in this analysis but because of the population shifts resulting from Hurricane's Katrina and Rita these data are too unstable at this time.

The data were analyzed to describe the movement of students in and out of the state public school system. No adjustments were made to the student records once the initial panel or cohort was established. Every student within the panel or cohort was counted and categorized and the results obtained were compared to each other and the two longitudinal estimates (CPI and AFGR). Data from the panel and cohort samples were examined to provide descriptive statistics about each group then various rates were calculated for the total sample and for subgroups.

The panel samples included those students who were identified as first time ninth graders in the Louisiana public school system. This included schools operated by the Local Education Agencies (LEAs) as well as the special school districts, charter schools and laboratory schools. Students in state correctional facilities are not part of the public school system and are not included. The cohort samples not only included first time ninth graders but also included any student who entered the public school system at the appropriate grade level within the following three years. As with the panels, the cohorts included all public schools and excluded the students in correctional institutions.

Results

Panel results are presented first followed by cohort results then a comparison is made between the two. Graduation rates created using the panel, cohort, CPI and the AFGR are compared at the end of this section.

Cohort to Panel Comparisons. Although the three cohorts average nearly 7,500 more students than the panels, the subgroup ratios are very similar. The average racial composition (See Table 1.) of the three panels is 42% African-American, 55% white and 3% other. The three cohorts are very similar with an average of 41% African-American, 55% white and 4%

other. The panels are slightly more impoverished with a free and reduced lunch population of 45% versus 42% for the cohort. The panels' also averaged one percentage point more students with disabilities and one percentage point more females than the cohorts. There is no average difference with regard to students with limited English speaking skills.

In general, the panels yield a higher percentage of graduates averaging 61.3%, while the cohorts average 57.6%. As Table 2 shows the ratio of completers (3.9%) is the same in both groups and there is a small difference in Attendees with the cohorts (8.8%) being slightly higher than the panels (7.5%.) A much higher percentage of the cohorts (12.3%) are identified as legitimate leavers than in the panels (7.9%.) The cohorts show on average almost two percentage points fewer dropouts than do the panels, 19.3% versus 17.5%.

When rates are compared where the legitimate leavers are excluded, the differences between the panels and cohorts decrease. The average panel graduation rate (66.6 %) is one percentage point higher than with the cohort (65.6 %, See Table 3.). There is only a tenth of a point difference for completers with the cohorts averaging 4.4% versus 4.3% for the panels. The largest difference is for attendees where the cohort average rate of 10.1% is two percentage points higher than the panel average rate of 8.1%. The dropout rate is almost a percentage point higher with the panels (20.9%) than with the cohorts (20.0%.)

The greatest difference between these two groups is in the overall size of the groups. The cohort samples are generally larger than the panels. The cohorts average 7,469 more students than do the panels (See Table 1.). The cohort should be larger because of including students new to the public school system after the initial 9th grade year. By not using the cohort method the state would not be accountable for nearly 7,500 students in this example.

External Comparisons. The panel and cohort graduation rates are compared to the AFGR (Seastrom, Hoffman, Chapman, and Stillwell, 2005) produced by NCES and the CPI (Edwards, 2006). To gain a better picture of comparative results a fourth cohort year is included. Figure 1 shows the variation in the rates for each of the four methods for four different time spans. For graduating classes of 2002 and 2003 there is very little difference among the four methods. Years 2004 and 2005 show a marked difference between the CPI and the other three methods. Using the cohort gradation rate as a reference point, one can see in Table 4 that the panel and AFGR methods average approximately a 1.1% higher rate across the four years. The CPI method produced a 6.7% average difference ranging from 1.5% higher in 2002 and 19.3% lower in 2005.

Conclusion

The cohort, panel, and AFGR methods of producing a high school graduation rate appear to produce stable statistics across time. The cohort statistics are on par with those produced using the panel method and the AFGR method produced by NCES. The CPI deviated significantly for two of the four years compared.

The AFGR, panel and cohort procedures are using data across four years, which appears to provide more stability. The CPI seems to be more susceptible to population fluctuations because it uses only two years of data. The increase in the 2004 graduation rate obtained through the panel, cohort and AFGR methods is probably due to the significant shift in population between the eighth and ninth grade in the 2000-01 school year as described earlier (See Figure 2.). This shift is due to the implementation of an eighth grade high stakes test for entrance into high school. This population change apparently had the opposite effect on the CPI. As the population begins to move back toward the center, the rates drop in 2005. The

drop for the CPI is much more dramatic because of a significant decline in enrollment in the 2005-06 school year due to hurricanes Katrina and Rita.

Implications

The development of student cohorts affords numerous opportunities to examine major policy changes within a state education system. What impact do these changes have on the completion rates of high schools? In Louisiana, researchers and policymakers have begun to examine the impact of high stakes tests on high school completion rates. The cohort also permits examinations of differences between those students who enter high school in the ninth grade versus those that enter later. Is there a difference in the achievement and behavior of those students who transfer into the cohort and those that are initial cohort members? Given that the cohort rate is consistently lower than the panel rate these differences need to be explored. The impact on subgroups of high schools, high stakes testing, and subsequent policy changes can and should be closely scrutinized with cohort data.

Table 1

Demographic Composition of Panels and Cohorts Expressed as Percentages

	Panel			Cohort		
	Graduating Year			Graduating Year		
Demographic Indicators	2003	2004	2005	2003	2004	2005
Total Membership	55610	45922	49757	62514	53400	57783
African-American	45	39	43	43	38	42
White	52	58	54	53	57	54
Other Race	3	3	3	4	5	4
Free/Reduced Priced Lunch	45	43	47	42	40	44
Paid Lunch	55	57	53	58	60	56
Female	50	50	49	50	49	49
Male	50	50	51	50	51	51
Limited English Proficient	1	1	1	1	1	1
Students with Disabilities	10	11	11	9	10	10

Table 2

Composition of Panels and Cohorts by Exit Categories Expressed as Percentages

	Panel			Cohort		
	Graduating Year			Graduating Year		
Exit Categories	2003	2004	2005	2003	2004	2005
Graduate	58.1	64.8	61.1	55.2	60.5	57.0
Completer	3.3	3.9	4.6	3.2	3.9	4.6
Attendee	9.1	6.5	6.9	10.4	7.6	8.5
Legitimate Leaver	7.9	8.0	7.9	12.0	12.4	12.4
Dropout	21.6	16.8	19.4	19.2	15.7	17.6

Table 3

Exit Rates of Panels and Cohorts Excluding Legitimate Leavers

	Panel			Cohort		
	Graduating Year			Graduating Year		
Exit Categories	2003	2004	2005	2003	2004	2005
Graduate	63.1	70.4	66.4	62.8	69.0	65.0
Completer	3.6	4.3	5.0	3.6	4.4	5.2
Attendee	9.9	7.0	7.5	11.8	8.7	9.7
Dropout	23.4	18.3	21.1	21.9	17.9	20.1

Table 4

A comparison of cohort graduation rates to panel, AFGR and CPI.

	Year			
	2002	2003	2004	2005
Panel Graduation Rate	-1.2	-.4	-1.4	-1.4
AFGR	-1.4	-1.3	-0.4	1.1
CPI	-1.5	1.1	7.9	19.3

Note: Negative values are higher and positive values are lower than cohort rates.

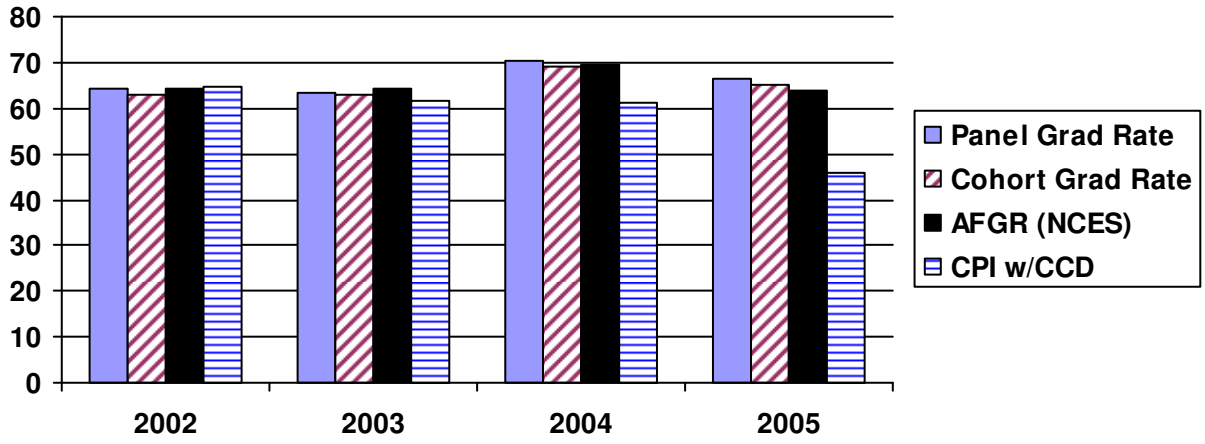

Figure 1. A comparison of four different methods for calculating a four year graduation rate.

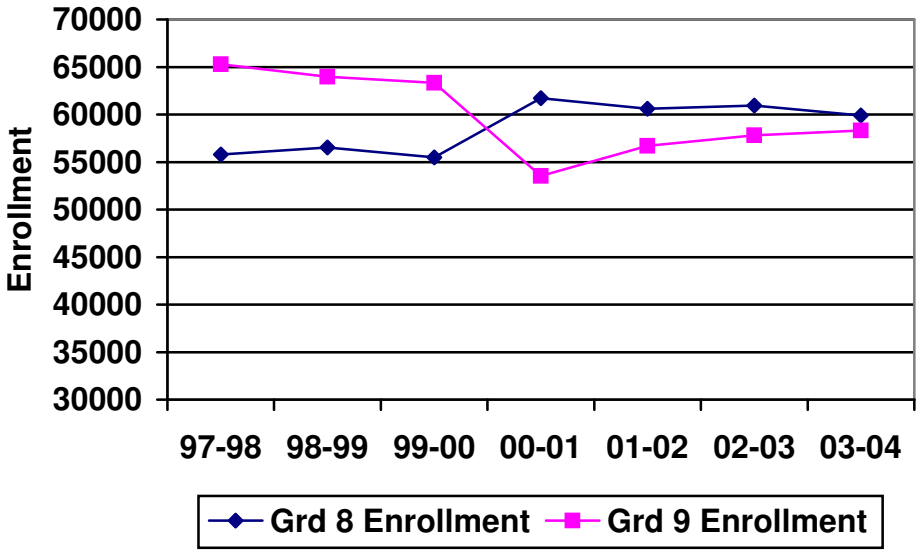

Figure 2. A comparison of enrollment data for grades eight and nine

References

Edwards, Virginia B. (Ed.). (2006, June 22). Diplomas count: An essential guide to graduation policy and rates. *Education Week, 25*(41S).

Author. and Kochan, Susan. (2000-2001). A statewide dropout tracking system. *Journal of Educational Technology Systems 29*(4), 313-339.

Greene, Jay P. & Forster, Greg. (2003). *Public high school graduation and college readiness rates in the United States.* New York, NY. Manhattan Institute for Policy Research

Green, Jay P. & Winters, Marcus A. (February, 2005). *Public high school graduation and college-readiness rates: 1991-2002.* New York, NY. Manhattan Institute for Policy Research

National Governors Association. (2005). *Graduation counts: A report of the NGA Task Force on State High School Graduation Data.* Washington, DC. National Governors Association.

National Institute of Statistical Sciences and the Education Statistics Services Institute. (2004). *National Institute of Statistical Sciences/Education Statistics Services Institute Task Force on Graduation, Completion, and Dropout Indicators: Final Report* (NCES Publication No. 2005-105). Washington, DC: U.S. Government Printing Office.

Seastrom, M., Hoffman, L., Chapman, C., and Stillwell, R. (2005). *The Averaged Freshman Graduation Rate for Public High Schools From the Common Core of Data: School Years 2001-02 and 2002-03* (NCES 2006-601). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Swanson, Christopher B. (2004, July 28). *The new math on graduation rates*. The Urban Institute. Retrieved February 16, 2006, from http://www.urban.org/url.cfm?ID=1000675

Swanson, Christopher B. & Chaplin, Duncan. (2003). *Counting high school graduates when graduates count: Measuring graduation rates under the high stakes of NCLB.* Washington, D.C.: The Urban Institute. (ERIC Document Reproduction Service No. ED474605)

Warren, John Robert. (2005). State-level high school completion rates: Concepts, Measures, and trends. *Education Policy Analysis Archives, 13*(51), 1-35.