DOCUMENT RESUME

ED 439 194 UD 033 411

AUTHOR	Winglee, Marianne; Marker, David; Henderson, Allison; Young, Beth Aronstamm; Hoffman, Lee
TITLE	A Recommended Approach to Providing High School Dropout and Completion Rates at the State Level. Technical Report.
INSTITUTION	National Center for Education Statistics (ED), Washington, DC.; Westat, Inc., Rockville, MD.
REPORT NO	NCES-2000-305
PUB DATE	2000-02-00
NOTE	59p.
AVAILABLE FROM	ED Pubs, P.O. Box 1398, Jessup, MD 20794-1398. Tel: 877-433-7827 (Toll Free).
PUB TYPE	Numerical/Quantitative Data (110) -- Reports - Evaluative (142)
EDRS PRICE	MF01/PC03 Plus Postage.
DESCRIPTORS	*Comparative Analysis; *Dropout Rate; Dropouts; *High School Students; High Schools; *Research Methodology; *School Statistics; Tables (Data)
IDENTIFIERS	Common Core of Data Program; National Center for Education Statistics

ABSTRACT

The National Center for Education Statistics (NCES) has explored the feasibility of adjusting nonstandard dropout reports to make them comparable with those from states using the standard Common Core of Data (CCD) definition. Between 1997 and 1999 staff from NCES and state education agencies worked with analysts from Westat to develop a methodology for adjusting nonconforming dropout data and to test a proposed high school completion rate. The analyses presented in this report find that the major types of nonstandard dropout reporting practices have statistically significant, but different, effects on the size of state dropout rates. The most common variant used a reporting calendar that takes a "snapshot" count of dropouts at the conclusion of the school year rather than the beginning. This typically leads to a small net increase in the number of dropouts reported, compared with the CCD reporting guidelines. This report recommends that data from states using an alternative calendar be published without adjustment, with data from states that conform to the CCD reporting calendar and a footnote to identify alternative reporting calendars. This would add 12 to the number of states with dropout data reported by the NCES. The effects of variations regarding the reporting of summer dropouts and the status of those moving to adult education are more variable, and it is recommended that NCES continue to withhold publication of data from states with these variations (10 states in 1995). Appendixes contain dropout and complete definitions and a map of Census regions. (Contains 12 tables.) (SLD)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 439 194

NATIONAL CENTER FOR EDUCATION STATISTICS

Technical Report

February 2000

A Recommended Approach to Providing High School Dropout and Completion Rates at the State Level

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

U.S. Department of Education
Office of Educational Research and Improvement

NCES 2000–305

BEST COPY AVAILABLE

UD033411

Customer Survey of OERI Publication Users

To help us improve future editions of this publication and give you better customer service, we would appreciate your comments on this survey form. Please check the appropriate box(es) below for each question. Responses will be kept completely confidential. You may return the survey by mail or FAX. It can be folded and taped closed to allow mailing to the address listed on the reverse side of this form, or it can be returned by FAX to 202–219–1321. **Many thanks for your customer feedback—it is very important to us!**

1a. Name of publication

1b. Publication number

1c. Author name

2. How did you receive a copy of this publication?

- ❑ Bought it
- ❑ Borrowed it
- ❑ Mailing list membership
- ❑ Telephone request
- ❑ Internet request
- ❑ Other (please describe)

3. Was this publication easy to get?

❑ Very ❑ Somewhat ❑ Not at all

4. How did you find out about this and other OERI publications? (Check all that apply.)

- ❑ Conferences
- ❑ Journal articles
- ❑ Teacher/educator
- ❑ Professional associations
- ❑ Internet (WWW)
- ❑ Publication announcement
- ❑ Received in mail
- ❑ OERI staff contact

5. For what purposes did you use this OERI publication? (Check all that apply.)

- ❑ Planning
- ❑ Policy or legislation
- ❑ Administrative decisions
- ❑ Teaching, class material
- ❑ Research/analysis
- ❑ General information
- ❑ Writing news articles, TV or radio material
- ❑ Marketing, sales, or promotion
- ❑ Other (please describe)

6. Did the publication help you accomplish whatever you needed it for?

❑ Yes ❑ No ❑ Partially

7. What is your occupation?

❑ Parent ❑ Teacher ❑ Administrator
❑ Librarian ❑ Researcher ❑ Statistician
❑ Journalist/writer ❑ Policy Analyst ❑ Student
❑ Program Planner
❑ Other (please specify)

8. How could this OERI publication (or other OERI publications) better meet your needs? (Check all that apply.)

- ❑ More important topics in education
- ❑ More timely release of data
- ❑ More text introductions to each section
- ❑ More research statistics
- ❑ Shorter reports (less than 10 pages)
- ❑ Other (please describe)

9. Overall, how satisfied are you with this product?	**Very Satisfied** ❑	**Satisfied** ❑	**Dissatisfied** ❑
a. Comprehensiveness of information	❑	❑	❑
b. Clarity of writing (readability, interpretability)	❑	❑	❑
c. Clarity of presentation (e.g., tables, charts)	❑	❑	❑
d. Timeliness of information	❑	❑	❑
e. Accuracy of information	❑	❑	❑
f. Clarity of technical notes	❑	❑	❑
g. Usefulness of resources and bibliography	❑	❑	❑
h. Organization	❑	❑	❑
i. Length	❑	❑	❑
j. Format	❑	❑	❑

PAPERWORK BURDEN STATEMENT
Office of Educational Research and Improvement (OERI)
Publication Customer Survey

According to the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number. The valid OMB control number for this information collection is 1880–0529. The time required to complete this information collection is estimated to average 10 minutes per response, [illegible] time to review instructions, search existing data resources, gather the data needed, and complete and review the information collection. If you have any comments [illegible] he accuracy of the time estimate(s), suggestions for improving this form, or concerns regarding the status of your individual submission of this form, write directly to: [illegible] oom 204, Media and Information Services, OERI, U.S. Department of Education, 555 New Jersey Avenue NW, Washington, DC 20208–5570.

BEST COPY AVAILABLE

OERI Publication Customer Survey
Media and Information Services
555 New Jersey Avenue NW—Rm. 202
Washington DC 20208–5570

OFFICIAL BUSINESS
PENALTY FOR PRIVATE USE, $300

NO POSTAGE
NECESSARY
IF MAILED
IN THE
UNITED STATES

BUSINESS REPLY MAIL
FIRST-CLASS MAIL PERMIT NO. 012935 WASHINGTON DC

POSTAGE WILL BE PAID BY US DEPARTMENT OF EDUCATION

V. Allen
Room 200
Media and Information Services
Office of Educational Research and Improvement
US Department of Education
555 New Jersey Avenue NW
Washington DC 20208–5570

Fold on line—TAPE CLOSED—DO NOT STAPLE

10. Do you have any suggestions regarding the content or format of future editions of this publication or other comments?

MIS 1999–6532

NATIONAL CENTER FOR EDUCATION STATISTICS

Technical Report February 2000

A Recommended Approach to Providing High School Dropout and Completion Rates at the State Level

Marianne Winglee
David Marker
Allison Henderson
Westat

Beth Aronstamm Young
Lee Hoffman
National Center for Education Statistics

U.S. Department of Education
Office of Educational Research and Improvement NCES 2000–305

U.S. Department of Education
Richard W. Riley
Secretary

Office of Educational Research and Improvement
C. Kent McGuire
Assistant Secretary

National Center for Education Statistics
Gary W. Phillips
Acting Commissioner

The National Center for Education Statistics (NCES) is the primary federal entity for collecting, analyzing, and reporting data related to education in the United States and other nations. It fulfills a congressional mandate to collect, collate, analyze, and report full and complete statistics on the condition of education in the United States; conduct and publish reports and specialized analyses of the meaning and significance of such statistics; assist state and local education agencies in improving their statistical systems; and review and report on education activities in foreign countries.

NCES activities are designed to address high priority education data needs; provide consistent, reliable, complete, and accurate indicators of education status and trends; and report timely, useful, and high quality data to the U.S. Department of Education, the Congress, the states, other education policymakers, practitioners, data users, and the general public.

We strive to make our products available in a variety of formats and in language that is appropriate to a variety of audiences. You, as our customer, are the best judge of our success in communicating information effectively. If you have any comments or suggestions about this or any other NCES product or report, we would like to hear from you. Please direct your comments to:

National Center for Education Statistics
Office of Educational Research and Improvement
U.S. Department of Education
555 New Jersey Avenue NW
Washington, DC 20208–5574

February 2000

The NCES World Wide Web Home Page is: ***http://nces.ed.gov***
The NCES World Wide Web Electronic Catalog is: ***http://nces.ed.gov/pubsearch/index.asp***

Suggested Citation
U.S. Department of Education. National Center for Education Statistics. *A Recommended Approach to Providing High School Dropout and Completion Rates at the State Level,* NCES 2000–305, by Marianne Winglee, David Marker and Allison Henderson. Project Officer, Beth Aronstamm Young. Washington, DC: 2000.

For ordering information on this report, write:

U.S. Department of Education
ED Pubs
P.O. Box 1398
Jessup, MD 20794–1398

or call toll free 1–877–4ED–Pubs.

Content Contact:
Beth Aronstamm Young
(202) 219–1562

ERIC

Executive Summary

The National Center for Education Statistics (NCES) has been collecting counts of public school dropouts through its Common Core of Data (CCD) survey since the 1991-1992 school year. However, not all states report dropout data in strict agreement with the CCD definition, with the result that data from these nonconforming states have been withheld from publication. This situation has led NCES to explore the feasibility of adjusting nonstandard dropout reports to make them comparable with those from states using the standard CCD definition.

The desire for comparable dropout statistics has been accompanied by considerable interest in developing a standard high school completion statistic based upon data available from the CCD. Between 1997 and 1999, staff from NCES and state education agencies worked with analysts from Westat to develop a methodology for adjusting nonconforming dropout data and to test a proposed high school completion rate.

The analyses presented in this report found that the major types of nonstandard dropout reporting practices have statistically significant, but different, effects on the size of state dropout rates. The most common variant practice uses a reporting calendar that effectively takes a "snapshot" count of dropouts at the conclusion of the school year rather than at the beginning of the next year. This typically leads to a small net increase in the number of dropouts reported, when compared with the CCD reporting guidelines. The effects of how summer dropouts (those who complete one school year but fail to enroll for the next) are reported, and whether students moving to adult education GED classes are considered dropouts (as required by the CCD) were stronger.

As a result, the report recommends that data from states using an alternative reporting calendar be published, without adjustment, with data from the states that conform to the CCD reporting calendar, and a footnote be used to identify states using an alternative reporting calendar. This would add 12 states to the number whose CCD dropout data are reported by NCES. Because the effects of the other two variations are stronger and more variable than the first, it is recommended that NCES continue to withhold publication of data from the states that follow these variations. There were 10 such states in 1995.

It appears from the analysis of dropout and completion data that the CCD can support a useful high school completion rate. This rate is the proportion of students who leave high school (grades 9 through 12) with a diploma or other credential to the total number of students who leave as completers or dropouts. High school equivalency recipients are excluded from the completer group because these data are not reported at the school district level, as are dropout and other high school graduation and completion counts.

The report recommends that a method using multiple years of dropout data be used in preference to a synthetic, or reconstructive, rate based on a single year of information.

ACKNOWLEDGMENTS

Thanks are due to many people who contributed to this work. First are the Common Core of Data (CCD) Coordinators and other State Education Agency staff who motivated and inspired us to do this work. This project would not have been successful without their contributions and practical thinking. Also, thanks go to Jonaki Bose and Bob Burton of NCES who contributed much to the creation of this work. And as always, there is a large debt of gratitude to John Sietsema who is all things CCD.

The authors would also like to thank Renee Roberts, Andrew Heller, and Ann Webber from Westat for their support in data preparation and programming, and Charlotte Lass for editing the report.

Many people spent quality time reviewing this report and the authors gratefully acknowledge the comments and suggestions of the reviewers. Reviewers from outside the U.S. Department of Education were: Tom Ogle, Missouri Department of Education; Bob Jones, Oregon Department of Education; David Smith, Nevada Department of Education; Melinda Fowler and Cherry Kugle, Texas Education Agency; Gary Tatlock, Indiana Department of Education; Dennis Cheek, Rhode Island Department of Education; Doris White, West Virginia Department of Education; and Linda Frazer, Department of Defense Dependents' Schools. From NCES the reviewers included Chris Chapman, Andrew Kolstad, Marilyn McMillen, and Paula Knepper. Department of Education reviewers included Mary Schifferli, Office of Civil Rights; Richard DiCola, Office of Vocational and Adult Education; and Judy Holt, Office of Special Education and Rehabilitative Services.

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

List of Appendixes

List of Tables

TABLE OF CONTENTS (continued)

List of Tables (continued)

List of Figures

1. INTRODUCTION

1.1 Study Goals

This report presents a possible approach to providing high school dropout and completion rates at the state level using data from the Common Core of Data (CCD) collected by the National Center for Education Statistics (NCES). The CCD is a national database of public elementary and secondary schools and school districts. State Education Agencies (SEAs) complete CCD surveys each year, and a dropout component has been included in the CCD beginning with the 1991 dropout counts[1] reported on the 1992 CCD Local Education Agency Universe Survey. A state's reporting of dropout statistics to the CCD is voluntary. While most states report dropout data to the CCD, some states do not conform to the CCD dropout definition. The variations in reporting practices affect the quality and comparability of the dropout data in the CCD. As a result, NCES publishes only the dropout rates from states that follow the CCD definition, and suppresses the dropout data from states that use nonconforming practices.

The two goals of this study are:

- Evaluate the quality of dropout data in the CCD to determine whether it is feasible to compensate for inconsistencies in states' reporting practices, thereby providing a comparable CCD dropout rate for more states then is now possible, and
- Explore two methods of calculating a high school completion rate using the CCD, and provide an annual completion rate by state.

This study used the CCD data to provide high school dropout and completion rates for 1993, 1994, and 1995. For states that did not conform to the dropout definition, an adjustment method was used to compensate for the effects of nonconformity and adjusted rates were estimated for these states to facilitate cross-state comparisons. Completion rates for each year were computed using a synthetic method (see Section 2.2). As dropout data collection continues in the CCD, the calculation of a longitudinal completion rate that emulates a true cohort rate is becoming possible.

[1] This report uses 1991 to refer to 1991-92 school year. Dropout data for 1991 were reported in the 1992 CCD file, after the academic year was completed.

1.2 Organization of the Report

This report consists of six chapters. Following this introduction, Chapter 2 provides a brief review of the background of the study, and describes the methods for computing high school dropout and completion rates using the CCD. Chapter 3 reviews the quality of the dropout data in the CCD, the extent and types of nonconformity, and the practices used by states. Chapter 4 presents an adjustment method that uses a model-based approach to improve data quality. Chapter 5 presents a method to calculate adjusted high school dropout and completion rates with compensation for nonconforming dropout data, and uses a graphical display to present dropout and completion rates for 1993, 1994, and 1995. Chapter 6 offers a summary and a discussion of the merits and limitations of the current research.

2. HIGH SCHOOL DROPOUT AND COMPLETION RATES FROM THE COMMON CORE OF DATA (CCD)

This chapter provides a brief overview of the dropout data collected through the CCD and summarizes the definition and calculation of high school dropout and completion rates using the CCD database. Section 2.1 reviews the development of the dropout computation for the CCD, and the method used to compute a high school dropout rate. Section 2.2 reviews several methods for calculating high school completion rates, and proposes the use of CCD data to provide an annual completion rate. Two methods to compute the proposed rate are discussed.

2.1 Defining and Calculating High School Dropout Rates

For the past seven years, the CCD has included a dropout statistic in the agency level data. Through the National Cooperative Education Statistics System, NCES worked with states and school districts to develop a dropout data collection and encourage the growth of the CCD as a national database for public school dropout information.

Standardizing dropout data collection through the CCD required a common dropout definition for uniform reporting by all states. The development of the CCD dropout definition was a collaborative effort. NCES worked with state representatives, CCD coordinators, educational researchers, and the academic community to agree upon a common dropout definition. The statistical analysis report *National Dropout Statistics Field Test Evaluation* (NCES, 1992) describes the development and field-testing of an initial definition. The methodology report, *State Dropout Data Collection Practices: 1991-92 School Year* (Hoffman, 1995) follows with the outcome and adjustments that were made after the first year of implementation.

NCES's annual reports on *Dropout Rates in the United States* (e.g., McMillen et al., 1997; McMillen and Kaufman, 1998; and Kaufman et al., 1999) provide statistics on high school dropout rates by academic year. These reports summarize the national dropout situation by analyzing data from several sources, including the Current Population Survey (CPS), the High School and Beyond Study (HS&B), the National Education Longitudinal Study of 1988 (NELS:88), and the CCD. However, while the CPS data provide national and regional information about dropouts, the sample is not large enough for reliable analysis at the state or school district levels. The CCD, as an annual universe collection, can provide a

count of all public school dropouts. (Appendix A provides the CCD standard definition for classifying students as dropouts or not as dropouts.)

NCES reports only dropout data that meet the quality and comparability standards necessary to support valid cross-state comparisons. The CCD annual event dropout rate is defined as the percentage of students classified as high school dropouts for a given school year and grade among all high school student members enrolled in that grade on October 1 of the same school year (including, of course, those who will drop out). Note that a composite grade 9-12 rate is based on the sum of dropouts and enrollments across these four grades.

The annual event rates describe the proportion of students who leave school each year without completing a high school program. For example, to compute the 1995 dropout rate, the denominator is the October 1, 1995 membership reported in the 1995 CCD Public Elementary/Secondary School Universe Survey; dropout data are the counts reported for 1995 in the 1996 Local Education Agency Universe Survey. The only adjustment made to this rate is prorating any ungraded student membership. In reporting dropouts, districts must assign the dropout a grade (i.e. there are no "ungraded" dropouts reported on the CCD). However, a district can report ungraded[2] student membership. This ungraded student membership is prorated into the denominator of the calculated dropout rate.

The state dropout rate is computed from the district level data in the Local Education Agency Universe Survey by using a composite estimate as follows:

$$r_s^{rep} = \frac{\sum_{i=1}^{n_s} y_i}{\sum_{i=1}^{n_s} M_i} \qquad (1)$$

where

r_s^{rep}	is the reported dropout rate for grades 9 to 12 in state s;
y_i	is the reported dropout count for grades 9 to 12 in district i;
M_i	is the reported student membership for grades 9 to 12 in district i; and
$i = 1,2,\ldots,n_s$	where n_s, is the number of districts in state s.

[2] Ungraded students are those students who are in classes or programs to which students are assigned without standard grade designations.

2.2 Defining and Calculating High School Completion Rates

High school completion rates go hand-in-hand with dropout rates. The CPS does collect high school completion rate information but has three limitations when it is used to measure outcomes in state public school systems. One is that the CPS does not have the sample size to support annual estimation of state completion rates; NCES uses a "rolling average" based on multiple years. Another limitation is that some CPS respondents may have attended school in a different state than their place of residence at the time of the CPS interview, thereby affecting the accuracy of state estimates. A third limitation is that the CPS does not distinguish between public and nonpublic school completers and dropouts.

The CCD is an alternative data source that can provide annual completion rates for public schools by state. Bose and Hoffman (1997) presented eight potential rates through which the CCD data can be used to give an estimate of high school completion. They compared each of these rates against a cohort rate (one obtained by following a group of students across four years of high school) and, with state support, suggested a completion rate based on counts of high school completers and dropouts. This is deliberately termed a "completion" rate. It is the proportion of students who leave high school with some completion credential compared to all students who leave school (completers and dropouts). Section 2.2.1 presents the longitudinal method of calculating this rate, and Section 2.2.2 shows the synthetic method. This report shows the results of using the two rates as well as discussing the pros and cons of each method.

2.2.1 Longitudinal Completion Rate

The longitudinal method of calculating a CCD completion rate, while not a true cohort approach in which students are followed through four years of high school, does emulate one. It accomplishes this by using data from grade 9 in year 1, grade 10 in year 2, and so on. For simplicity, the following expression shows the method of calculation using composite data at the state level. The longitudinal completion rate, c_{st}^{long} for state s at year t is :

$$c_{st}^{long} = \frac{g_{st}}{g_{st} + d_{st}^{12} + d_{s(t-1)}^{11} + d_{s(t-2)}^{10} + d_{s(t-3)}^{9}} \quad (2)$$

where for state s

g_{st} is the number of graduates at year t;

d_{st}^{12} is the number of grade 12 dropouts at year t;

$d_{s(t-1)}^{11}$ is the number of grade 11 dropouts at year t-1;

$d_{s(t-2)}^{10}$ is the number of grade 10 dropouts at year t-2; and

$d_{s(t-3)}^{9}$ is the number of grade 9 dropouts at year t-3.

The number of graduates is taken from the district level data in the CCD Local Education Agency Universe Survey. The data include students who graduate with a regular or other diploma and other high school completers. (High school equivalency recipients are reported on the state file and thus are excluded from this district-based rate. For a definition of the high school completer categories, see Appendix A.) For example, in order to compute the 1995 longitudinal completion rate, the required data elements are: the counts of 1995 graduates, the 1995 grade 12 dropouts, the 1994 grade 11 dropouts, the 1993 grade 10 dropouts, and the 1992 grade 9 dropouts.

Given the data available in the CCD, this method provides an estimate of completions that is close to the true cohort rate. Using four years of data allows the state or district rate to be less affected by one-year changes than a synthetic rate (based on a single year's data) would allow. One constraint of this method is that it requires dropout data for four consecutive years. Another constraint is that new district or school programs must be in place for several years to have a visible impact on this rate. Also, those districts that have a sizable net loss in students over time due to migration out of the district may be at a slight disadvantage since their dropouts from four years ago were drawn from a larger student body than that providing the number of completers in the current year.

2.2.2 Synthetic (or Reconstructive) Completion Rate

An alternative method to compute completion rates is a synthetic, or reconstructive, method. Instead of following something like a cohort of students over four years, this method uses the students of the current year as the synthetic cohort. For example, to compute the 1995 synthetic completion rate, the 1995 grade 9, 10, and 11 dropouts are used to reconstruct the 1994 grade 11 dropouts, the 1993 grade 10

dropouts, and the 1992 grade 9 dropouts. Using this reconstructed cohort, the calculation of a synthetic completion rate requires only one year of dropout data. The expression in equation (2) simplifies as follows:

$$c_{st}^{syn} = \frac{g_{st}}{g_{st} + d_{st}^{12} + d_{st}^{11} + d_{st}^{10} + d_{st}^{9}} \qquad (3)$$

The synthetic completion rate is easier to compute than the longitudinal rate because it is not necessary to cumulate dropout data over four years. When the dropout situation and the student population are fairly consistent across the four years, the synthetic rate will provide a reasonable approximation to the longitudinal rate. However, when there are policy interventions and large demographic changes during the four years, the synthetic rate may not be appropriate. Also, if multiple years of data are available, a synthetic method may add more confusion to the already complex nature of reporting high school completers and dropouts.

3. QUALITY OF DROPOUT DATA

This chapter reviews the quality of the dropout data in the CCD. The issue is nonconformance with the standard CCD definition and reporting practices. Section 3.1 discusses the extent of variation. Section 3.2 reviews the types of nonconformance in states' reporting of dropout data. Section 3.3 shows the practices used by states for 1993, 1994, and 1995. Since the calculations of both dropout and completion rates depend on dropout counts, improvements in the quality of dropout data in the CCD affect both statistics.

3.1 Extent of Nonconformance

Table 3-1 shows the total number of states that reported dropout counts and the number that reported using the CCD definition between 1991 and 1995. Some 45 or 46 states reported dropout data each year. For 1991, one-third of the states that reported dropout data were in conformance with the CCD definition. The number increased to about half by 1995. States' continuing efforts to report dropouts show their willingness to provide the information. The concern is that some states may never change to the CCD definition because data collection systems are difficult (and expensive) to change, as are local administrative practices and state policies. States that have not conformed express no immediate plans to change their reporting practices. States that have not reported at all do not currently collect dropout data in a way that can be given to NCES.

Table 3-1. Number of states reporting dropout data to the CCD, 1991 to 1995

Dropout year	Number of states*	
	Reported	Reported with CCD definition
1991	45	15
1992	46	20
1993	46	21
1994	46	25
1995	45	23

* Including the District of Columbia but not outlying areas.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Common Core of Data, Local Education Agency Universe Survey, unpublished internal working files, 1992-1996.

3.2 Types of Nonconforming Practices

Staff from NCES or the Bureau of the Census, which collects and edits the CCD data, contact the state CCD coordinator each year to verify the reporting practices for dropout data and to record the ways in which state practices differ from the CCD definition. Table 3-2 summarizes the differences between the CCD definition and the nonconforming practices.

Table 3-2. Types of nonconforming practices

Type of practice	CCD Definition	Nonconforming practice
Alternative reporting calendar (A)	October-cycle Reporting year begins on the first day of school	June-cycle Reporting year ends on the last day of school
Summer dropout (S)	Reported as dropout in grade and year for which the student failed to return	Reported as dropout in grade and year completed
Adult GED (G)	Reported student in this program as dropout	Did not report student in this program as dropout

SOURCE: U.S. Department of Education, National Center for Education Statistics, Common Core of Data, Local Education Agency Universe Survey, unpublished internal working files, 1992-1996.

Alternative reporting calendar. The most common nonconforming practice is associated with the use of an alternative reporting calendar. The CCD dropout definition is based on an October-September reporting cycle with the summer months considered part of the school year preceding them. This approach is based on the cross-sectional perspective of taking a snapshot on October 1 of each year. It counts as dropouts those students who were enrolled in school at some time during the previous school year but are not enrolled by October 1 of the current school year. Students who return to school after October 1 (the snapshot day) are counted as dropouts for the previous school year. Students who left school during the previous school year but are re-enrolled on or before October 1 are not counted as dropouts. In contrast, some states use a July-June calendar cycle, in which the snapshot is taken on the last day of school, defined for convenience as June 30.

Summer dropout. The second difference occurs with students who drop out between school years. The CCD definition classifies students who completed the previous school year but fail to enroll by October 1 as the dropouts from the grade and school year for which they fail to return. Some states, in contrast, count these students as dropouts from the grade and school year that they just completed.

Table 3-3 shows the assignment of summer dropouts based on the CCD definition and the nonconforming practices. For example, students who completed grade 8 in 1994 but failed to enroll in grade 9 in October 1, 1995, are counted as grade 9 dropouts in 1995 under the CCD definition. In contrast, the nonconforming practice counts these students as grade 8 dropouts in 1994.

Table 3-3. The assignment of summer dropout students by grade and year: CCD definition and nonconforming practices

Summer dropout	Grade completed	Assignment of dropouts by grade and year					
		CCD Definition			Nonconforming practice		
		1994	1995	1996	1994	1995	1996
1995	8		9		8		
	9		10		9		
	10		11		10		
	11		12		11		
	12		12		12		
1996	8			9		8	
	9			10		9	
	10			11		10	
	11			12		11	
	12			12		12	

SOURCE: U.S. Department of Education, National Center for Education Statistics, Common Core of Data, Local Education Agency Universe Survey, unpublished internal working files, 1992-1996.

Adult GED. The third difference concerns students in adult education programs preparing for the General Educational Development (GED) test. The CCD definition reports all students who transfer to adult GED programs as dropouts unless the school district tracks these students and reports them as dropouts should they leave the adult GED program. Students enrolled in secondary school programs preparing for GED are not dropouts. The nonconforming practice does not report any adult education GED students as dropouts; in other words, a student enrolled in any GED program is treated as a transfer.

3.3 Reporting Practices by State

Table 3-4 lists the type of nonconforming practices by state for 1993, 1994, and 1995. Michigan, New Hampshire, and Washington did not report any dropout data for these years and are not included in this analysis. Texas reported in these years, but since it had a unique variation from the CCD procedures (dropouts who re-enrolled were never counted as dropouts again), the state's data were

excluded from this research in order to make the adjustment model as robust as possible. Beginning with 1996, Texas is reporting in conformance with the CCD definition. Louisiana changed its dropout data collection system for 1995; for reasons of comparability, the 1993 and 1994 data for Louisiana were excluded from this research. Alaska, the District of Columbia, and Montana only reported dropout data for some of the years.

Most states that differ from the CCD definition do so because of one nonconforming practice. For example, in 1995, 16 states followed an alternative reporting calendar[3]; five states (Hawaii, Indiana, Kentucky, North Carolina, and Virginia) differed from the CCD standard in the way in which they reported summer dropouts; and six states (California, Florida, Indiana, New Mexico, Oregon, and South Carolina) did not conform in their reporting of adult GED. Florida, Oregon, and South Carolina were different because of both alternative reporting calendar and adult GED discrepancies. Virginia was nonconforming because of alternative reporting calendar and summer dropout differences; and Indiana was nonconforming because of summer dropout and adult GED differences.

[3] These states were Alabama, Alaska, Arizona, Colorado, Florida, Idaho, Illinois, Maryland, New Jersey, Oregon, South Carolina, South Dakota, Tennessee, Vermont, Virginia, and Wisconsin.

Table 3-4. Types of nonconforming practices by state, 1993 to 1995 (A=Alternative reporting calendar, S=Summer dropout, G=Adult GED)

State	Types of nonconforming practice		
	1993	1994	1995
Alabama	A	A	A
Alaska	*	*	A
Arizona	A	A	A
Arkansas			
California	G	G	G
Colorado	A	A	A
Connecticut			
Delaware			
District of Columbia			*
Florida	A,G	A,G	A,G
Georgia			
Hawaii	S	S	S
Idaho	A	A	A
Illinois	A	A	A
Indiana	S,G	S,G	S,G
Iowa			
Kansas			
Kentucky	S	S	S
Louisiana	*	*	
Maine			
Maryland	A	A	A
Massachusetts			
Minnesota			
Mississippi			
Missouri			
Montana	*	*	
Nebraska			
Nevada			
New Jersey	A	A	A
New Mexico	G	G	G
New York			
North Carolina	S	S	S
North Dakota			
Ohio	A,G		
Oklahoma	A	A	*
Oregon	A,G	A,G	A,G
Pennsylvania			
Rhode Island			
South Carolina	A,G	A,G	A,G
South Dakota	A	A	A
Tennessee	A	A	A
Utah	A		
Vermont	A	A	A
Virginia	A,S	A,S	A,S
West Virginia	A,G		
Wisconsin	A	A	A
Wyoming	A		
Number of A	20	16	16
Number of S	5	5	5
Number of G	8	6	6

*= data not available or not included in this research

SOURCE: U.S. Department of Education, National Center for Education Statistics, Common Core of Data, Local Education Agency Universe Survey, internal records on dropout verification.

4. DEVELOPING AN APPROACH TO COMPENSATE FOR NONCONFORMANCE

This chapter describes the research to determine the effects of nonconformance in reporting dropout data and presents a model-based approach to compensate for differences. Section 4.1 discusses the concepts of a modeling approach, and the merits and limitations of this method. Section 4.2 describes the process of model development, the data used for modeling, and preliminary exploratory data analyses. Section 4.3 summarizes the estimation of the model parameters and the results of applying the estimation model to three years of dropout data: 1993, 1994, and 1995. The effects of nonconformance are modeled both on the combined grades 9 through 12 dropout data and separately by grade for each of grades 9 through 12. The estimation by grade is needed to calculate a high school completion rate using the longitudinal method.

4.1 A Possible Approach to Investigate the Effects of Nonconformance

This section formulates the mathematical equations to express the deviations associated with nonconformance and discusses a modeling approach to estimate the unknowns in the equations. Section 4.1.1 states the types of deviations from the CCD definition while Section 4.1.2 presents the estimation model and discusses the pros and cons of a multivariate modeling approach to estimate the unknowns in the equations.

4.1.1 Deviations From CCD Definition

One method to examine the effects of nonconformance is to consider these effects as deviations from the CCD definition. Let y_i denote the reported dropout count in the CCD agency file for district i. Since the reporting of y_i is not standardized, it can be expressed as an arithmetic sum of components affected by reporting practices. An expression for y_i is the following:

$$y_i = y_i^{CCD} + \sum_{j=1}^{3} \Delta y_{ij} \qquad (4)$$

where

y_i^{CCD}	is the dropout count by using the CCD definition;
Δy_{ij}	is the deviation due to the nonconforming practice j; and
j=1,2,3	for the practices: alternative reporting calendar, summer dropout, and adult GED.

For districts in states that follow the CCD definition, the reported dropout counts are the CCD-definition counts; that is, $y_i = y_i^{CCD}$. For districts that reported nonconforming data, the reported counts are not the CCD-definition count; that is, $y_i \neq y_i^{CCD}$. The counts of these districts deviate from the CCD-definition count, and the extent of deviation depends on the type of nonconforming practice. For example, for districts that reported using the alternative reporting calendar practice, the deviation is denoted as Δy_{i1}, and $y_i = y_i^{CCD} + \Delta y_{i1}$. For districts that reported using the summer dropout practice, the deviation is denoted as Δy_{i2}, and $y_i = y_i^{CCD} + \Delta y_{i2}$. Likewise, for districts that reported using the adult GED practice, the deviation is denoted as Δy_{i3}, and $y_i = y_i^{CCD} + \Delta y_{i3}$.

For districts that are different because of more that one variant practice, the effect of each practice is assumed to be independent and additive. For example, for districts that use all three nonconforming practices, $y_i = y_i^{CCD} + \Delta y_{i1} + \Delta y_{i2} + \Delta y_{i3}$. This formulation assumes no interaction effects between the nonconforming practices. This assumption is reasonable since each practice refers to a separate subset of dropout students, and the use of one practice does not presuppose the concurrent use of another practice.

For districts reporting with nonconformance, the CCD-definition dropout count, y_i^{CCD} is unknown. One way to provide an estimate, $\hat{y}_i^{CCD}$, for these districts is to begin by first estimating the deviation components $\Delta\hat{y}_{ij}$. Then, an adjusted $\hat{y}_i^{CCD}$ can be computed from equation (4) as follows:

$$\hat{y}_i^{CCD} = y_i - \sum_{j=1}^{3} \Delta\hat{y}_{ij} \qquad (5)$$

In this equation, the $\Delta\hat{y}_{ij}$ are unknown and need to be estimated.

4.1.2 A Modeling Approach to Estimate the Effects of Nonconformance

One possible method to estimate the effects of nonconformance in equation (5) is a multivariate modeling approach. This approach ascertains the net effects of nonconformance after holding constant the influence of other contributing factors. Factors that are known to influence dropout counts include differences by regions and Census divisions of the country, state educational policies, district characteristics, and the composition of schools and students within the districts. Therefore, it is necessary to control for these factors when comparing the differences associated with the standard and nonconforming practices. By employing a multivariate model, we can use the data available in the CCD files from all reporting states, test for significant differences between data that do and do not adhere to the CCD definition, and estimate the magnitude of the differences.

Estimation model. A possible estimation model is the following:

$$lt(r_i) = \beta_0 + \sum_{j=1}^{3} \beta_j x_{ij} + \sum_{k=1}^{23} \gamma_k z_{ik} + \varepsilon_i \qquad (6)$$

where

r_i	is the dropout rate for district i;
x_{ij}	are indicator variables, $x_{ij} = 1$ if district i reports with nonconforming practice j; $x_{ij} = 0$ if district reports with the CCD definition;
β_0	is the intercept parameter for the regression line;
β_j's	are the coefficients associated with the three indicators of nonconformity;
γ_k's	are the coefficients associated with the other explanatory covariates;
z_{ik}'s	are the other explanatory covariates, with k=1,2,..., 23 covariates in the model; and
ε_i	is the error parameter, assumed to have an independent $N(0,\sigma^2)$ distribution.

The dependent variable is specified as the logit (lt) dropout rate. This transformation is selected to conform with the model assumption that the dependent variable follows an approximately normal distribution. (The actual dropout rate is constrained between 0 and 1 and unlikely to follow a normal distribution.) The dropout rate, r_i, is computed as the percentage of dropouts for a given school year among student members enrolled as of October 1 of that school year. The logit transformation

$lt(r_i) = 0.5 * log\left(\frac{r_i}{1 - r_i}\right)$ was applied to yield quantities that are approximately normally distributed (see Johnson and Wichern, 1982).

The estimation model is defined at the district level because the CCD Local Education Agency Universe Survey contains district-level data, and district data are more suitable than state aggregate data for estimation. There are many more districts than states, hence there are more degrees of freedom and power in the analyses, and more explanatory variables can be included in the model for estimation.

The process of model development requires careful model specification, data management, and investigations to test, refine, evaluate, and improve the model. Winglee et al. (1997) discussed the initial efforts to develop an adjustment model and to determine the set of suitable explanatory variables in the model. The results of this work, and the current continuing efforts to improve the model, were presented to SEA representatives at the NCES Summer Data Conference in Washington DC, in 1997 and 1998. Both meetings were well attended, and SEA representatives at these meetings provided valuable feedback and comments to enhance the adjustment model.

A limitation of the modeling approach is that the model outcomes depend on the data available to include in the estimation, and all models are subject to errors. To clearly present dropout estimates derived through modeling, this study derived approximate estimates of standard errors (see Section 5.1) to reflect uncertainties in the model estimates. Confidence intervals are used to show the range of values that is likely to contain the true values.

4.2 Processes to Develop the Estimation Model

This section describes the processes used to develop an estimation model for the effects of the three nonconforming practices. Section 4.2.1 describes the dropout data and explanatory variables available to support modeling. Section 4.2.2 summarizes the exploratory analyses conducted on the data and the necessary steps to prepare the data to include in the model.

4.2.1 Data for Modeling

This study used the dropout data from the CCD, Public Elementary/Secondary Agency Universe Survey unpublished internal working files that included data from both standard and nonconforming states. Table 4-1 shows the number of districts with high school dropout data for 1993, 1994, and 1995. These are districts that reported high school students (students in any one of grades 9 to 12) and dropout counts (value could be zero, but not missing)[4].

The choice of which data to include as explanatory variables in the adjustment model depends on extant state and district level data sources. The sources used in this study included the 1990 School District Databook, the 1991-92 F-33 fiscal information file, *The Digest of Education Statistics* for 1993 through 1995, and other auxiliary sources that provide information about state policies. Analysts familiar with these databases reviewed the content and suggested an initial list of variables for investigation.

Table 4-2 shows the explanatory variables included in the final models. These variables were based on the results of the feasibility study by Winglee et al. (1997). In addition, several enhancements were added based on the suggestions of SEA representatives and other reviewers. For example, new variables were added to reflect the influence of state educational policies on the age of compulsory school attendance and the requirement for exit examinations before graduation.

The current model specification offers several improvements over the initial research by Winglee et al. With better understanding of the types of nonconforming practices, the original model was respecified to avoid collinearities and redundancies in the estimation parameters.

[4] For 1993, the dropout data come from the 1994 CCD Local Education Agency Universe Survey, and the membership data are from the 1993 Public Elementary/Secondary School Universe Survey, aggregated to district level. (The situation is similar for the following two years.)

Table 4-1. Number of districts reporting high school students and dropout data, 1993-1995

State†	Number of districts reporting dropout counts		
	1993	1994	1995
Alabama	123	127	127
Alaska	*	*	55
Arizona	105	103	103
Arkansas	309	311	314
California	441	441	459
Colorado	175	174	180
Connecticut	123	122	120
Delaware	19	19	19
District of Columbia	1	1	*
Florida	70	71	72
Georgia	173	173	172
Hawaii	1	1	1
Idaho	103	105	105
Illinois	513	507	515
Indiana	291	292	287
Iowa	341	344	352
Kansas	299	297	302
Kentucky	172	172	172
Louisiana	*	*	66
Maine	117	114	117
Maryland	24	24	24
Massachusetts	252	247	259
Minnesota	352	338	311
Mississippi	157	158	157
Missouri	448	448	449
Montana	*	*	154
Nebraska	304	281	300
Nevada	15	16	15
New Jersey	255	252	260
New Mexico	87	88	88
New York	644	650	650
North Carolina	119	119	118
North Dakota	186	183	186
Ohio	610	608	611
Oklahoma	433	432	*
Oregon	178	180	177
Pennsylvania	505	506	509
Rhode Island	33	33	33
South Carolina	93	93	93
South Dakota	174	175	179
Tennessee	123	124	125
Utah	39	38	41
Vermont	61	63	62
Virginia	131	131	131
West Virginia	56	56	56
Wisconsin	365	367	381
Wyoming	47	47	47
Total	9,067	9,031	8,954

†Washington, Michigan, New Hampshire never reported and the data from Texas are excluded from model because of data issues.
*Not reported or not used in this study.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Common Core of Data, Local Education Agency Universe Survey, unpublished internal working files, 1994-1996.

Table 4-2. List of explanatory variables

Variable	Description
I	Intercept (regression line)
	State practices:
x_A	Nonconformity: Alternative reporting calendar
x_S	Nonconformity: Summer dropout
x_G	Nonconformity: Adult GED
z_1 - EXAM	Requires graduation examination
z_2 - AGE 17	Compulsory age of school attendance is 17 years (versus 16 years)
z_3 - AGE 18	Compulsory age of school attendance is 18 years (versus 16 years)
	Census division* (versus Northwest central):
z_4 - PC	Pacific
z_5 - SWC	Southwest central
z_6 - SEC	Southeast central
z_7 - NEC	Northeast central
z_8 - SA	South Atlantic
z_9 - MA	Mid-Atlantic
z_{10} - NE	New England
z_{11} - MNT	Mountain
	District characteristics:
z_{12} - MSA	Serves metropolitan areas (MSAs versus other areas)
z_{13} - REG	Regular district (compared to other district type)
z_{14} - LARGE	Large student membership (over 1,000 students)
z_{15} - SMALL	Low student membership (fewer than 200 students)
z_{16} - INC	Median household income (log scale)
z_{17} - PUBASST	Percent of households receiving public assistance (log scale)
	Student composition:
z_{18} - BLACK	Black over 20 percent
z_{19} - HISP	Hispanic over 20 percent
z_{20} - INDIAN	American Indian over 5 percent
z_{21} - HIGHSP	Special education students over 14 percent
z_{22} - LOWSP	Special education students under 7 percent
z_{23} - RATIO	Pupil-teacher ratio (square root scale)

SOURCE: U.S. Department of Education, National Center for Education Statistics, Common Core of Data, Local Education Agency Universe Survey unpublished internal working files, 1994-1996; the 1990 School District Databook, the 1991-92 F-33 fiscal information file, the *Digest of Education Statistics* for 1993 through 1995, and other auxiliary sources.

*A map showing the states in each census division is included in Appendix B.

4.2.2 Exploratory Data Analyses and Data Preparation

The feasibility study by Winglee et al. (1997) accomplished most of the background work of constructing the variables, conducting exploratory data analyses to review the statistical properties of the variables, and testing for suitable transformations of the variables to attain normality in the data distributions. The exploratory analyses in that study included distributional plots to check for outlying (or extreme) values and tests for collinearity between variables.

The variables included in the model are constructed to satisfy the model requirements and the underlying model assumptions. For example, variables with discrete categories are coded as binary indicator variables with values of 0 or 1. To represent Census divisions, nine indicator variables were defined, one for each of the nine census divisions. For each district, only one of these nine variables has a value of 1, denoting that the district is contained in the division; the other eight variables have a value of 0, indicating that the district is not in these divisions. Eight of the nine Census division variables are included in the estimation model; one division (selected arbitrarily) is left out of the model as the reference group for comparisons.

The data distributions for continuous variables (such as income) were examined to identify outlying values that may have a large influence on the estimation. Extreme outlying values were either suppressed or set to a maximum value. The distributions of the data were tested for normality, and those distributions that failed the test were considered for transformation or recoding. For example, the median household income variable was normalized using the natural logarithm transformation. The variables on the percentage of students who are black, Hispanic, and American Indian were coded as indicator variables, to indicate high or low percentages of such students. These variables were recoded because there was no suitable transformation to attain normality. An initial set of explanatory variables was entered in the estimation equation and a method of backward elimination was used to delete variables that had no significant contribution to the variation in the predicted variable.

4.3 Estimation and Results

This section summarizes the method used to estimate the model parameters and the results. The parameters of the model are estimated using a weighted least square method (Draper and Smith, 1981). This method takes into account the variations in the size of the student population in the school districts. Some districts contain a much larger student population than other districts, and the weighted

method provides a greater contribution from larger districts. Section 4.3.1 shows the results of the estimation models on the combined grades 9 to 12 dropout data for 1993, 1994, and 1995. Section 4.3.2 summarizes the results of the same models applied by grade for the three years.

4.3.1 Models of the Combined Grades 9 to 12 Dropout Data

Table 4-3 shows the estimated regression coefficients, and the standard errors of the estimates, from models on the combined grades 9 to 12 dropout data. The estimates and the standard errors of estimates that are not statistically significant at $\alpha = 0.05$ (using a two-tailed test) are shown as "n.s." The overall performances of the estimation models for the three years 1993, 1994, and 1995 are fairly comparable, accounting for about 43 to 46 percent of the variability in the dropout data. This level of explanatory power is quite good given the nature of the data.

4.3.2 Models by Grade

Table 4-4 shows the estimated coefficients for the nonconforming practices, and the standard errors of the coefficients from the models by grade and year. These models used the same explanatory variables as the models on the combined grade analyses and only the results for the nonconforming practices are shown here.

The results of the analyses by grade are fairly consistent across the three years. Overall, the models for grades 9, 10, and 11 explained about 35 to 42 percent of the variations in the dropout rates. The models for grade 12, however, have less explanatory power, accounting for about 22 to 23 percent of the variation in dropout rates. The estimated effects of nonconformance are fairly comparable for 1994 and 1995. For both years, the largest effect was due to the GED practice. The alternative reporting calendar has consistently significant but small effects. The summer dropout practices had mostly nonsignificant effects by grade. The estimation for 1993 showed nonsignificant summer dropouts for grades 10 and 11.

Table 4-3. Estimates of regression coefficients and standard errors for the combined grades 9 to 12 adjustment models, 1993-1995

Variable	Regression coefficient			Standard error		
	1993	1994	1995	1993	1994	1995
I	4.63	3.56	4.10	0.25	0.22	0.26
x_A	0.03	0.03	0.06	0.01	0.01	0.01
x_S	n.s.	0.11	0.07	n.s.	0.01	0.02
x_G	-0.16	-0.26	-0.30	0.01	0.01	0.02
z_1 - EXAM	0.04	0.05	0.08	0.01	0.01	0.01
z_2 - AGE 17	-0.07	-0.03	n.s.	0.01	0.01	n.s.
z_3 - AGE 18	-0.23	-0.20	-0.24	0.02	0.01	0.02
z_4 - PC	0.24	-0.24	0.27	0.03	0.03	0.03
z_5 - SWC	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
z_6 - SEC	-0.20	-0.18	-0.27	0.03	0.02	0.03
z_7 - NEC	0.05	n.s.	n.s.	0.02	n.s.	n.s.
z_8 - SA	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
z_9 - MA	-0.22	-0.25	-0.28	0.02	0.02	0.02
z_{10} - NE	n.s.	-0.06	-0.07	n.s.	0.02	0.02
z_{11} - MNT	0.18	0.17	0.17	0.02	0.02	0.02
z_{12} - MSA	0.15	0.15	0.16	0.01	0.01	0.01
z_{13} - REG	n.s.	0.10	n.s.	n.s.	0.04	n.s.
z_{14} - LARGE	0.18	0.18	0.17	0.01	0.01	0.01
z_{15} - SMALL	-0.20	-0.18	-0.19	0.03	0.03	0.03
z_{16} - INC	-0.58	-0.50	-0.54	0.03	0.02	0.03
z_{17} - PUBASST	0.17	0.22	0.22	0.02	0.02	0.03
z_{18} - BLACK	0.11	0.14	0.15	0.01	0.01	0.01
z_{19} - HISP	0.16	0.13	0.13	0.01	0.01	0.01
z_{20} - INDIAN	n.s.	0.10	n.s.	n.s.	0.02	n.s.
z_{21} - HIGHSP	n.s.	-0.06	n.s.	n.s.	0.01	n.s.
z_{22} - LOWSP	0.04	0.02	0.05	0.01	0.01	0.01
z_{23} - RATIO	-0.32	n.s.	-0.81	0.05	n.s.	0.03
Model R^2	0.43	0.43	0.46			

n.s. = not significant at $\alpha = 0.05$ for a two-tailed test.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Common Core of Data, Local Education Agency Universe Survey, unpublished internal working files, 1994-1996.

Table 4-4. Estimates of regression coefficients and standard errors for adjustment models by grade, 1993-1995

Nonconforming practice	Regression coefficient* Grade 9	10	11	12	Standard error of coefficients Grade 9	10	11	12
1993								
x_A: Alternative reporting calendar	0.05	0.03	0.05	0.08	0.02	0.01	0.01	0.01
x_S: Summer dropout	0.16	n.s.	n.s.	-0.14	0.03	n.s.	n.s.	0.02
x_G: Adult GED	-0.20	-0.16	-0.19	-0.20	0.03	0.02	0.02	0.02
Model R^2	0.37	0.41	0.36	0.22				
1994								
x_A: Alternative reporting calendar	0.04	0.05	0.05	0.04	0.01	0.01	0.01	0.01
x_S: Summer dropout	0.18	0.13	0.10	n.s.	0.02	0.02	0.02	n.s.
x_G: Adult GED	-0.24	-0.22	-0.25	-0.25	0.02	0.02	0.02	0.02
Model R^2	0.39	0.39	0.35	0.22				
1995								
x_A: Alternative reporting calendar	0.07	0.06	0.07	0.09	0.02	0.01	0.01	0.01
x_S: Summer dropout	0.17	0.07	0.03	-0.07	0.03	0.02	0.02	0.02
x_G: Adult GED	-0.28	-0.27	-0.29	-0.25	0.02	0.02	0.02	0.02
Model R^2	0.42	0.42	0.37	0.23				

*The estimates for the covariates are not shown in this table.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Common Core of Data, Local Education Agency Universe Survey, unpublished internal working files, 1994-1996.

4.3.3 Summary of Results from Model Estimation

This study applied the same estimation model to examine the effects of nonconformity in 1993, 1994, and 1995. The major findings are summarized below. The results are fairly consistent across the three years, in that there are statistically significant differences associated with each of the alternative calendar and adult GED nonconforming practices. The sizes of the differences, however, vary to some extent by year and by grade.

- The effects associated with the alternative reporting calendar practice are small but consistently positive, suggesting that this practice is associated with a slight overreporting of dropouts relative to the CCD definition.
- The effects of summer dropouts are less consistent and the magnitude of the effect is not always significant by grade.

- The effects due to adult GED are consistently statistically significant and negative in all three years. Districts that are nonconforming because of the adult GED practice are reporting dropout rates lower than districts that conform with the CCD definition.

- The estimated effects associated with the nonconforming practices are fairly consistent for 1994 and 1995. The slight variation in the size of the estimates by year may be a function of the change in reporting status for some states. Another possibility is that the explanatory variables based on Census information of the average family income and percentage of families receiving public assistance may be dated as the model ages. Further research may use a modeling approach that combines these two years of data to derive a more stable estimate.

- Overall, the statistically significant effects of nonconformance highlight the need for compensation before the dropout data from conforming and nonconforming states can be compared.

5. AN ADJUSTMENT METHOD TO COMPENSATE FOR NONCONFORMANCE

This chapter describes a method to compute adjusted CCD dropout counts with compensation for the effects of nonconforming practices. High school dropout and completion rates are computed with adjustments to facilitate state-level comparisons. Section 5.1 discusses the method that was used. Section 5.2 shows the results of the adjusted dropout and completion rates for 1993, 1994, and 1995, and the differences between the adjusted and reported rates.

5.1 Method to Adjust for the Effects of Nonconformance

A method to adjust for nonconformance is to apply the $\hat{\beta}_j$ parameters estimated for equation (6) to estimate the deviation parameters $\Delta\hat{y}_{ij}$ in equation (5). To do this, the first step is to apply the $\hat{\beta}_j$ in equation (6) to compute an adjusted logit dropout rate as follows: $lt(\hat{r}_i) - \sum_{j=1}^{3} \hat{\beta}_j x_{ij}$. Then the logit transformation is reversed to obtain an adjusted dropout rate, and the adjusted rate is multiplied by the student membership to obtain an adjusted CCD-definition dropout count, $\hat{y}_i^{CCD}$.

To estimate the modeling error, the variance and covariance of the $\hat{\beta}_j$ can be used to compute the 95 percent confidence interval of the adjusted logit dropout rate as follows: $\left[lt(\hat{r}_i) - \sum_{j=1}^{3} \hat{\beta}_j x_{ij} \right] \pm 1.96 \sqrt{var\left(\sum_{j=1}^{3} \hat{\beta}_j x_{ij} \right)}$. Then, the above sequence of steps is repeated to reverse the logit transformation in order to obtain the 95 percent confidence interval of the adjusted dropout rate, and to multiply the dropout rate confidence bounds by their membership to obtain the confidence interval of dropout counts.

Using adjusted dropout counts, an adjusted CCD-definition dropout rate can be computed as follows:

$$\hat{r}_s^{adj} = \frac{\sum_{i=1}^{n_s}\left(y_i - \sum_{j=1}^{3}\Delta\hat{y}_{ij}\right)}{\sum_{i=1}^{n_s} M_i} = \frac{\sum_{i=1}^{n_s}\hat{y}_i^{CCD}}{M_s} \qquad (7)$$

where M_s is the total student membership in state s. This equation applies to both the aggregate grade and the individual grade cases. For the aggregate grade case, the y_i and the M_i are the sums of the counts for grades 9-12. For the individual grade cases, they correspond to the counts for a given grade.

Estimating Variance. The variance of the adjusted state dropout rates, ignoring the covariance of districts within states, is approximated as follows:

$$var\left(\hat{r}_s^{adj}\right) = Var\left(\frac{\sum_{i=1}^{n_s}\hat{y}_i^{CCD}}{M_s}\right) \cong \frac{1}{M_s^2}\sum_{i=1}^{n_s} Var\left(\hat{y}_i^{CCD}\right) \qquad (8)$$

The modeling error of the adjusted state dropout rate is the square root of this variance and it reflects the error associated with the model estimation.

Likewise, the adjusted completion rates for the nonconforming states can be computed using the same method by substituting the adjusted dropout counts instead of the reported counts into equations (2) or (3).

5.2 Reported and Adjusted High School Dropout and Completion Rates by State

Tables 5-1 through 5-3 show the high school dropout rate and the synthetic completion rates for 1993, 1994, and 1995. The adjusted dropout rates for grades 9-12 were computed using the parameter estimates from the aggregate models, and the adjusted completion rates used the estimates by grade. For consistency, all rates are shown with two significant digits, and their standard errors with one significant

digit. These tables show both a reported rate and an adjusted rate for each statistic. The adjusted rate is an estimate of the CCD-definition rate derived using the method described in Section 5.1. For states that conform to the CCD definition, the reported and the adjusted rate are the same because no adjustment is necessary. For states that did not conform, the reported rates are affected by the reporting practices. The adjusted rates for these states are estimates of the CCD definition rate, and approximate estimates of the modeling error are shown in parentheses to indicate the errors associated with model estimation. For example, in 1993 West Virginia reported an annual dropout rate at 3.8 percent; when adjusted for nonconformity this became 4.8 percent. In 1995, West Virginia again reported a 3.8 percent dropout rate; this was not adjusted since the state was now in agreement with the CCD definition.

These tables contain over 100 adjusted estimates for individual states. To test whether each of these adjustments are significantly different from the reported rates, it is necessary to control for the effect of multiple comparisons. Using the conservative Bonferroni procedure, differences were required to exceed 2.8 times the estimated standard error to be considered statistically significant at the 95 percent confidence level. Comparisons that are significantly different are indicated with an asterisk. With the exception of a few states, the adjusted estimates are not significantly different from those reported. In fact, a closer examination shows that the differences are not significant for the states that are nonconforming simply because of an alternative reporting year or summer dropout. The significant differences are in states that do not follow the CCD procedures for adult GED.

In general, the percentage change in the dropout rate due to adjustment is small for the alternative reporting calendar practice, but more substantial for the summer dropout and the adult GED practices. Depending on the variant practice, the percentage difference between the adjusted and the reported dropout rates ranges from -23 to 72 percent. The relative change associated with the adjustments on the completion rates is smaller; here, the percentage change ranges from about -19 to 12 percent. Since relatively few states use the summer dropout or the adult GED practice, which account for the larger adjustments, the overall effect of adjustment nationally across all reporting states is small. It ranges from between 7 and 10 percent for the dropout rate, and between -2 and -3 percent for the completion rate.

Completion rates for 1995. The synthetic completion rates for 1995 (shown in Table 5-3) are shown along with the longitudinal rate in Table 5-4. For most states, the two rates are very similar. The longitudinal rate is theoretically more precise because it uses a pseudo-cohort approach, and it takes into account the reporting practice of dropout data each year. The synthetic rate, however, is a close approximation for most states even though it relies on a single year of data.

Longitudinal completion rates are computed only for 1995 because of data limitations. This rate requires dropout data for four consecutive years. However, unpublished internal working files with dropout data from both standard and nonconforming states were available for 1993 through 1995 only. Therefore, the 1995 longitudinal rate shown here used the dropout counts from grade 12 for 1995, grade 11 for 1994, and both grades 9 and 10 for 1993. (The grade 9 dropout count should have been that from 1992.) As dropout data collection continues in the CCD, the calculation of a longitudinal completion rate becomes more feasible.

Table 5-1. Reported and adjusted completion and dropout rates by state, 1993

State (nonconforming practice)	Annual dropout rate			Percentage Change	Synthetic completion rate			Percentage change
	Reported	Adjusted	(s.e.)		Reported	Adjusted	(s.e.)	
Alabama (A)	5.9	5.5	(0.2)	-7	76	77	(1)	1
Arizona (A)	13.7	12.7	(0.4)	-7	53	55	(1)	4
Arkansas	5.3	5.3			79	79		
California (G)	5.3	7.2	(0.5)*	36	77	71	(2)*	-8
Colorado (A)	7.3	6.9	(0.3)	-5	73	74	(1)	1
Connecticut	4.9	4.9			81	81		
Delaware	4.6	4.6			80	80		
District of Columbia	9.6	9.6			65	65		
Florida (A,G)	5.5	7.0	(0.4)*	27	76	71	(2)*	-7
Georgia	8.7	8.7			67	67		
Hawaii (S)	5.1	5.0	(0.4)	-2	80	80	(1)	0
Idaho (A)	8.5	8.0	(0.4)	-6	69	71	(1)	3
Illinois (A)	6.9	6.5	(0.3)	-5	74	75	(1)	1
Indiana (S,G)	4.6	6.1	(0.6)*	33	81	76	(2)*	-6
Iowa	3.2	3.2			87	87		
Kansas	5.0	5.0			80	80		
Kentucky (S)	5.5	5.4	(0.4)	-2	69	69	(1)	0
Maine	3.1	3.1			87	87		
Maryland (A)	5.3	4.9	(0.2)	-8	79	80	(1)	1
Massachusetts	3.7	3.7			84	84		
Minnesota	5.1	5.1			80	80		
Mississippi	6.1	6.1			76	76		
Missouri	7.1	7.1			73	73		
Nebraska	4.6	4.6			82	82		
Nevada	9.8	9.8			63	63		
New Jersey (A)	4.2	4.0	(0.2)	-5	83	84	(1)	1
New Mexico (G)	8.1	10.8	(0.6)*	33	68	61	(2)*	-10
New York	4.0	4.0			82	82		
North Carolina (S)	5.9	5.9	(0.4)	0	76	76	(1)	0
North Dakota	2.7	2.7			89	89		
Ohio (A,G)	4.7	6.0	(0.4)*	28	80	76	(1)*	-5
Oklahoma (A)	4.6	4.3	(0.2)	-5	81	82	(1)	1
Oregon (A,G)	7.2	9.2	(0.6)*	28	73	68	(2)*	-7
Pennsylvania	3.8	3.8			84	84		
Rhode Island	4.9	4.9			80	80		
South Carolina (A,G)	3.0	3.9	(0.4)*	30	86	83	(2)*	-4
South Dakota (A)	5.3	5.0	(0.2)	-5	80	81	(1)	1
Tennessee (A)	4.9	4.6	(0.2)	-5	78	79	(1)	1
Utah (A)	3.2	3.0	(0.1)	-5	86	87	(1)	1
Vermont (A)	4.8	4.6	(0.2)	-5	81	82	(1)	1
Virginia (A,S)	4.8	4.5	(0.4)	-6	81	82	(1)	1
West Virginia (A,G)	3.8	4.8	(0.4)*	26	85	81	(1)*	-5
Wisconsin (A)	3.1	2.9	(0.1)	-7	86	87	(1)	1
Wyoming (A)	6.5	6.1	(0.3)	-5	76	77	(1)	1
All reporting states	5.4	5.8	(0.2)	7	78	76	(1)	-3

s.e. = Estimates of standard error; A = Alternative reporting calendar, S = Summer dropouts, G = Adult GED.

* The difference between the reported and adjusted rates is statistically significant at the 95 confidence level using a Bonferroni test adjusting for multiple comparisons.

NOTE: This table presents a possible approach to adjusting state-reported nonconforming dropout rate and calculating completion rates.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Common Core of Data, Local Education Agency Universe Survey, unpublished internal working files.

BEST COPY AVAILABLE

Table 5-2. Reported and adjusted completion and dropout rates by state, 1994

State (nonconforming practice)	Annual dropout rate			Percentage Change	Synthetic completion rate			Percentage Change
	Reported	Adjusted	(s.e.)		Reported	Adjusted	(s.e.)	
Alabama (A)	6.2	5.8	(0.2)	-7	76	77	(1)	1
Arizona (A)	9.5	9.0	(0.3)	-5	63	64	(1)	2
Arkansas	5.0	5.0			80	80		
California (G)	4.4	7.1	(0.4)*	61	80	71	(2)*	-11
Colorado (A)	7.3	6.9	(0.3)	-6	72	74	(1)	3
Connecticut	4.9	4.9			80	80		
Delaware	4.6	4.6			79	79		
District of Columbia	10.6	10.6			62	62		
Florida (A,G)	4.8	7.2	(0.5)*	50	77	69	(1)*	-10
Georgia	9.0	9.0			66	66		
Hawaii (S)	4.9	4.0	(0.4)	-18	80	83	(2)	4
Idaho (A)	9.2	8.7	(0.4)	-5	69	70	(1)	1
Illinois (A)	6.6	6.2	(0.3)	-6	75	76	(1)	1
Indiana (S,G)	4.6	6.1	(0.6)*	33	81	77	(2)*	-5
Iowa	3.5	3.5			86	86		
Kansas	5.1	5.1			80	80		
Kentucky (S)	5.9	4.8	(0.4)	-19	77	80	(2)	4
Maine	3.4	3.4			86	86		
Maryland (A)	5.2	4.8	(0.2)	-8	80	81	(1)	1
Massachusetts	3.6	3.6			85	85		
Minnesota	5.2	5.2			79	79		
Mississippi	6.4	6.4			75	75		
Missouri	7.1	7.1			73	73		
Nebraska	4.5	4.5			83	83		
Nevada	10.3	10.3			62	62		
New Jersey (A)	4.0	3.7	(0.2)	-8	84	85	(1)	1
New Mexico (G)	8.5	13.3	(0.6)*	57	66	55	(1)*	-17
New York	4.1	4.1			81	81		
North Carolina (S)	6.0	4.9	(0.4)	-18	76	80	(1)	5
North Dakota	2.5	2.5			90	90		
Ohio	5.3	5.3			77	77		
Oklahoma (A)	5.8	5.4	(0.2)	-7	77	79	(1)	3
Oregon (A,G)	7.1	10.7	(0.5)*	51	74	65	(2)*	-12
Pennsylvania	4.1	4.1			83	83		
Rhode Island	4.6	4.6			81	81		
South Carolina (A,G)	3.1	4.8	(0.5)*	55	85	79	(2)*	-7
South Dakota (A)	5.3	5.0	(0.2)	-6	79	80	(1)	1
Tennessee (A)	5.0	4.7	(0.2)	-6	79	80	(1)	1
Utah	3.6	3.6			85	85		
Vermont (A)	4.7	4.4	(0.2	-6	82	83	(1)	1
Virginia (A,S)	5.2	4.0	(0.4)	-23	79	83	(1)	5
West Virginia	4.2	4.2			83	83		
Wisconsin (A)	2.7	2.5	(0.1)	-7	88	89	(1)	1
Wyoming	6.7	6.7			75	75		
Reporting states	5.2	5.7	(0.2)	10	79	77	(1)	-2

s.e. = Estimates of standard error; A = Alternative reporting calendar, S = Summer dropouts, G = Adult GED.

* The difference between the reported and adjusted rates is statistically significant at the 95 confidence level using a Bonferroni test adjusting for multiple comparisons.

NOTE: This table presents a possible approach to adjusting state-reported nonconforming dropout rate and calculating completion rates.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Common Core of Data, Local Education Agency Universe Survey, unpublished internal working files.

Table 5-3. Reported and adjusted completion and dropout rates by state, 1995

State (nonconforming practice)	Annual dropout rate			Percentage Change	Synthetic completion rate			Percentage Change
	Reported	Adjusted	(s.e.)		Reported	Adjusted	(s.e.)	
Alabama (A)	5.6	5.0	(0.2)	-11	77	79	(1)	2
Alaska (A)	5.6	5.0	(0.2)	-11	76	78	(1)	3
Arizona (A)	9.5	8.6	(0.2)	-10	61	64	(1)	5
Arkansas	4.1	4.1			83	83		
California (G)	3.9	6.7	(0.3)*	72	82	72	(2)*	-12
Colorado (A)	6.7	6.0	(0.3)	-11	74	76	(1)	3
Connecticut	4.7	4.7			81	81		
Delaware	4.5	4.5			80	80		
Florida (A,G)	4.9	7.7	(0.4)*	57	77	68	(2)*	12
Georgia	8.4	8.4			67	67		
Hawaii (S)	4.7	4.1	(0.4)	-13	81	83	(2)	3
Idaho (A)	8.0	7.1	(0.4)	-11	72	74	(1)	3
Illinois (A)	6.5	5.8	(0.3)	-11	75	77	(1)	3
Indiana (S,G)	3.5	5.4	(0.6)*	54	85	78	(2)*	-8
Iowa	3.1	3.1			87	87		
Kansas	4.7	4.7			80	80		
Kentucky (S)	5.8	5.0	(0.4)	-14	77	79	(1)	3
Louisiana	11.6	11.6			60	60		
Maine	3.1	3.1			87	87		
Maryland (A)	4.8	4.3	(0.2)	-10	80	82	(1)	2
Massachusetts	3.3	3.3			86	86		
Minnesota	5.2	5.2			79	79		
Mississippi	6.2	6.2			75	75		
Missouri	6.6	6.6			74	74		
Montana	5.6	5.6			78	78		
Nebraska	4.5	4.5			82	82		
Nevada	9.7	9.7			63	63		
New Jersey (A)	4.0	3.6	(0.2)	-10	84	85	(1)	2
New Mexico (G)	8.3	13.9	(0.6)*	68	67	54	(1)*	-19
New York	3.7	3.7			83	83		
North Carolina (S)	6.0	5.2	(0.4)	-13	75	78	(1)	4
North Dakota	2.5	2.5			90	90		
Ohio	5.4	5.4			78	78		
Oregon (A,G)	6.9	10.6	(0.6)*	54	74	65	(2)*	-12
Pennsylvania	4.0	4.0			84	84		
Rhode Island	4.6	4.6			81	81		
South Carolina (A,G)	2.9	5.0	(0.5)*	72	86	79	(2)*	-8
South Dakota (A)	5.7	5.0	(0.2)	-12	78	80	(1)	2
Tennessee (A)	5.0	4.5	(0.2)	-10	79	80	(1)	2
Utah	4.4	4.4			81	81		
Vermont (A)	5.3	4.7	(0.2)	-11	79	81	(1)	2
Virginia (A,S)	4.7	3.6	(0.3)*	-23	81	84	(1)*	4
West Virginia	3.8	3.8			85	85		
Wisconsin (A)	2.4	2.1	(0.1)	-12	89	90	(1)	1
Wyoming	5.7	5.7			77	77		
Reporting states	5.1	5.6	(0.2)	10	79	77	(1)	-3

s.e. = Estimates of standard error; A = Alternative reporting calendar, S = Summer dropouts, G = Adult GED.

* The difference between the reported and adjusted rates is statistically significant at the 95 confidence level using a Bonferroni test accounting for multiple comparisons.

NOTE: This table presents a possible approach to adjusting state-reported nonconforming dropout rate and calculating completion rates.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Common Core of Data, Local Education Agency Universe Survey, unpublished internal working files.

BEST COPY AVAILABLE

Table 5-4. Reported and adjusted synthetic and longitudinal completion rates by state, 1995

State (nonconforming practice)	Synthetic Completion rate			Percentage Change	Longitudinal Completion rate			Percentage Change
	Reported	Adjusted	(s.e.)		Reported	Adjusted	(s.e.)	
Alabama (A)	77	79	(1)	2	77	78	(1)	1
Alaska (A)	76	78	(1)	3	n.a.	n.a.		
Arizona (A)	61	64	(1)	5	57	60	(1)	4
Arkansas	83	83			80	80		
California (G)	82	72	(2)*	-12	78	71	(2)*	-9
Colorado (A)	74	76	(1)	3	74	76	(1)	3
Connecticut	81	81			81	81		
Delaware	80	80			81	81		
Florida (A,G)	77	68	(2)*	12	77	71	(2)*	-8
Georgia	67	67			67	67		
Hawaii (S)	81	83	(2)	3	80	82	(2)	3
Idaho (A)	72	74	(1)	3	71	73	(1)	2
Illinois (A)	75	77	(1)	3	75	77	(1)	3
Indiana (S,G)	85	78	(2)*	-8	82	77	(1)*	-6
Iowa	87	87			87	87		
Kansas	80	80			80	80		
Kentucky (S)	77	79	(1)	3	79	80	(1)	1
Louisiana	60	60			n.a.	n.a.		
Maine	87	87			87	87		
Maryland (A)	80	82	(1)	2	80	82	(1)	2
Massachusetts	86	86			85	85		
Minnesota	79	79			80	80		
Mississippi	75	75			75	75		
Missouri	74	74			74	74		
Montana	78	78			n.a.	n.a.		
Nebraska	82	82			83	83		
Nevada	63	63			63	63		
New Jersey (A)	84	85	(1)	2	84	85	(1)	1
New Mexico (G)	67	54	(1)*	-19	68	59	(2)*	-13
New York	83	83			82	82		
North Carolina (S)	75	78	(1)	4	76	79	(2)	4
North Dakota	90	90			90	90		
Ohio	78	78			80	78	(1)	-3
Oregon (A,G)	74	65	(2)*	-12	73	67	(2)*	-8
Pennsylvania	84	84			84	84		
Rhode Island	81	81			81	81		
South Carolina (A,G)	86	79	(2)*	-8	86	82	(2)*	-5
South Dakota (A)	78	80	(1)	2	79	81	(1)	2
Tennessee (A)	79	80	(1)	2	79	81	(1)	2
Utah	81	81			86	86		
Vermont (A)	79	81	(1)	2	81	83	(1)	2
Virginia (A,S)	81	84	(1)*	4	81	83	(2)	3
West Virginia	85	85			84	82	(1)	-2
Wisconsin (A)	89	90	(1)	1	88	89	(1)	1
Wyoming	77	77			76	77	(1)	1
Reporting states	79	77	(1)	-3	79	77	(1)	-3

s.e. = estimates of standard error; A = Alternative reporting calendar, S = Summer dropouts, G = Adult GED. n.a. = not available.

* The difference between the reported and adjusted rates is statistically significant at the 95 confidence level using a Bonferroni test accounting for multiple comparisons.

NOTE: This table presents a possible approach to adjusting state-reported nonconforming dropout rate and calculating completion rates.

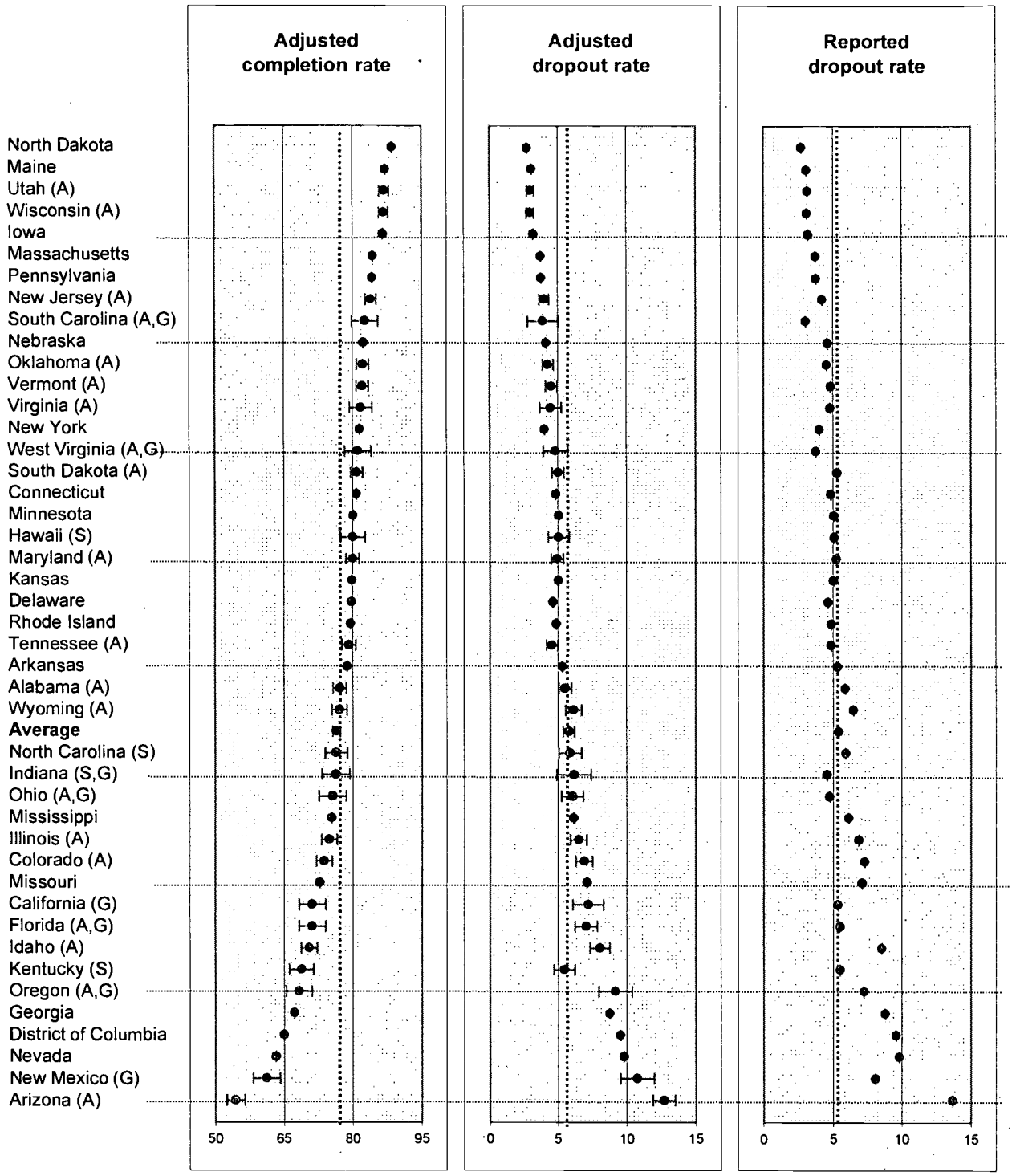
SOURCE: U.S. Department of Education, National Center for Education Statistics, Common Core of Data, Local Education Agency Universe Survey, unpublished internal working files.

5.3 Displays of Completion and Dropout Rates by State

Graphical displays are useful tools to summarize related statistics, facilitate the comparisons of multiple statistics by states, and convey additional information that is difficult to present in table format. Figures 5-1 through 5-3 display the data shown in the corresponding tables 5-1 through 5-3. The statistics shown include the adjusted synthetic high school completion rate, the adjusted dropout rate, and the reported dropout rate by states for 1993, 1994, and 1995. Likewise, figure 5-4 shows the data in table 5-4 on the adjusted longitudinal completion rate, the adjusted synthetic completion rate, and the reported longitudinal rates for 1995. The longitudinal method is preferred as dropout data become available in the CCD. The states are shown in order, sorted by the statistics in the first column. That is, states with higher completion rates are shown on top and ranked in descending order. The completion rate was used to determine the sort sequence because this statistic is less affected by the quality of the dropout data.

In each data panel, the dotted line shows the average for all reporting states. The distance of the data point from this line shows whether a state's rate is very different from the overall average of all reporting states. The 95 percent confidence interval of the adjusted rate (the error bars to the left and right of values for nonconforming states) shows the range of values that is likely to contain the true standardized rate for the state. (The intervals shown are for each individual estimate and do not include the Bonferroni adjustment for multiple comparisons.) The ranking applied to the adjusted completion rate corresponds reasonably well with the adjusted dropout rates. There are more irregularities in the reported rate panel where the effects of adjustment are large.

Figure 5-1. Reported and adjusted high school dropout (annual) and completion rates, 1993

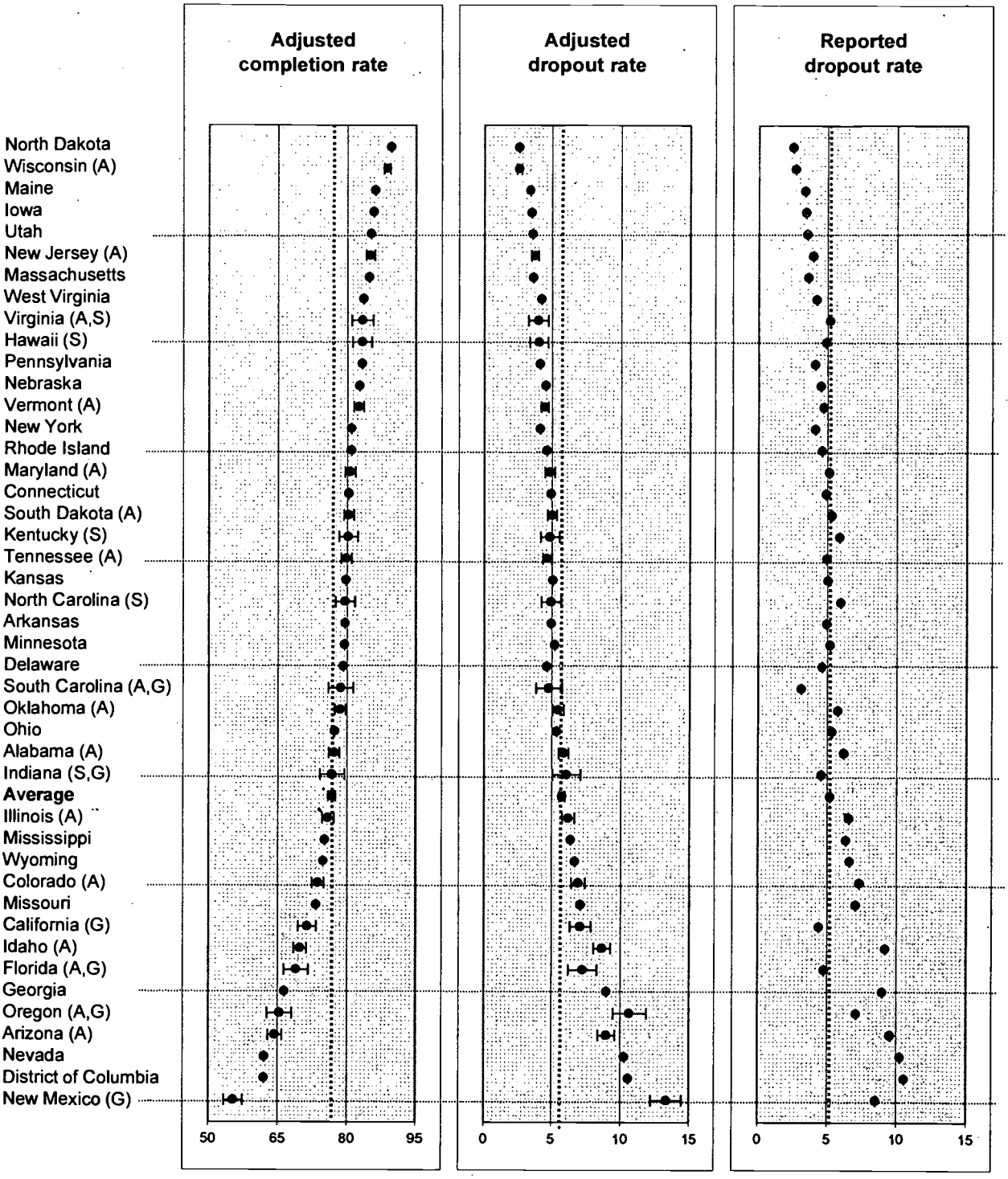


A=Alternative reporting calendar, S = Summer dropouts, G = Adult GED

Note:This figure presents a possible approach to adjusting state reported nonconforming dropout rates and calculating completion rates.

Source: National Center of Education Statistics, Common Core of Data, Public Elementary/Secondary Agency Universe Survey, unpublished files.

Figure 5-2. Reported and adjusted high school dropout (annual) and completion rates, 1994

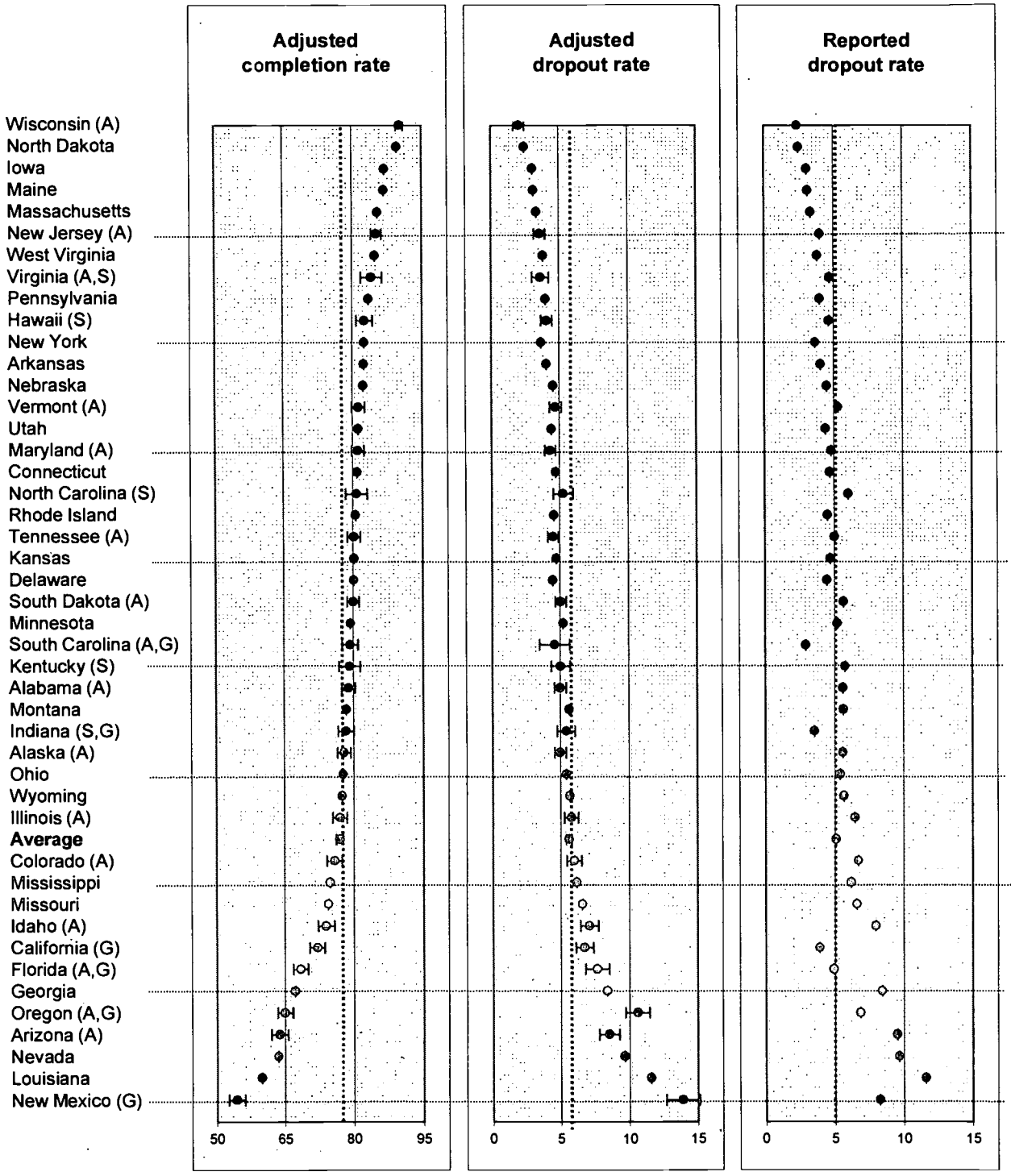


A=Alternative reporting calendar, S = Summer dropouts, G = Adult GED

Note:This figure presents a possible approach to adjusting state reported nonconforming dropout rates and calculating completion rates.

Source: National Center of Education Statistics, Common Core of Data, Public Elementary/Secondary Agency Universe Survey,unpublished file

Figure 5-3. Reported and adjusted high school dropout (annual) and completion rates, 1995

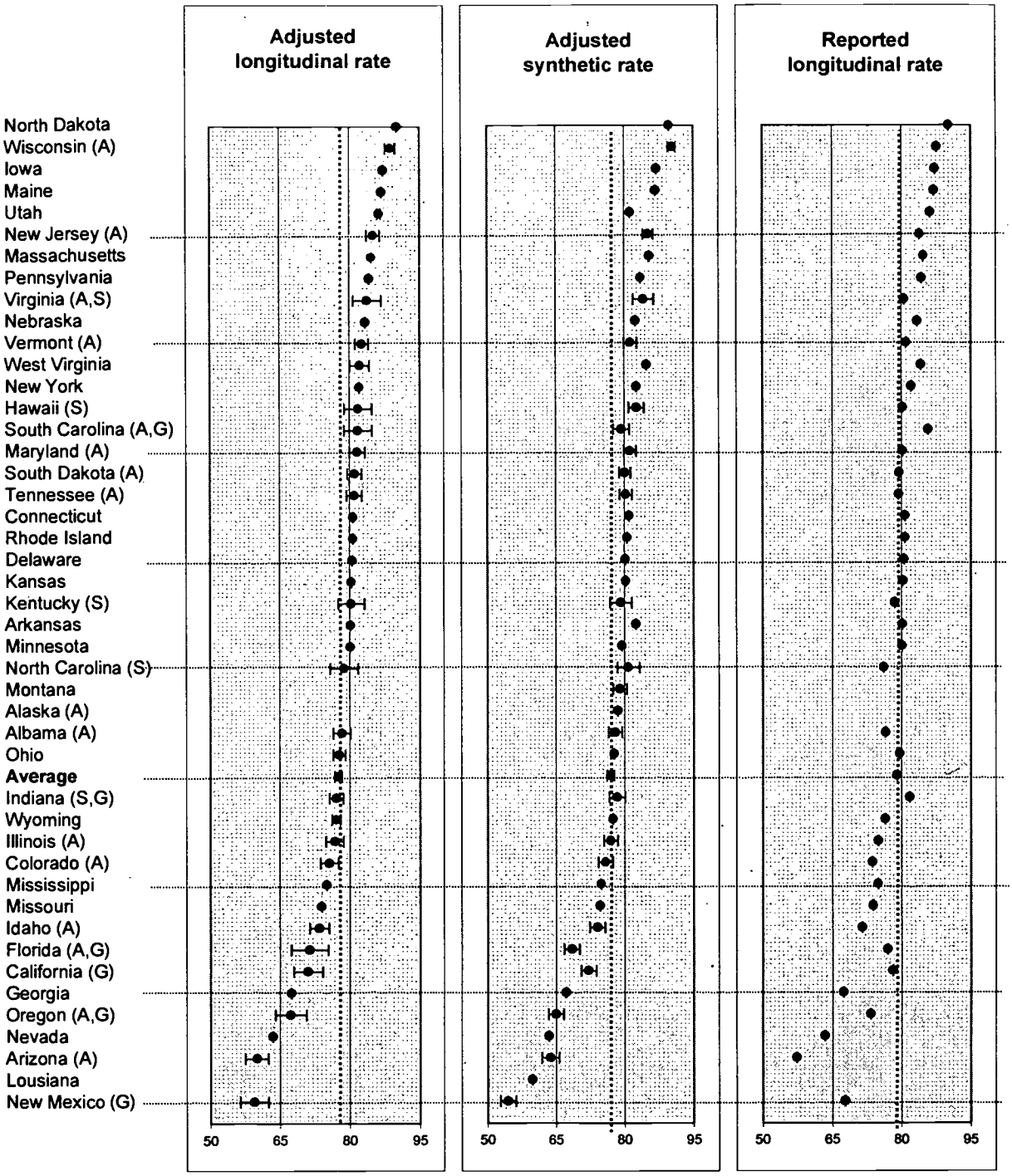


A=Alternative reporting calendar, S = Summer dropouts, G = Adult GED

Note:This figure presents a possible approach to adjusting state reported nonconforming dropout rates and calculating completion rates.

Source: National Center of Education Statistics, Common Core of Data, Public Elementary/Secondary Agency Universe Survey, unpublished file

Figure 5-4. Reported and adjusted high school completion rates, 1995



A=Alternative reporting calendar, S = Summer dropouts, G = Adult GED

Note:This figure presents a possible approach to adjusting state reported nonconforming dropout rates and calculating completion rates.

Source: National Center of Education Statistics, Common Core of Data, Public Elementary/Secondary Agency Universe Survey, unpublished file

6. CONCLUSION AND RECOMMENDATIONS

6.1 Summary of Findings

Quality of dropout data on the CCD. The quality concern addressed in this study of dropout data was the issue of definitional comparability. Reporting dropout data is quite complex. The CCD provides a standard dropout definition, and states are urged to follow this definition so that the reported data can be used for comparisons across states. Since the CCD began collecting dropout data with the 1991 school year, many states have reported the statistic each year. However, not all states have conformed with the reporting definition (see table 3.1). There have been improvements over the years, and the number of states reporting in agreement with the CCD has doubled between 1991 and 1995. However, the concern remains that some states are unable to change their reporting practice or are not making plans to do so. Therefore, a situation exists in which many states are willing to provide dropout data, but not all of the reported data can be used because of the issue of comparability.

Types of nonconformance. NCES has continued to monitor data quality and the types and extent of nonconformance are recorded each year. Staff at NCES or at the Census Bureau contact the CCD coordinators at State Education Agencies every year and ask about the state's reporting practices. The major nonconforming practices can be summarized as three distinct types: the alternative reporting calendar, the summer dropout, and the adult GED practices (table 3-2). The alternative reporting calendar practice affects the reporting of students who leave school and then re-enroll. The adult GED practice excludes from the dropout reports those students in adult education programs for General Educational Development. The summer dropout practice is more complex. It affects the grade and year to which those who drop out between school years are assigned (see table 3-3). Most nonconforming states differ from the CCD definition because of one of these practices; some states however, incorporate two variant practices (table 3-4).

A model-based approach to compensate for nonconforming practices. A possible method to compensate for the effects of definitional variation is through a multivariate modeling approach (Chapter 4), by applying model-based results to derive adjusted data after compensating for the effects of nonconformance. This approach was used to compare the differences between the data reported following the CCD definition and the data reported following variant practices after holding constant the influence of other extraneous factors. A multivariate modeling approach was possible because the extent

and the type of nonconforming practices were known, and data were available to support the modeling. The advantage of modeling is that it is possible to test for systematic differences between data reported with the CCD definition and the nonconforming practices, and to estimate the magnitude of these differences.

Processes of model development. The processes of model development require careful model specification, data management, and iterative steps of testing, refinement, and evaluation to improve the model. To prepare the data for modeling was not a trivial process. The process began with a review of extant data sources to select data suitable to include in the model; this was followed by exploratory data analyses to look at the data properties, presence of outliers, missing data, and the relationships between related data fields. Tests were then conducted to examine the distributions of the data, and the normality of the data distributions. For data items that failed this test, mathematical transformations were used to normalize the data such that the input data fit with the model requirements and underlying assumptions. Preliminary models were fitted and refitted successively. At each fitting, a backwards elimination procedure was used to delete, one at a time, those variables that failed to meet the threshold for statistically significant contribution to explaining the variation in the reported dropout data. The final set of explanatory variables was reviewed and further enhanced based on the recommendations of reviewers and SEA representatives.

Estimation. The overall results of estimation show systematic statistically significant differences between the CCD definition data and the nonconforming practices. The difference associated with the alternative reporting calendar practice is relatively small. For state estimates, the adjustment for this difference results in percentage changes of about 10 percent in the high school dropout rates by states, and less than 5 percent change in completion rates.

The differences associated with reporting variations for summer dropout and adult GED are larger. The adjustment for summer dropout results in percentage changes that range from 0 to 19 percent. The adjustment for adult GED results in a percentage change ranging from 33 to 72 percent. Since there are relatively few states that use either the summer dropout or the adult GED practice, very few states had large adjustments. More years of data and repeated tests of these estimates would be needed before recommending their use.

Calculating completion rates with CCD data. Through a cooperative process with states, NCES developed a CCD completion rate calculation. This rate uses grades 9 through 12 dropout and

completion data to produce an average completion rate; that is, of those students who leave school, what percentage of them graduated. This rate excludes those students who are still enrolled in school from that year's rate since they are neither dropouts nor completers.

Longitudinal Versus Synthetic Completion Rates. This study explored two methods to calculate high school completion rates using CCD data. The preferred way is a longitudinal rate that uses four years of dropout data. The intent is to mimic a true cohort approach by following consecutive grades through four years. A simpler method is a synthetic or a "reconstructive" cohort approach. This approach uses the current year data to reconstruct the past. The results from the two methods were similar in most states.

6.2 Conclusions

Reporting dropout rates of nonconforming states. One of the major purposes of this report was to discuss possible approaches to adjusting dropout rates. The results of the analysis support NCES using state dropout data of nonconforming states and calculating a state-level high school completion rate. There are a few possible approaches to how NCES could report dropout data.

The Status Quo. NCES could continue to report only those state dropout rates that conform with the original NCES definition. This would require no manipulation of the reported numbers. However, when verifying states' dropout reports, the majority of the states that do not conform reported no future plans to do so. This approach would mean that NCES could report only approximately 23 state dropout rates.

Use the regression model to report nonconforming states. The model used in this analysis can be used to adjust the dropout data of the non-conforming states so that their data can be reported with the other states. That procedure would allow NCES to publish data for all but those states that do not report any dropout data. On the other hand, the model development was difficult, the GED and Summer effects were large, and the model's explanatory power was moderate. Also, using a regression model to adjust reported figures can be difficult to explain to data users.

Report dropout data from states that use an alternative reporting calendar. The final option is to use unadjusted data from states that use the July-June reporting calendar as their only area of nonconformance. This would add 12 more states' dropout data to public files and reports. The difference associated with the alternative reporting calendar practice was found to have small impact on state-level comparisons. Hence for simplicity, no adjustment would be made to the reported data.

Further, those states who collect their data on the alternative reporting calendar might have the most difficulty in adjusting their systems to conform with the current definition since this would require modifying the state's entire reporting model as opposed to reclassifying some smaller group of students (e.g. adult GED participants) within an existing system. The major argument against accepting these nonstandard data would be that they would add a small amount of noncomparability to the NCES file.

Reporting state-level completion rates. The second purpose of this report was to develop a high school completion rate that could be used for state-level reports. Using available CCD data, the proposed completion rate seems a reasonable measure of how many students leave high school successfully. The synthetic version of this rate requires only one year of dropout and completion data to calculate. However, the longitudinal rate reflects changes in dropout rates from year to year and more accurately presents a state's completions.

6.3 Recommendations

Three factors were considered in reviewing the findings of this study: data quality, the feasibility of obtaining data, and the usefulness of resulting information. On the whole, accepting dropout data from states that use the July-June calendar appears the best compromise among all three. The effect of this variation is small. It tends to produce dropout rates that are slightly higher than those produced under the standard definition, which should encourage nonconforming states to change their reporting. Adding 12 states to the dropout file would greatly increase the data's usefulness.

While developing and testing the dropout adjustment model proved beneficial to the basic understanding of state reporting, the estimations were not stable enough over time, nor did the model account for enough of the variance to strongly support its use in adjusting reported data. These models should be tested again when more years of CCD dropout data are available. The dropout data are high-stakes statistics, and reporting potentially complicated "calculated" dropout numbers does not seem feasible at this time.

Finally, it appears feasible to produce a high school completion rate using CCD dropout and graduation data. The longitudinal rate is recommended because it is less affected by single-year or one-time changes in dropout rates within a district and thus would likely be more useful in assessing the need for, or success of, intervention programs.

6.4 Caveats

This analysis used the Local Education Agency Survey (school district) count of high school completers in calculating the completion rate because the dropout counts, which are also used in the completion rate, are reported on this survey. The sum of school district high school completers for a state may not match the numbers reported on the State Nonfiscal Survey of Public Elementary/Secondary Education. This is because high school equivalency recipients are reported only on the state-level survey and because there may be other legitimate differences – such as the granting of diplomas independent of any local school district – between the district and state surveys. As a result, variables found only in the state level survey cannot be readily used in analyses of the proposed completion rate.

Data users should also be cautious in applying either the dropout or completion rates to very small school districts. The fewer students a district enrolls, the more impact a single dropout or graduate can have on the total rate. Under these conditions, differences of one or two students can give the impression of considerable change.

REFERENCES

Draper, N., and Smith, H. (1987). *Applied Regression Analysis*, 2nd Edition, John Wiley and Sons, Inc.

Bose, J., and Hoffman, L. (1997). Developing a High School Graduation Rate Using CCD Data. Discussion paper April 1997, EIAC Meeting.

Hoffman, L.M. (1995). *State Dropout Data Collection Practices: 1991-1992 School Year*, NCES 95-690.

Johnson, R.A., and Wichern, D.W. (1982). *Applied Multivariate Statistical Analysis*, Prentice-Hall Inc., Englewood Cliffs, New Jersey 07632.

Kaufman, P., Klein, S., and Frase, M. (1999). *Dropout Rates in the United States: 1997.* NCES 1999-082.

McMillen, M., and Kaufman, P. (1998). *Dropout Rates in the United States: 1996*, NCES 98-250.

McMillen, M.M., Kaufman, P. and Klein, S. (1997). *Dropout Rates in the United States: 1995.* NCES 97-473.

National Center for Education Statistics, (1992). *National Dropout Statistics Field Test Evaluation*, Contractor Report, NCES 92-051.

Winglee, M., Marker, D., Henderson, A., and Webber, A. (1997). *A Dropout Rate Study* (Prepared under contract to the National Center for Education Statistics). Rockville, MD: Westat.

APPENDIX A. DEFINITIONS OF DROPOUTS AND COMPLETERS

CCD DROPOUT DEFINITION

The CCD dropout definition is based on a "snapshot" count of students at the beginning of the school year: A dropout is an individual who:

1. Was enrolled in school at some time during the previous school year* and was not enrolled on October 1 of the current school year; or
2. Was not enrolled on October 1 of the previous school year although expected to be in membership (i.e., was not reported as a dropout the year before); and
3. Has not graduated from high school or completed a state- or district-approved educational program and
4. Does not meet any of the following exclusionary conditions:
    i. Transfer to another public school district, private school, or state- or district-approved education program;
    ii. Temporary school-recognized absence due to suspension or illness; or
    iii. Death.

For purposes of applying this dropout definition, the following definitions also apply:

School year is the 12-month period beginning on October 1 and ending September 30. Thus, it includes the summer following the regular school year.

School completer is an individual who has graduated from high school or completed some other education program that is approved by the state or local education agency.

Students who completed a school year and failed to return to school in the subsequent year were counted as dropouts from the grade and school year for which they failed to enroll.

The event dropout rate was calculated as the number of dropouts for a given school year divided by membership on October 1 of that school year.

CCD HIGH SCHOOL COMPLETER CATEGORIES

There are three high school completion count categories on the Local Education Agency Universe Survey: regular diploma recipients, other diploma recipients, and other high school completers. These counts are taken at the end of the school year and the end of summer prior to that school year.

* A student who was enrolled in September, 1997; dropped out of school in February, 1998; and was not enrolled in school on October 1, 1998 would be reported as a 1997-1998 dropout (*previous* year) on the 1998-1999 CCD (*current* year).

BEST COPY AVAILABLE

These counts do not include high school General Education Development (GED) credential recipients. The definitions of these categories are as follows:

Regular Diploma Recipients. Individuals who received a regular diploma.

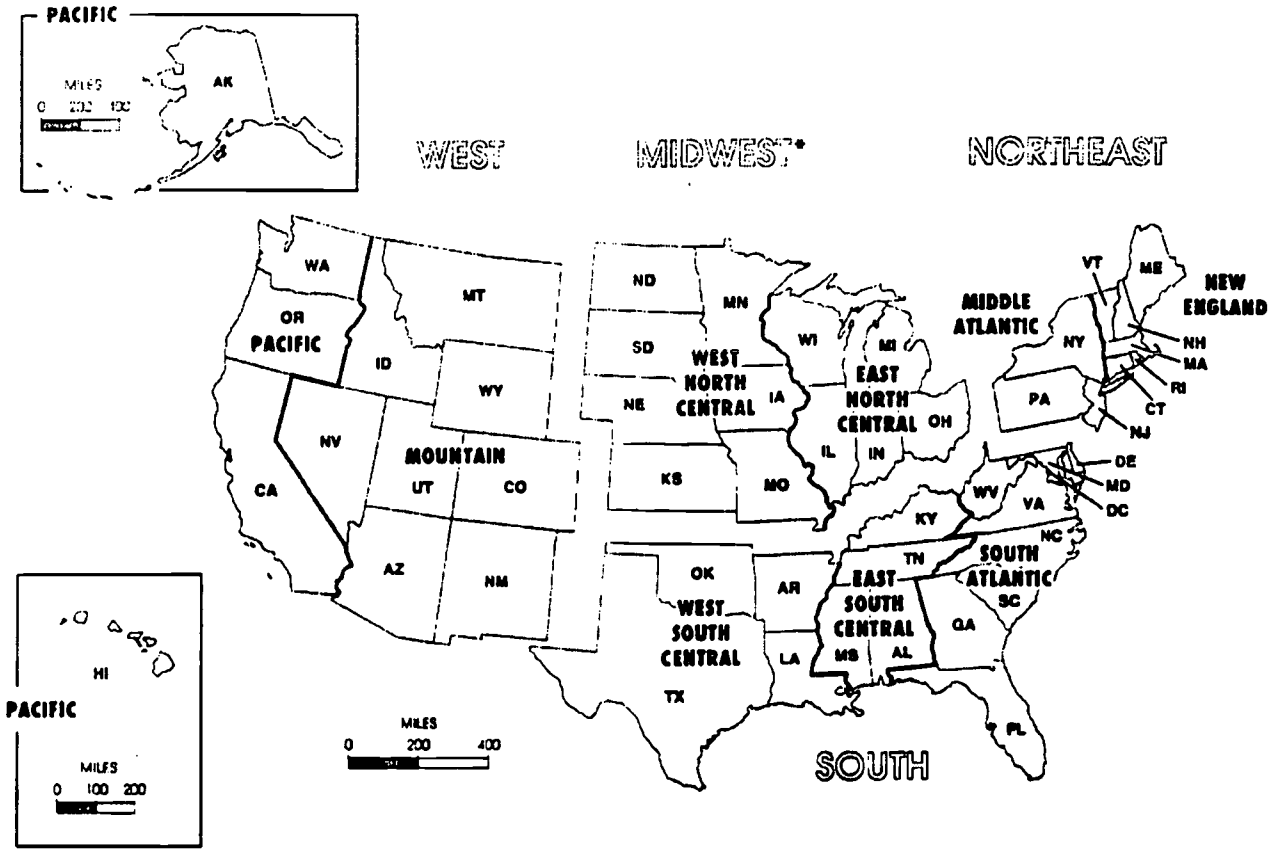
Other Diploma Recipients. Individuals who received a diploma from other than their regular school program.

Other High School Completers. Individuals who received a certificate of attendance, or other certificate of completion, in lieu of a diploma.

There are four high school completion count categories on the State Nonfiscal Survey of Public Elementary/Secondary Education: regular diploma recipients, other diploma recipients, other high school completers, and high school equivalency recipients.

High School Equivalency Recipients. Individuals age 19 years or younger who received a high school equivalency certificate. A high school equivalency certificate is a formal document certifying that an individual met the state requirements for high school graduation equivalency by obtaining satisfactory scores on an approved examination, and meet other performance requirements (if any) set by a state education agency or other appropriate body.

Appendix B. Census Regions and Divisions of the United States



* The Midwest Region was designated as the North Central Region until June 1984.

BEST COPY AVAILABLE

☆U.S. GOVERNMENT PRINTING OFFICE: 2000 463-878/20014

United States
Department of Education
Washington, DC 20208–5651

Official Business
Penalty for Private Use, $300

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

Standard Mail (A)

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

NOTICE

REPRODUCTION BASIS

☐ This document is covered by a signed "Reproduction Release (Blanket) form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

☑ This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").