DATA

For our analysis, we use 2003, 2005, and 2007 NAEP data for the mathematics and reading assessments. The NAEP items that we use are derived from the SD Background Questionnaire.[15] This questionnaire is filled out by the special education teacher or staff member who is most familiar with the student for each student who is both selected for NAEP participation and designated either as a student with disabilities or as an English language learner. For all SDs sampled for participation on NAEP, the questionnaire gathers information about the type of disability, the extent of the student's disability, and the type of instruction the student receives. Whether or not an SD can participate in the NAEP assessment is determined by the child's school, and the decision is supported by information in the SD Background Questionnaire.

The sample analyzed in this report is limited to SDs who are not also English language learners.[16] In addition to whether a student is included on the NAEP assessment, we use information on the type of disability the student has, the severity level of his or her disability, and indicators for the grade level of instruction compared with that of other students without a disability. For a portion of our analysis, we also use information on the type of accommodations, if any, the student received on state assessments.

Student disability type is included in our analysis as an explanatory variable for inclusion because some disabilities are easier to accommodate on NAEP assessments and some disabilities hinder learning more than others. NAEP allows disabilities to be classified according to 12 categories. In the 2003 NAEP, the teacher or staff member who filled out the SD Background Questionnaire was asked to pick one category that *best describes the student's primary disabilit*y. In 2005 and 2007, respondents were asked to indicate which of the 12 *describes the student's identified disability(ies)* and to check all that applied. Because the focus of analysis is the application to change from 2005 to 2007, the types of student disabilities are treated as not mutually exclusive in all years.

Although disabilities in NAEP are classified into 12 categories, many categories have very few students. In our model, we attempt to estimate separate effects for each type of disability, severity level, and grade level of instruction combination. For statistical power, we need an adequate amount of observations (students in our analysis) in each category. However, many of the 12 disability categories contained small numbers, so we explored options for collapsing them. We first consulted with an expert to help us combine categories thematically, grouping like disabilities together. This, however, still left us with some small categories and, since even 'like' categories were still heterogeneous, we felt this approach was problematic. Finally, our solution was to retain the largest 4 of the 12 categories and collapse the others, along with observations with missing disability type information, into a fifth category. In this approach, the largest categories are left homogeneous and only one heterogeneous category captures the remaining ones. Here, the effects for the largest categories will not be influenced by small categories having been folded into them. However, since so many heterogeneous categories are combined into the "other" category, its estimated effect will be an average of those combined categories and, therefore, one cannot draw conclusions about students in disability

15 The 2005 and 2007 versions of these questionnaires are included in appendix D. Additional information about the Background Questionnaires is available at http://nces.ed.gov/nationsreportcard/bgquest.asp.

16 As mentioned earlier, in the 2005 and 2007 mathematics and reading NAEP assessments, students with disabilities who were also English language learners made up 14 to 15 percent of all grade 4 students with disabilities and 16 to 19 percent of all grade 8 students with disabilities. The matter of how the exclusion of SDs who are also ELLs from the sample could affect the analyses in this report is, therefore, of importance. We expect SDs who are also ELLs to be included on NAEP under a very different process; as such, we expect that the model and, possibly, results will change by including them. This is the subject of the next set of reports which will be addressing the inclusion of English language learners and the joint SD and ELL populations.

categories that are folded into this broad category.[17] The categories used for analysis are listed in table 7.

Table 7. Categories of variables in regression equation

Student disability types (not mutually exclusive)	
1	Specific learning
2	Speech impairment
3	Mental retardation
4	Emotional disturbance
5	Other disability; Not reported [1]
Disability severity level	
1	Severe
2	Moderate
3	Mild
4	Not reported
Grade level of instruction	
1	Instruction at or above grade level
2	One year below grade level
3	Two or more years below grade level
4	Not reported; Not receiving instruction in this subject
Received accommodation on state assessment not allowed on NAEP	
0	Did not receive such an accommodation
1	Did receive such an accommodation

[1] Disabilities included in this category are: hearing impairment/deafness, visual impairment/blindness, orthopedic impairment, traumatic brain injury, autism, developmental delay (age 9 or younger), other health impairment, and any other nonspecified disabilities.

In our analysis, we also include a measure of the severity of a student's disabilities. The measure of severity, as asked in the SD Background Questionnaire, is an overall assessment of the degree of a student's disabilities and is not subject specific. When we hold a student's type of disability constant, the more severe the disability, the more likely it is that the student will not be included. The categories of the severity level are given in table 7. The major criticism of the severity measure is that it is not subject specific; the impact of the disability might be more severe for one subject than another. A second criticism of the severity measure is that it may be subjective to the SD Background Questionnaire respondent.

A subject-specific measure of severity, which also provides a more objective assessment of how the school treats a student, is the indicator for the grade level of instruction the student receives in the subject tested: *Is the content the same as that given to students in the same grade? Is it one grade level of instruction below, or is it two or more grades below?* A small number of students were reported as *not receiving* instruction in the subject being tested, mathematics or reading. Students not receiving instruction or with this information missing were collapsed into a fourth category because there were too few cases to analyze them

17 The treatment of missing or "not reported" data is a common problem in empirical research. The inclusion of a separate categorical variable for observations with "not reported"/missing as well as practicing listwise deletion of incomplete observations have both been shown to induce bias in estimated coefficients (Vach and Blettner 1991; Jones 1996). Imputation procedures that reduce bias exist. However, our focus in this report is not on the estimated coefficients or their interpretation but in using the coefficients as benchmarks for controlling for shifts in demographics of states' SD populations over time. We included a separate indicator variable for missing categorical data in order to use all observations in the calculation of state-level inclusion rates and avoid complications from employing imputation procedures for completing observations.

separately. The response levels to this question are given in table 7.[18] Although the use of this variable addresses criticism of the severity measure, it is not without its own criticism. The major criticism is that under the *Individuals with Disabilities Education Act* (IDEA) and the *No Child Left Behind Act*, students with disabilities are to be given instruction and goals based on grade-level standards.[19] An answer that the student is receiving instruction at any level other than *at or above grade level* could be interpreted as indicating that the school is breaking the law. This calls into question the validity of responses to this item on the SD Background Questionnaire. Nevertheless, we use this variable, understanding that it is possibly measured with error. The concern over this variable suggests that the items in the NAEP SD Questionnaire be brought more into alignment with the language of IDEA.

The SD Background Questionnaire collects details about accommodations the student received on the state assessments. For our analysis, we use this information to construct a variable that indicates whether the student received an accommodation on the state assessment that is not allowed on NAEP.[20] For 2003 and 2005 data, we match the list of accommodations allowed on NAEP with the list of accommodations reported on the questionnaire for state assessments to identify which students received an accommodation on the state assessment that is not allowed on NAEP. In 2007, an additional item was added to the Background Questionnaire that explicitly asked whether the student received any accommodation on the state assessment that is not allowed on NAEP.[21] For 2007 data, we use this item to construct our variable.

For mathematics, the accommodations not allowed on NAEP include using a tape recorder to submit answers, using a calculator, and testing over several days. For reading, the accommodations not allowed on NAEP include those for mathematics plus having reading passages and test questions read aloud and having test questions asked in sign language. A further discussion of the role of this information in our analysis was provided above.

Standard errors for all results are calculated using a modification of NAEP's recommended procedure for calculating standard errors with the provided jackknife weights. NAEP's recommended procedure needed to be modified because that procedure is meant to provide standard errors for statistics estimated using data from a single NAEP administration. In both approaches described above, regression coefficients estimated using one NAEP administration are applied to a second NAEP administration. Hence, there are two potential sources of error: one from the coefficients and one from the data to which the coefficients are applied. The recommended NAEP procedure for calculating standard errors was modified to take both of these sources of error into account. The procedure we use is described further in appendix A.

SUMMARY STATISTICS

To demonstrate the variation in factors used in analysis, tables 8a, 8b, and 8c provide, for each state, the (weighted) percentage of sampled SDs in each type of disability, the percentage included on the NAEP assessment, the percentages in each severity level, the percentages in each grade level of instruction, and the percentage receiving an accommodation on the state assessment that is not allowed on NAEP. Additionally, the number of SDs sampled in each state is provided in the first column of 8a. Using Alabama in table 8a as an example, we see that for the 2007 NAEP grade 4 mathematics assessment, approximately 390 of the students sampled from that state were classified as an SD who is not

18 In previous versions of our analysis, we used curriculum level, whether the student received the same curriculum as nondisabled students, because it asked explicitly whether the student was given special treatment because of a disability. However, in 2007, this item was discontinued, but the item about grade level of instruction continued.

19 Individuals with Disabilities Education Act Legislation, retrieved 6/24/2008 from http://www.ed.gov/policy/speced/leg/edpicks.jhtml?src=ln.

20 For summary information on the accommodations practices in each state see the profiles of state assessment standards at http://nces.ed.gov/nationsreportcard/researchcenter/statemapping.asp

21 See questionnaires in appendix D.

an ELL. For the non-ELL SD population represented by those 390 students, 88.3 percent participated on NAEP, 52.8 percent had *a specific learning disability,* 23.9 percent had *a speech impairment*, and so on. From table 8b, 12.2 percent had their disabilities classified as *severe* and 45.3 percent had their disabilities classified as *mild*. Finally, table 8c reports in the second column that 45.3 percent of the non-ELL SDs in Alabama received instruction in mathematics at or above grade level. The last column reports that 6.1 percent of the same group received an accommodation on the state assessment that was not allowed on NAEP.

Table 8a. Sample size of public school students with disabilities in NAEP grade 4 mathematics assessment, percentage included on the assessment, and percentage with each disability type: By state, 2007

State	Sample size[1]	Included students	Learning disability	Speech impairment	Mental retardation	Emotional disturbance	Other disabilities
Alabama	390	88.3	52.8	23.9	7.9	1.2	19.3
Alaska	420	91.3	49.4	26.8	5.0	3.5	27.2
Arizona	340	83.4	51.1	34.2	4.4	4.0	27.1
Arkansas	360	79.8	38.8	29.1	11.7	1.4	31.6
California	620	83.9	42.5	33.0	3.6	3.4	29.2
Colorado	360	88.2	51.0	26.0	5.5	9.8	30.5
Connecticut	370	89.7	36.9	20.9	4.1	2.8	42.7
Delaware	550	73.1	60.0	14.9	2.9	3.1	29.6
District of Columbia	260	66.3	63.3	15.0	6.0	6.0	19.5
Florida	790	88.3	56.4	23.8	5.0	5.3	30.2
Georgia	540	83.5	31.6	29.2	6.6	13.3	33.0
Hawaii	340	90.5	59.4	7.2	6.3	4.1	27.3
Idaho	350	86.3	43.9	27.3	7.2	5.9	30.8
Illinois	630	77.6	46.2	29.9	5.3	5.9	29.6
Indiana	520	85.7	39.6	41.0	12.3	5.3	18.8
Iowa	390	90.0	62.1	13.8	2.9	6.9	34.4
Kansas	350	78.9	49.3	27.7	8.2	6.0	28.3
Kentucky	500	84.2	14.7	37.2	11.7	3.2	42.1
Louisiana	530	87.8	40.6	23.7	5.9	2.6	34.5
Maine	530	83.7	39.5	21.5	4.1	4.6	39.5
Maryland	430	72.7	42.6	30.3	3.4	4.7	30.9
Massachusetts	750	74.3	50.3	23.0	1.6	5.6	40.1
Michigan	450	76.1	45.3	33.9	8.7	7.2	20.2
Minnesota	470	85.9	33.7	35.5	8.4	9.9	33.3
Mississippi	350	92.3	43.8	32.9	3.7	2.4	30.2
Missouri	480	76.6	33.6	41.7	9.3	3.6	26.4
Montana	370	81.7	54.6	32.8	5.9	5.4	28.0
Nebraska	470	85.7	42.2	48.0	9.7	2.1	24.5
Nevada	420	84.5	37.3	24.1	2.9	3.8	35.3
New Hampshire	620	88.7	49.6	23.0	2.7	5.3	40.6
New Jersey	460	88.0	49.8	12.3	2.0	4.0	40.0
New Mexico	320	83.2	53.2	31.3	5.4	6.6	27.1
New York	580	91.2	42.1	24.4	2.2	2.7	38.3
North Carolina	760	89.7	39.5	18.9	6.9	3.3	37.0
North Dakota	420	76.3	45.3	36.0	7.8	6.4	24.5
Ohio	640	71.5	50.5	23.5	9.3	4.4	29.0
Oklahoma	470	67.8	50.8	23.2	8.2	5.5	25.1
Oregon	500	85.8	48.1	33.3	4.5	3.7	28.4
Pennsylvania	560	86.1	54.1	30.4	7.2	4.7	24.8
Rhode Island	570	91.0	48.9	29.8	3.8	5.2	34.1
South Carolina	470	88.1	53.4	37.7	9.0	1.5	16.4
South Dakota	460	92.2	43.9	34.0	6.2	3.8	25.3
Tennessee	460	59.4	40.5	33.4	8.6	2.3	33.8
Texas	960	62.8	49.2	29.8	2.5	7.6	38.1
Utah	390	84.4	50.6	29.9	3.1	3.2	23.2
Vermont	440	86.4	46.7	16.3	4.4	12.5	39.8
Virginia	530	74.1	39.9	30.1	5.2	4.3	33.0
Washington	510	85.7	46.0	25.8	5.9	4.5	30.4
West Virginia	540	91.6	31.9	40.3	13.2	1.3	23.9
Wisconsin	450	85.5	43.9	36.1	5.4	11.6	23.1
Wyoming	390	89.6	48.0	36.2	3.6	5.2	29.7

[1] Rounded to the nearest 10 for confidentiality.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 2007 Mathematics Assessment.

Table 8b. Percentages of public school students with disabilities in NAEP grade 4 mathematics assessment with each severity level of disabilities: By state, 2007

State	Severe disability	Moderate disability	Mild disability	Severity not reported
Alabama	12.2	35.4	45.3	7.2
Alaska	9.7	35.0	42.5	12.7
Arizona	4.4	24.8	48.2	22.6
Arkansas	10.5	38.3	43.8	7.5
California	9.3	28.3	49.3	13.1
Colorado	10.5	42.8	36.7	10.1
Connecticut	8.2	38.0	34.7	19.2
Delaware	7.7	37.4	48.7	6.2
District of Columbia	9.1	49.9	31.5	9.5
Florida	6.9	39.1	45.5	8.5
Georgia	3.6	22.4	59.2	14.8
Hawaii	7.1	34.5	51.1	7.3
Idaho	7.8	40.2	44.2	7.8
Illinois	3.7	32.2	50.0	14.1
Indiana	3.3	14.6	72.5	9.7
Iowa	4.1	25.9	63.5	6.5
Kansas	12.7	39.2	42.0	6.0
Kentucky	7.0	30.5	54.4	8.2
Louisiana	2.5	12.3	64.7	20.5
Maine	7.6	41.2	41.4	9.7
Maryland	7.0	38.8	40.6	13.6
Massachusetts	5.9	57.6	28.0	8.4
Michigan	10.9	38.1	46.3	4.7
Minnesota	12.2	31.9	51.2	4.7
Mississippi	3.3	23.9	53.3	19.4
Missouri	9.1	39.0	49.4	2.5
Montana	13.2	41.6	37.2	8.0
Nebraska	2.5	18.4	74.5	4.6
Nevada	7.0	24.2	42.6	26.2
New Hampshire	11.1	51.5	28.8	8.5
New Jersey	2.9	45.9	42.2	8.9
New Mexico	14.7	41.0	33.9	10.4
New York	8.0	41.7	37.7	12.7
North Carolina	6.5	36.9	49.2	7.5
North Dakota	9.1	39.1	45.8	6.0
Ohio	9.4	34.7	46.5	9.4
Oklahoma	5.7	28.3	61.1	5.0
Oregon	7.2	43.0	43.8	6.0
Pennsylvania	5.1	34.7	52.2	8.0
Rhode Island	2.3	36.5	49.9	11.3
South Carolina	5.5	27.4	59.6	7.5
South Dakota	4.1	19.7	65.6	10.6
Tennessee	10.5	34.8	45.4	9.3
Texas	10.0	41.6	40.4	7.9
Utah	9.3	35.5	47.6	7.6
Vermont	15.0	48.5	25.9	10.6
Virginia	9.2	35.5	48.7	6.6
Washington	5.4	36.0	48.9	9.7
West Virginia	2.6	18.0	70.2	9.2
Wisconsin	7.5	40.8	47.1	4.6
Wyoming	11.4	41.6	40.3	6.8

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 2007 Mathematics Assessment.

Table 8c. Percentages of public school students with disabilities in NAEP grade 4 mathematics assessment at each grade level of instruction, and percentage who received a non-NAEP accommodation: By state, 2007

State	Same or above grade	One year below	Two years or more below	Not reported	Non-NAEP accommodation
Alabama	45.3	21.2	25.8	7.7	6.1
Alaska	36.1	22.6	25.5	15.8	10.0
Arizona	38.6	16.1	25.5	19.8	11.2
Arkansas	41.2	13.3	35.0	10.6	16.4
California	51.4	13.2	21.0	14.4	10.8
Colorado	37.3	24.8	28.1	9.7	7.6
Connecticut	43.1	18.0	20.8	18.1	2.0
Delaware	48.0	25.6	19.5	6.9	47.5
District of Columbia	18.4	22.7	44.3	14.6	28.2
Florida	52.5	22.0	16.0	9.5	5.2
Georgia	51.5	20.9	17.9	9.6	19.3
Hawaii	34.1	23.1	34.9	7.8	8.0
Idaho	45.9	18.4	25.8	9.9	10.6
Illinois	46.2	19.4	19.4	15.0	13.6
Indiana	62.8	15.3	15.3	6.6	18.5
Iowa	44.7	26.9	19.7	8.7	16.5
Kansas	44.5	26.2	20.6	8.7	12.0
Kentucky	57.8	15.4	16.1	10.8	15.6
Louisiana	55.9	14.8	20.5	8.8	30.5
Maine	43.6	22.4	21.0	13.0	16.1
Maryland	49.1	20.2	19.4	11.3	23.0
Massachusetts	52.7	19.9	15.7	11.8	31.2
Michigan	42.2	21.9	29.4	6.4	11.8
Minnesota	45.2	20.1	26.2	8.5	2.4
Mississippi	61.5	10.0	7.8	20.7	5.5
Missouri	53.5	16.6	24.0	5.9	22.5
Montana	37.6	20.0	28.6	13.7	19.4
Nebraska	59.5	16.4	19.8	4.2	17.3
Nevada	35.9	17.5	20.4	26.2	8.7
New Hampshire	46.4	24.9	18.5	10.2	7.2
New Jersey	46.9	27.0	16.1	10.0	21.5
New Mexico	37.1	24.2	26.5	12.1	5.5
New York	48.9	21.3	17.8	12.1	9.5
North Carolina	54.0	17.8	19.3	8.9	4.2
North Dakota	55.2	15.0	18.9	10.8	17.7
Ohio	44.0	18.0	28.3	9.7	30.5
Oklahoma	40.4	22.5	28.5	8.7	20.6
Oregon	39.3	25.0	25.0	10.8	20.9
Pennsylvania	45.1	21.0	23.4	10.5	18.8
Rhode Island	48.4	22.3	18.5	10.8	8.5
South Carolina	59.9	12.2	18.7	9.2	8.4
South Dakota	54.2	18.4	15.2	12.3	11.0
Tennessee	36.1	20.6	30.1	13.2	42.3
Texas	52.7	18.2	20.9	8.1	7.3
Utah	39.6	18.6	27.4	14.5	13.1
Vermont	35.1	26.4	27.2	11.3	8.8
Virginia	62.1	15.8	13.2	8.9	29.1
Washington	37.5	21.1	30.7	10.7	9.9
West Virginia	63.9	10.0	16.4	9.7	6.0
Wisconsin	47.8	20.1	21.5	10.6	15.0
Wyoming	41.1	24.3	24.5	10.1	11.8

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 2007 Mathematics Assessment.

CAVEATS AND CAUTIONS IN INTERPRETATION

SUBJECTIVITY AND MEASUREMENT ERROR

Some degree of subjectivity exists in the variables providing information on a student's disability characteristics. For example, the SD Background Questionnaire respondents who classify the students may have different interpretations of the disability classifications or of how to code the severity level of a student's disability. Reschly (1996) analyzes the subjective nature of these widely used systems of classifying SDs. If the subjective interpretation of a control variable is random across all observations, it is akin to measurement error. The statistical consequence of measurement error in a control variable is that the coefficient estimated for that variable in a regression model will tend toward zero. In the extreme case, if all our control variables were pure error, we would not be able to detect any differences among states in terms of those variables. In this extreme case, the measure of change over time would simply be the actual observed change over time because none of that change would be attributed to changes in control variables. Similarly, the starting point measure, which measures differences between states in the initial period, would simply be the difference between a state's actual inclusion rate and the average overall inclusion rate because none of the variation would be explained by control variables.

In our analysis, we cannot know how much our variables are measured with error. To the extent that a control variable is measured with error, its ability to explain differences in inclusion rates is reduced. Because the measure of change captures the portion of change that is not explained by the control variables, as the ability of the control variables to explain differences in inclusion rates is reduced, the magnitude of the measure of change will rise. For example, suppose that State A has an inclusion rate of 85 percent in 2005 and an inclusion rate of 92 percent in 2007. The actual change in inclusion rate for State A is 7 percentage points. If our control variables were perfect, they might, say, explain 4 of those 7 percentage points and the change measure would be 3 percentage points. If however, the control variables are measured with error, they will explain less of the change in inclusion rates. Hence, the control variables measured with error might explain 2 of the 7 percentage points, resulting in a change measure of 5 percentage points. Similar logic can be applied to the starting point measure, which captures unexplained differences in inclusion rates among states.

If the subjective interpretation of a control variable is not completely random but, to some extent, differs systematically and is correlated with some observable or non-observable characteristic, bias will occur in the estimated coefficients. In our analysis, the potentially subjective variables, type of disability and severity level, are control variables and are not variables of interest. What is of interest are the state-level predictions we obtain from applying the model to data. The subjectivity, therefore, will be of concern if it is correlated somehow with states or a state-level characteristic. For example, we would be concerned if we saw systematic differences in the definition of autism across states. Such a systematic difference will cause bias in our estimates of change.

The bias is a concern in the regression models where the model is estimated using data from all states at the same time: the nation-based measure of change, the nation-based starting point measure, and the state-specific starting point measure. The bias from systematic subjectivity is not a concern in the state-specific approach for measuring change because here the regression model is estimated separately for each state. Subjectivity within the state will still cause measurement error, as discussed above, but the bias in calculating state-level statistics will be removed. For the state-specific approach's change measure, however, it will be a concern if the subjective interpretation of a variable is thought to change over time within a state.

An additional source of error that affects models for measuring the starting point and the nation-based model for measuring change, but not the state-specific model for measuring change, is small differences between the 2005 and 2007 SD questionnaires. As discussed above, the 2007 SD questionnaire includes a question that asks directly whether the student receives an accommodation on the state assessment that is not allowed on NAEP, which enables a more accurate gauge of who receives such an accommodation. The 2005 questionnaire, on the other hand, has a number of "other accommodation" items that create ambiguity. The effect is that the coefficient on "received an accommodation on the state assessment not allowed on NAEP" will have some error due to ambiguity as to whether all the appropriate students were included in estimating this effect. This will result in an added element of unreliability in the estimated measures that we cannot predict. As stated above, the state-specific measure of change is, however, not affected by this problem as the variable does not enter that model.

CHANGES IN IDENTIFICATION RATES OF STUDENTS WITH DISABILITIES

In this report, we do not control explicitly for changes in identification rates of students with disabilities but random changes will automatically be accounted for by our control variables. In our method, the inclusion rates of SDs in the initial period set expectations for the inclusion of SDs in the second period. The assumption is that the group of students with a given set of characteristics in the first period is not different from the group of students with the same characteristics in the second period. As long as this assumption holds, we can apply the expectations set by the initial period whether the proportion of students who are SDs in a state changes or not. If, however, changes in the proportion of students who are SD in a state are due to changes in identification policy, this can cause some inaccuracies in our method.

Say, for example, that a state decides that a particular group of students who in the initial period would have not been identified as SD, and hence are all included, are in the second period identified as students with disabilities, and that they all have the following characteristics: they have a specific learning disability; the disability is classified as mild; they are receiving *at or above* grade-level instruction in the subject being assessed; they did not receive an accommodation on the state assessment that was not allowed on NAEP. According to our method, these students are expected to be treated as other students with those characteristics. If they are instead more mildly disabled than other students with those characteristics and hence, for example, all included, our partitioning methods might conclude that this state was improving in its inclusion of SDs. This conclusion would be based on the fact that students of this type are included more often. The underlying reason, however, is not that the state is being more inclusive, but rather that they increased their identification rate.

Such a shift in policy might cause a jump in the change measure. But, it will also raise the expectation of inclusion for that state for measuring change over the following period. If the state were to reverse this policy in the third period, then the state would see a jump in the opposite direction from period 2 to period 3. This discussion applies equally to the nation-based and state-specific methods.

CAUTION AGAINST NORMATIVE INTERPRETATIONS

The purpose of this report is to develop measures of changes in inclusion rates. As such, we develop expected or predicted inclusion rates. It is important to emphasize that predicted inclusion rates are not to be interpreted as normative. At the same time in this report, positive change in inclusion rates is interpreted as a desirable result, in alignment with the intent and language of the NAEP legislation (National Assessment of Educational Progress Authorization Act of 2002), which requires NAEP to report scores for SD and ELL populations, and the NCLB act.

NCLB regulations require states to conduct academic assessments for accountability purposes and aspire to the ideal of having all students participate. A Federal Register summary of the regulations (U.S. Department of Education 2002) describes it thus:

> *The final IDEA regulations that are included in these regulations provide that a State's (or in the case of district-wide assessments, an LEA's) guidelines require each child to be validly assessed and identify, for each assessment, any accommodations that would result in an invalid score. Consistent with Title I, a student with disabilities must receive a valid score in order to be counted as a participant under the IDEA.*

These regulations apply to state assessments that are used for accountability purposes and do not apply directly to NAEP. Nevertheless, we provide this as justification that higher participation rates on NAEP can generally be thought of as a *good.* We do not claim, however, that the ideal participation rate is 100 percent. Given that NAEP's construct is fixed, certain accommodations would invalidate the NAEP score. Hence, if a student required such an accommodation to be assessed, he or she would need to be excluded from NAEP.

Although in this report we generally refer to higher inclusion rates on NAEP as better, we make no claim about which students should or should not be assessed on NAEP. For measuring change, a comparison point needs to be set. Although it is tempting to interpret such comparison points as norms for inclusion, they should not be interpreted as such. We can set such points and still not interpret them as normative.

As an example, consider a completely naïve approach to measuring change in inclusion rates. The approach would say that a state's inclusion rate in one period is what we expect it to be in the next period. This would be a legitimate, though naïve, way of measuring change, and it places no normative interpretation on the benchmark, the state's initial inclusion rate, that is set for measuring change. Similarly, in our approach that sets the benchmarks to hold differences in the distribution of SD characteristics constant, the benchmarks are set to measure change and should not be interpreted normatively as target, correct, or desirable inclusion rates.

Although our benchmarks for measuring change should not be normatively interpreted, the measure of change itself is in this report. Under NCLB regulations, all students, including SDs, should be assessed by states. This creates a norm that is here applied to NAEP. In the naïve example given above, our crude measure of change could call an increase in inclusion rates good. The more complex approach we take here is to factor out changes in the distribution of SD characteristics when deciding whether the change is really good or not. Hence, though the benchmarks do not themselves provide normative content, the change measures that they allow do.

RESULTS

NATION-BASED RESULTS

Results for the nation-based approach are provided in tables 9 through 12 for grades 4 and 8, mathematics and reading, for changes between the 2005 school year and the 2007 school year.[22,23] Fit statistics indicate that the regressions fit the data adequately. Logit pseudo R^2 statistics were .35 for mathematics grade 4, .30 for mathematics grade 8, .33 for reading grade 4, and .30 for reading grade 8.[24] Likelihood ratio tests were all statistically significant at the .01 level.

The first row in table 9 (first two columns) shows that Alabama had an 89.6 percent actual inclusion rate for 2005 and an 88.3 percent for 2007. According to the nation-based model results in columns 3 and 4, Alabama had state-level predicted inclusion rates of 79.7 percent in 2005 and 83.0 percent in 2007. This predicted inclusion rate is based on the proportions of students with different types and severities of disabilities and accommodations offered by the state for its own state assessment tests that are not allowed on NAEP. The resulting differences between actual and predicted inclusion rates are 9.9 and 5.4 in 2005 and 2007, respectively. The change between 2005 and 2007 was thus measured to be –4.6 percentage points, as shown in column 7. In other words, of students with the same disability profile, 4.6 percent fewer were included in Alabama in 2007 than in 2005. This change was not statistically significant. As shown in the last column, Alabama's 2005 starting point measure, the difference from predicted for 2005, placed it in the top quartile (i.e., the 4th quartile), with a change score of 0, for a composite index score of (4,0).

Tables 9 through 12 also show wide variation in state-level predicted inclusion rates (columns 3 and 4). These rates represent the inclusion rate we would expect the state to have because of the characteristics of the SDs in the state and the availability of an accommodation on the state assessment that is not allowed on NAEP. These predicted inclusion rates illustrate that on the sole basis of different distribution of SD characteristics across states, we expect inclusion rates on the grade 4 NAEP mathematics assessment to range from 71.6 (District of Columbia) to 87.1 (Mississippi) in 2005 and from 73.3 (Tennessee) to 87.4 (Mississippi) in 2007.

As can be seen in tables 9 through 12, most states did not make statistically significant changes from 2005 in their rate of inclusion. In about one-third of the states, there were significant changes in inclusion rates for the mathematics assessments; in about half the states, there were significant changes for the reading assessments.[25] States that were less inclusive of SDs in 2007 than 2005 outnumbered states that were more inclusive:

- For grade 4 mathematics, 7 states were significantly more inclusive in 2007 than in 2005, and 8 states were significantly less inclusive.
- For grade 8 mathematics, 2 states were significantly more inclusive and 17 states were significantly less inclusive.
- For grade 4 reading, 8 states were significantly more inclusive, and 18 were significantly less inclusive.

22 Results for changes from 2003 to 2005 using the nation-based method are in appendix tables B-1 through B-4.

23 Tables 13 through 16 present results for the state-specific approach.

24 Pseudo R^2 statistics reported for logistic regressions are approximations to OLS R^2 statistics but are not the same and a number of different approximations exist. Here we report Nagelkerke's (1991) re-centered pseudo R^2.

25 Tables 19 through 22 show which states were in each starting point quartile and whether or not they had positive, negative, or no significant change from 2005 to 2007.

- For grade 8 reading, 4 states were significantly more inclusive, and 21 states were significantly less inclusive.

Tables 17 and 18 present the counts of states with positive change, negative change, and no significant change by the quartile of their initial inclusion rate. Across each grade and subject in the nation-based results, states that were more inclusive in 2007—in other words, that had significant positive change—were in the lowest two quartiles of the starting point measure, meaning that they had lower initial inclusion rates than half the states. For states that became significantly less inclusive in 2007, in mathematics there was no obvious association with initial inclusion rates. In reading, however, more states were prevalent in the upper quartiles of the starting point measure. Tables 19 through 22 show which states were in each quartile of initial inclusiveness and whether or not they had positive, negative, or no significant change from 2005 to 2007.

Table 9. Actual and predicted percentages of students with disabilities in NAEP grade 4 mathematics assessments, starting point and change measures, using nation-based approach: By state, 2005 and 2007

State	Actual rates		Predicted rates		Diff. from predicted[1]				
	2005 (a)	2007 (b)	2005 (c)	2007 (d)	2005 (e) (a–c)	2007 (f) (b–d)	Change 2005–07 (f–e)	Standard error	Composite Index[2]
Alabama	89.6	88.3	79.7	83.0	9.9	5.4	–4.6	2.83	(4,0)
Alaska	93.7	91.3	84.1	81.5	9.7	9.8	0.2	1.83	(4,0)
Arizona	80.3	83.4	82.7	82.0	–2.4	1.4	3.8	3.28	(1,0)
Arkansas	86.3	79.8	78.1	77.8	8.2	2.0	–6.2 *	2.81	(4,–1)
California	81.0	83.9	83.6	83.1	–2.6	0.8	3.3	2.55	(1,0)
Colorado	83.7	88.2	79.2	79.9	4.5	8.3	3.8	2.77	(3,0)
Connecticut	87.9	89.7	86.0	82.5	1.9	7.2	5.2 *	2.24	(2,1)
Delaware	58.4	73.1	80.1	78.2	–21.8	–5.1	16.7 *	2.65	(1,1)
District of Columbia	67.7	66.3	71.6	75.1	–3.9	–8.8	–4.9	3.17	(1,0)
Florida	88.6	88.3	83.9	85.3	4.6	3.0	–1.6	2.66	(3,0)
Georgia	87.9	83.5	84.1	83.5	3.8	0.0	–3.8	2.53	(3,0)
Hawaii	85.2	90.5	79.5	81.0	5.7	9.5	3.9	2.58	(4,0)
Idaho	92.2	86.3	80.5	80.7	11.7	5.7	–6.0 *	2.29	(4,–1)
Illinois	86.5	77.6	81.9	83.1	4.6	–5.5	–10.2 *	2.77	(3,–1)
Indiana	91.7	85.7	86.1	85.1	5.6	0.6	–5.0	2.64	(4,0)
Iowa	88.3	90.0	85.0	82.9	3.3	7.1	3.8	2.60	(3,0)
Kansas	85.7	78.9	80.3	79.1	5.5	–0.2	–5.6	2.99	(4,0)
Kentucky	83.7	84.2	80.4	83.6	3.3	0.6	–2.7	2.90	(2,0)
Louisiana	83.8	87.8	84.7	84.1	–0.9	3.7	4.6	2.91	(2,0)
Maine	83.0	83.7	81.0	81.2	2.0	2.5	0.5	2.66	(2,0)
Maryland	79.4	72.7	81.7	81.3	–2.3	–8.6	–6.3 *	3.10	(1,–1)
Massachusetts	83.9	74.3	83.9	80.7	0.0	–6.3	–6.3	3.45	(2,0)
Michigan	74.6	76.1	78.4	81.2	–3.8	–5.1	–1.3	3.31	(1,0)
Minnesota	86.3	85.9	82.4	80.9	4.0	5.0	1.0	2.93	(3,0)
Mississippi	80.4	92.3	87.1	87.4	–6.6	4.9	11.5 *	2.95	(1,1)
Missouri	87.2	76.6	82.7	81.4	4.5	–4.8	–9.2 *	3.17	(3,–1)
Montana	83.4	81.7	80.6	77.3	2.7	4.4	1.7	3.17	(2,0)
Nebraska	88.7	85.7	85.1	85.2	3.6	0.5	–3.2	2.59	(3,0)
Nevada	80.7	84.5	83.3	82.3	–2.6	2.1	4.7	3.47	(1,0)
New Hampshire	90.3	88.7	82.0	82.3	8.3	6.4	–1.9	2.15	(4,0)
New Jersey	87.7	88.0	83.9	82.5	3.8	5.5	1.7	3.21	(3,0)
New Mexico	89.3	83.2	81.9	80.2	7.4	3.0	–4.4	3.55	(4,0)
New York	83.7	91.2	83.2	85.2	0.5	6.0	5.5 *	2.56	(2,1)
North Carolina	87.5	89.7	83.7	84.6	3.8	5.1	1.3	1.87	(3,0)
North Dakota	85.5	76.3	84.9	81.9	0.6	–5.6	-6.2 *	2.32	(2,–1)
Ohio	73.0	71.5	78.9	77.3	–5.9	–5.8	0.1	4.43	(1,0)
Oklahoma	78.8	67.8	78.5	79.9	0.3	–12.1	–12.5 *	3.27	(2,–1)
Oregon	78.7	85.8	80.0	80.3	–1.3	5.5	6.8 *	2.91	(2,1)
Pennsylvania	85.3	86.1	80.5	81.5	4.8	4.6	–0.1	3.07	(3,0)
Rhode Island	87.9	91.0	84.6	84.0	3.3	7.0	3.7	2.45	(2,0)
South Carolina	73.8	88.1	83.4	85.7	–9.6	2.4	12.1 *	2.56	(1,1)
South Dakota	91.0	92.2	86.1	85.2	4.9	7.0	2.1	1.51	(3,0)
Tennessee	76.1	59.4	73.3	73.3	2.8	–13.9	–16.7 *	4.32	(2,–1)
Texas	65.2	62.8	82.5	82.8	–17.3	–19.9	–2.6	3.04	(1,0)
Utah	88.6	84.4	81.8	82.0	6.8	2.4	–4.4	2.56	(4,0)
Vermont	80.4	86.4	79.1	78.2	1.3	8.2	6.9 *	2.36	(2,1)
Virginia	71.6	74.1	82.4	82.8	–10.8	–8.7	2.1	3.33	(1,0)
Washington	85.4	85.7	78.1	80.5	7.3	5.3	–2.0	2.58	(4,0)
West Virginia	88.5	91.6	84.3	87.4	4.1	4.2	0.1	1.83	(3,0)
Wisconsin	88.3	85.5	81.3	82.7	7.0	2.8	–4.2	2.88	(4,0)
Wyoming	91.8	89.6	83.2	81.1	8.6	8.6	#	2.02	(4,0)

* 2005–2007 change over time is significantly different from zero ($p < .05$).

Estimate rounds to zero.

[1] The 2005 *difference from predicted* is also the *starting point* measure.

[2] The *composite index* (*q*,*s*) is the quartile of the *starting point, q* (from 1, the lowest, to 4, the highest), and statistical significance of the *change* score, *s*, where *s* is –1 if the change is negative and statistically significant, 1 if the change is positive and statistically significant, and 0 if changes are not statistically different from zero ($p < .05$).

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 2005 and 2007 Mathematics Assessments.

Table 10. Actual and predicted percentages of students with disabilities in NAEP grade 8 mathematics assessments, starting point and change measures, using nation-based approach: By state, 2005 and 2007

State	Actual rates		Predicted rates		Diff. from predicted[1]				
	2005 (a)	2007 (b)	2005 (c)	2007 (d)	2005 (e) (a–c)	2007 (f) (b–d)	Change 2005–07 (f–e)	Standard error	Composite Index[2]
Alabama	92.3	76.8	78.9	76.7	13.4	0.1	–13.3 *	4.01	(4,–1)
Alaska	84.2	63.1	78.0	72.3	6.2	–9.2	–15.4 *	3.41	(3,–1)
Arizona	71.4	75.3	77.5	78.2	–6.1	–2.9	3.2	5.33	(1,0)
Arkansas	80.2	81.7	74.0	76.1	6.3	5.6	–0.7	3.39	(3,0)
California	84.4	83.0	78.6	81.4	5.9	1.6	–4.3	2.54	(3,0)
Colorado	83.9	87.2	77.6	77.6	6.3	9.6	3.3	2.75	(3,0)
Connecticut	83.7	90.7	80.9	81.3	2.9	9.4	6.6 *	2.09	(2,1)
Delaware	34.1	57.1	71.2	72.2	–37.1	–15.1	22.0 *	3.25	(1,1)
District of Columbia	71.2	45.7	70.5	69.9	0.7	–24.2	–24.9 *	3.48	(2,–1)
Florida	85.6	83.5	78.6	80.8	7.0	2.7	–4.3	2.80	(4,0)
Georgia	82.0	50.2	79.5	79.6	2.5	–29.4	–31.9 *	3.78	(2,–1)
Hawaii	86.1	90.8	79.2	79.5	6.9	11.3	4.5	2.29	(3,0)
Idaho	86.3	86.6	76.2	76.8	10.1	9.7	–0.3	2.59	(4,0)
Illinois	83.7	65.6	76.7	74.3	6.9	–8.7	–15.7 *	4.10	(3,–1)
Indiana	76.5	63.8	77.3	77.5	–0.8	–13.7	–12.9 *	4.46	(2,–1)
Iowa	83.8	84.2	76.7	80.6	7.1	3.6	–3.5	3.33	(4,0)
Kansas	76.6	69.5	73.4	72.1	3.3	–2.6	–5.9	4.04	(2,0)
Kentucky	72.7	51.5	76.0	72.7	–3.3	–21.1	–17.9 *	4.41	(1,–1)
Louisiana	70.7	74.1	75.1	74.9	–4.4	–0.8	3.6	5.43	(1,0)
Maine	75.2	71.9	73.7	77.0	1.4	–5.1	–6.5	3.43	(2,0)
Maryland	67.2	38.3	74.1	73.2	–6.9	–34.9	–27.9 *	5.04	(1,–1)
Massachusetts	68.6	49.4	76.3	73.2	–7.7	–23.8	–16.0 *	4.40	(1,–1)
Michigan	69.4	69.1	75.3	74.1	–5.9	–4.9	1.0	4.25	(1,0)
Minnesota	85.5	83.4	76.4	76.3	9.1	7.1	–1.9	2.78	(4,0)
Mississippi	68.6	78.5	80.8	83.0	–12.3	–4.5	7.8	5.09	(1,0)
Missouri	72.5	65.4	73.0	73.2	–0.5	–7.8	–7.3	3.99	(2,0)
Montana	84.2	77.3	77.8	74.8	6.5	2.5	–4.0	3.19	(3,0)
Nebraska	91.1	83.0	81.3	81.9	9.8	1.0	–8.8 *	2.68	(4,–1)
Nevada	82.4	73.8	80.6	76.9	1.7	–3.1	–4.9	3.51	(2,0)
New Hampshire	87.9	83.3	78.4	80.0	9.6	3.3	–6.2 *	2.43	(4,–1)
New Jersey	83.1	82.7	79.5	78.5	3.6	4.2	0.6	4.01	(3,0)
New Mexico	87.1	83.5	78.2	76.3	8.9	7.2	–1.7	2.63	(4,0)
New York	81.2	79.3	79.9	80.9	1.3	–1.7	–3.0	3.28	(2,0)
North Carolina	85.8	86.7	79.5	82.8	6.3	3.9	–2.4	2.39	(3,0)
North Dakota	74.0	58.1	75.2	75.3	–1.2	–17.1	–16.0 *	3.12	(2,–1)
Ohio	59.9	53.6	70.8	67.6	–10.9	–14.0	–3.0	5.14	(1,0)
Oklahoma	76.5	44.9	73.4	73.7	3.1	–28.8	–31.9 *	4.46	(2,–1)
Oregon	82.1	78.0	73.6	76.4	8.5	1.6	–6.9 *	3.08	(4,–1)
Pennsylvania	80.8	77.9	75.9	78.0	4.9	–0.1	–5.0	4.15	(3,0)
Rhode Island	85.1	88.0	81.0	82.4	4.1	5.6	1.5	2.21	(3,0)
South Carolina	59.2	60.2	78.8	79.6	–19.5	–19.4	0.1	4.73	(1,0)
South Dakota	82.9	78.3	75.7	75.4	7.2	2.9	–4.3	2.47	(4,0)
Tennessee	68.7	47.0	71.3	71.0	–2.6	–24.0	–21.4 *	4.70	(2,–1)
Texas	61.0	58.5	78.3	78.6	–17.3	–20.1	–2.8	3.48	(1,0)
Utah	82.4	77.1	74.2	74.6	8.2	2.5	–5.7	3.06	(4,0)
Vermont	79.2	78.0	75.8	74.3	3.3	3.7	0.4	2.41	(3,0)
Virginia	70.9	58.9	79.5	78.5	–8.6	–19.6	–11.0 *	4.14	(1,–1)
Washington	83.0	73.1	74.3	74.0	8.7	–0.9	–9.6 *	3.88	(4,–1)
West Virginia	83.0	88.9	80.0	83.0	3.0	5.9	2.9	2.20	(2,0)
Wisconsin	78.7	73.6	72.1	74.2	6.6	–0.5	–7.1	3.66	(3,0)
Wyoming	89.4	84.7	78.8	79.4	10.6	5.3	–5.2	2.77	(4,0)

* 2005–2007 change over time significantly different from zero ($p < .05$).

[1] The 2005 difference from predicted is also the starting point measure.

[2] The *composite index* (*q*,*s*) is the quartile of the *starting point, q* (from 1, the lowest, to 4, the highest), and statistical significance of the *change* score, *s*, where *s* is –1 if the change is negative and statistically significant, 1 if positive and statistically significant, and 0 if changes are not statistically different from zero ($p < .05$).

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 2005 and 2007 Mathematics Assessments.

Table 11. Actual and predicted percentages of students with disabilities in NAEP grade 4 reading assessments, starting point and change measures, using nation-based approach: By state, 2005 and 2007

State	Actual rates		Predicted rates		Diff. from predicted[1]				
	2005 (a)	2007 (b)	2005 (c)	2007 (d)	2005 (e) (a–c)	2007 (f) (b–d)	Change 2005–07 (f–e)	Standard error	Composite Index[2]
Alabama	85.3	78.1	65.3	71.6	20.0	6.6	–13.4 *	3.01	(4,–1)
Alaska	84.5	80.9	66.3	69.2	18.2	11.7	–6.5	3.55	(4,0)
Arizona	66.7	74.9	66.2	70.7	0.5	4.1	3.6	4.40	(2,0)
Arkansas	53.3	55.0	60.0	67.2	–6.7	–12.2	–5.5	4.59	(1,0)
California	74.9	77.3	71.8	73.7	3.1	3.6	0.5	3.02	(2,0)
Colorado	78.6	79.9	64.0	69.3	14.7	10.6	–4.0	3.34	(4,0)
Connecticut	78.5	85.1	69.4	72.8	9.0	12.2	3.2	3.30	(3,0)
Delaware	28.6	46.0	59.3	58.5	–30.7	–12.4	18.3 *	2.80	(1,1)
District of Columbia	57.7	28.6	54.1	57.2	3.6	–28.6	–32.2 *	3.66	(3,–1)
Florida	76.0	77.3	71.2	75.7	4.8	1.7	–3.2	3.51	(3,0)
Georgia	59.7	43.4	62.9	65.9	–3.3	–22.5	–19.3 *	4.12	(2,–1)
Hawaii	82.5	77.1	68.5	68.8	14.0	8.4	–5.7	3.66	(4,0)
Idaho	72.2	77.2	64.1	69.5	8.2	7.6	–0.5	4.31	(3,0)
Illinois	65.3	67.5	64.4	70.9	1.0	–3.3	–4.3	4.26	(2,0)
Indiana	75.2	77.6	67.6	72.6	7.6	5.0	–2.6	3.89	(3,0)
Iowa	65.2	71.2	60.2	65.9	5.0	5.3	0.3	5.12	(3,0)
Kansas	77.4	62.4	62.2	70.4	15.3	–8.0	–23.2 *	4.32	(4,–1)
Kentucky	45.4	53.1	58.8	66.3	–13.4	–13.1	0.3	3.26	(1,0)
Louisiana	41.2	79.3	64.9	70.7	–23.7	8.6	32.3 *	4.79	(1,1)
Maine	65.8	69.5	66.0	67.8	–0.2	1.7	1.9	3.51	(2,0)
Maryland	63.3	51.4	67.3	66.4	–3.9	–15.0	–11.1 *	3.55	(2,–1)
Massachusetts	67.2	72.1	67.1	67.5	0.2	4.5	4.3	3.27	(2,0)
Michigan	52.3	68.2	61.9	67.6	–9.7	0.6	10.3 *	4.51	(1,1)
Minnesota	84.5	77.1	70.6	72.9	13.9	4.1	–9.7 *	3.02	(4,–1)
Mississippi	66.9	77.6	74.5	76.6	–7.6	1.0	8.6 *	3.93	(1,1)
Missouri	55.8	78.9	63.2	73.2	–7.4	5.6	13.1 *	3.37	(1,1)
Montana	61.1	64.8	59.0	63.6	2.1	1.1	–1.0	4.42	(2,0)
Nebraska	75.2	72.1	70.5	72.1	4.7	0.0	–4.7	3.63	(3,0)
Nevada	60.4	70.1	64.2	69.2	–3.8	1.0	4.7	4.62	(2,0)
New Hampshire	83.0	79.8	68.4	72.7	14.6	7.0	–7.5 *	3.60	(4,–1)
New Jersey	73.8	61.7	64.7	66.8	9.0	–5.1	–14.2 *	5.22	(3,–1)
New Mexico	64.4	54.2	64.2	69.5	0.2	–15.2	–15.4 *	5.05	(2,–1)
New York	74.5	72.5	68.8	71.1	5.7	1.4	–4.3	4.03	(3,0)
North Carolina	83.2	87.6	69.1	74.0	14.2	13.6	–0.6	2.44	(4,0)
North Dakota	64.7	46.2	67.1	65.7	–2.4	–19.5	–17.1 *	2.92	(2,–1)
Ohio	40.8	50.6	54.4	61.1	–13.5	–10.4	3.1	4.86	(1,0)
Oklahoma	72.5	58.8	63.9	67.6	8.6	–8.9	–17.5 *	4.22	(3,–1)
Oregon	69.4	74.3	67.0	69.2	2.3	5.1	2.7	3.29	(2,0)
Pennsylvania	73.3	72.0	63.5	68.9	9.7	3.1	–6.6	4.09	(4,0)
Rhode Island	88.0	82.9	70.1	72.8	17.9	10.1	–7.8 *	2.61	(4,–1)
South Carolina	61.1	73.7	70.0	72.5	–8.9	1.2	10.2 *	3.54	(1,1)
South Dakota	71.4	66.2	65.1	71.1	6.3	–4.9	–11.2 *	2.72	(3,–1)
Tennessee	38.4	37.0	53.6	58.7	–15.2	–21.7	–6.5	5.30	(1,0)
Texas	58.5	51.7	72.7	71.3	–14.2	–19.6	–5.4	3.25	(1,0)
Utah	72.0	62.8	66.0	68.5	6.0	–5.7	–11.7 *	4.01	(3,–1)
Vermont	68.4	67.0	59.7	64.3	8.6	2.7	–6.0 *	2.96	(3,–1)
Virginia	36.8	55.5	59.7	70.0	–22.9	–14.5	8.4 *	3.85	(1,1)
Washington	76.8	72.5	63.3	68.5	13.5	4.0	–9.5 *	4.16	(4,–1)
West Virginia	69.1	91.0	69.4	74.7	–0.3	16.4	16.7 *	2.77	(2,1)
Wisconsin	71.7	72.4	61.5	70.1	10.2	2.3	–7.9 *	3.95	(4,–1)
Wyoming	90.4	78.6	67.8	68.4	22.7	10.3	–12.4 *	2.22	(4,–1)

* 2005–2007 change over time significantly different from zero ($p < .05$).

[1] The 2005 *difference from predicted* is also the *starting point* measure.

[2] The *composite index* (*q*,*s*) is the quartile of the *starting point, q* (from 1, the lowest, to 4, the highest), and statistical significance of the *change* score, *s*, where *s* is –1 if the change is negative and statistically significant, 1 if positive and statistically significant, and 0 if changes are not statistically different from zero ($p < .05$).

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 2005 and 2007 Reading Assessments.

Table 12. Actual and predicted percentages of students with disabilities in NAEP grade 8 reading assessments, starting point and change measures, using nation-based approach: By state, 2005 and 2007

State	Actual rates		Predicted rates		Diff. from predicted[1]				
	2005 (a)	2007 (b)	2005 (c)	2007 (d)	2005 (e) (a–c)	2007 (f) (b–d)	Change 2005–07 (f–e)	Standard error	Composite Index[2]
Alabama	88.0	74.2	72.7	74.0	15.3	0.2	–15.1 *	4.55	(4,–1)
Alaska	88.0	84.1	72.8	72.5	15.2	11.6	–3.6	2.57	(4,0)
Arizona	73.3	67.3	67.9	73.3	5.5	–6.0	–11.5 *	4.17	(3,–1)
Arkansas	62.0	62.1	60.8	71.8	1.2	–9.7	–11.0 *	5.22	(2,–1)
California	80.1	80.3	71.5	77.1	8.7	3.2	–5.4	2.92	(3,0)
Colorado	76.7	78.1	68.3	74.9	8.4	3.2	–5.3	3.93	(3,0)
Connecticut	84.4	87.0	75.4	75.7	9.1	11.3	2.3	2.42	(3,0)
Delaware	33.3	62.6	62.0	64.4	–28.7	–1.8	26.9 *	3.22	(1,1)
District of Columbia	62.1	33.6	62.6	59.6	–0.5	–26.1	–25.5 *	3.11	(2,–1)
Florida	80.2	83.5	72.3	77.9	7.9	5.5	–2.3	2.52	(3,0)
Georgia	62.4	45.0	69.2	69.7	–6.8	–24.8	–18.0 *	4.46	(2,–1)
Hawaii	82.4	87.6	72.8	74.2	9.6	13.4	3.8	2.40	(4,0)
Idaho	81.5	74.9	70.4	74.9	11.1	0.0	–11.1 *	3.41	(4,–1)
Illinois	72.4	72.4	66.9	70.3	5.6	2.1	–3.5	4.49	(3,0)
Indiana	73.5	70.9	70.4	72.4	3.1	–1.5	–4.5	3.93	(2,0)
Iowa	75.9	71.7	69.6	73.0	6.4	–1.3	–7.7 *	3.73	(3,–1)
Kansas	72.6	65.9	64.3	71.6	8.4	–5.8	–14.1 *	4.07	(3,–1)
Kentucky	44.8	42.4	60.7	63.5	–15.9	–21.1	–5.2	4.17	(1,0)
Louisiana	49.8	80.1	67.0	74.5	–17.2	5.6	22.8 *	5.56	(1,1)
Maine	65.9	67.9	67.3	72.0	–1.4	–4.1	–2.7	3.18	(2,0)
Maryland	69.1	47.8	66.2	67.2	2.8	–19.4	–22.2 *	4.60	(2,–1)
Massachusetts	69.0	68.7	70.3	71.9	–1.3	–3.2	–1.8	4.05	(2,0)
Michigan	56.1	63.4	65.7	71.3	–9.7	–7.8	1.9	4.33	(1,0)
Minnesota	82.8	73.2	71.0	73.4	11.8	–0.2	–12.0 *	3.34	(4,–1)
Mississippi	58.0	63.2	74.1	75.9	–16.1	–12.7	3.5	5.59	(1,0)
Missouri	49.5	76.1	62.7	73.8	–13.2	2.4	15.5 *	4.22	(1,1)
Montana	66.6	69.2	66.4	71.5	0.2	–2.3	–2.5	3.78	(2,0)
Nebraska	77.2	73.8	67.5	74.7	9.7	–0.9	–10.6 *	3.07	(4,–1)
Nevada	74.7	69.5	71.6	70.6	3.1	–1.1	–4.2	3.78	(2,0)
New Hampshire	87.7	80.7	73.8	76.8	13.9	3.9	–10.0 *	2.29	(4,–1)
New Jersey	78.5	64.7	72.6	71.7	5.9	–7.0	–12.9 *	4.55	(3,–1)
New Mexico	68.7	60.5	70.5	73.0	–1.8	–12.5	–10.6 *	3.75	(2,–1)
New York	65.2	66.8	69.4	74.8	–4.1	–8.0	–3.8	4.31	(2,0)
North Carolina	80.9	83.2	71.1	76.9	9.8	6.3	–3.5	3.24	(4,0)
North Dakota	55.9	38.7	65.5	68.2	–9.6	–29.4	–19.9 *	3.24	(1,–1)
Ohio	50.8	50.9	63.9	63.1	–13.1	–12.2	1.0	4.50	(1,0)
Oklahoma	74.2	59.3	65.9	71.2	8.4	–11.8	–20.2 *	3.84	(3,–1)
Oregon	77.4	81.5	68.0	70.7	9.4	10.8	1.4	2.83	(4,0)
Pennsylvania	79.3	73.8	69.0	73.8	10.3	0.1	–10.3 *	4.18	(4,–1)
Rhode Island	84.6	85.3	74.0	78.2	10.6	7.0	–3.6	1.91	(4,0)
South Carolina	52.2	57.2	68.7	72.2	–16.5	–15.0	1.5	4.51	(1,0)
South Dakota	73.0	50.6	62.7	70.3	10.3	–19.6	–29.9 *	3.93	(4,–1)
Tennessee	43.4	40.7	60.1	62.8	–16.7	–22.1	–5.4	4.66	(1,0)
Texas	63.1	56.6	73.7	72.8	–10.6	–16.1	–5.6	3.20	(1,0)
Utah	72.1	61.2	67.6	67.6	4.5	–6.4	–10.9 *	4.17	(2,–1)
Vermont	77.5	73.7	69.3	73.9	8.2	–0.2	–8.4 *	2.37	(3,–1)
Virginia	54.9	55.9	69.6	73.3	–14.7	–17.4	–2.7	3.62	(1,0)
Washington	72.5	67.5	65.6	71.8	7.0	–4.3	–11.3 *	4.17	(3,–1)
West Virginia	62.7	86.6	67.1	74.9	–4.4	11.7	16.1 *	3.43	(2,1)
Wisconsin	68.3	62.9	63.6	66.4	4.7	–3.5	–8.2	4.48	(3,0)
Wyoming	81.6	76.7	71.3	68.9	10.3	7.8	–2.5	2.83	(4,0)

* 2005–2007 change over time significantly different from zero ($p < .05$).

[1] The 2005 *difference from predicted* is also the *starting point* measure.

[2] The *composite index* (*q*,*s*) is the quartile of the *starting point, q* (from 1, the lowest, to 4, the highest), and statistical significance of the *change* score, *s*, where *s* is –1 if the change is negative and statistically significant, 1 if positive and statistically significant, and 0 if changes are not statistically different from zero ($p < .05$).

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 2005 and 2007 Reading Assessments.

Table 13. Actual and second-period predicted percentages of students with disabilities in NAEP grade 4 mathematics assessments, starting point and change measures, using state-specific approach: By state, 2005 and 2007

State	Actual rates		Starting point measure—2005		Predicted	2005–07 Change measure[1]		Composite Index[2]
	2005	2007	Estimated	Std error	2007	Change	Std error	
Alabama	89.6	88.3	7.5 *	1.98	89.5	–1.2	2.62	(4,0)
Alaska	93.7	91.3	8.0 *	1.32	92.6	–1.3	1.96	(4,0)
Arizona	80.3	83.4	–4.2	2.52	79.6	3.8	3.96	(1,0)
Arkansas	86.3	79.8	5.7 *	1.63	85.3	–5.5	3.28	(4,0)
California	81.0	83.9	–4.4 *	1.86	78.5	5.4 *	2.52	(1,1)
Colorado	83.7	88.2	2.8	2.03	81.8	6.4	3.64	(3,0)
Connecticut	87.9	89.7	0.2	1.58	85.6	4.1	2.68	(2,0)
Delaware	58.4	73.1	–23.5 *	1.95	57.1	16.0 *	2.62	(1,1)
District of Columbia	67.7	66.3	–7.3 *	1.97	71.7	–5.4	3.03	(1,0)
Florida	88.6	88.3	3.1	1.84	88.6	–0.3	2.59	(3,0)
Georgia	87.9	83.5	2.3	1.54	89.2	–5.7 *	2.69	(3,–1)
Hawaii	85.2	90.5	3.1	1.90	86.1	4.4	3.55	(3,0)
Idaho	92.2	86.3	9.9 *	1.45	92.2	–5.9 *	2.41	(4,–1)
Illinois	86.5	77.6	3.4 *	1.34	85.9	–8.3 *	3.30	(3,–1)
Indiana	91.7	85.7	4.9 *	1.63	91.0	–5.3 *	2.70	(4,–1)
Iowa	88.3	90.0	2.4	1.25	84.3	5.7 *	2.79	(3,1)
Kansas	85.7	78.9	4.1 *	1.42	83.9	–5.0	3.02	(4,0)
Kentucky	83.7	84.2	1.4	2.21	84.3	–0.1	2.93	(2,0)
Louisiana	83.8	87.8	–1.8	2.06	84.0	3.8	3.16	(2,0)
Maine	83.0	83.7	–0.2	1.95	84.3	–0.6	2.57	(2,0)
Maryland	79.4	72.7	–4.4 *	2.19	79.0	–6.3	3.32	(1,0)
Massachusetts	83.9	74.3	–1.7	1.97	80.6	–6.3	3.75	(2,0)
Michigan	74.6	76.1	–5.9 *	2.17	72.9	3.2	3.12	(1,0)
Minnesota	86.3	85.9	2.6	2.41	83.9	2.0	2.48	(3,0)
Mississippi	80.4	92.3	–8.1 *	2.37	85.3	7.0 *	2.99	(1,1)
Missouri	87.2	76.6	2.6	2.32	85.7	–9.1 *	3.38	(3,–1)
Montana	83.4	81.7	1.1	2.25	82.9	–1.2	3.49	(2,0)
Nebraska	88.7	85.7	2.8	1.53	88.0	–2.3	2.58	(3,0)
Nevada	80.7	84.5	–4.8	2.64	82.1	2.4	3.25	(1,0)
New Hampshire	90.3	88.7	6.4 *	1.26	89.2	–0.5	2.02	(4,0)
New Jersey	87.7	88.0	1.9	2.37	87.0	1.0	3.30	(3,0)
New Mexico	89.3	83.2	5.3 *	2.24	86.5	–3.3	3.62	(4,0)
New York	83.7	91.2	–1.9	1.98	86.1	5.1 *	2.36	(2,1)
North Carolina	87.5	89.7	2.4	1.25	87.9	1.8	1.90	(3,0)
North Dakota	85.5	76.3	–0.9	1.34	80.6	–4.3	3.02	(2,0)
Ohio	73.0	71.5	–7.9 *	3.28	75.4	–3.9	4.44	(1,0)
Oklahoma	78.8	67.8	–1.6	1.73	80.1	–12.3 *	3.85	(2,–1)
Oregon	78.7	85.8	–3.1	2.41	76.7	9.1 *	3.27	(2,1)
Pennsylvania	85.3	86.1	3.2	2.61	85.5	0.6	3.34	(3,0)
Rhode Island	87.9	91.0	1.9	2.05	86.5	4.5	2.51	(2,0)
South Carolina	73.8	88.1	–10.9 *	2.14	76.8	11.3 *	2.62	(1,1)
South Dakota	91.0	92.2	3.9 *	1.06	89.4	2.8	1.89	(4,0)
Tennessee	76.1	59.4	0.7	2.93	75.8	–16.4 *	4.79	(2,–1)
Texas	65.2	62.8	–19.3 *	1.79	65.7	–2.9	2.85	(1,0)
Utah	88.6	84.4	5.2 *	1.50	88.6	–4.2	3.00	(4,0)
Vermont	80.4	86.4	–1.0	1.80	77.3	9.1 *	2.90	(2,1)
Virginia	71.6	74.1	–12.4 *	2.59	72.9	1.2	3.46	(1,0)
Washington	85.4	85.7	5.1 *	1.72	87.0	–1.3	2.56	(4,0)
West Virginia	88.5	91.6	3.0 *	1.42	90.6	1.0	1.64	(3,0)
Wisconsin	88.3	85.5	5.2 *	1.94	88.2	–2.7	2.92	(4,0)
Wyoming	91.8	89.6	7.2 *	1.37	89.4	0.2	2.14	(4,0)

* Significantly different from zero ($p < .05$).

[1] In the state-specific approach, the *change measure* is, by construction, the difference between the 2007 actual and 2007 predicted rates.

[2] The *composite index* (*q*,*s*) is the quartile of the *starting point, q* (from 1, the lowest, to 4, the highest), and statistical significance of the *change* score, *s*, where *s* is –1 if the change is negative and statistically significant, 1 if positive and statistically significant, and 0 if changes are not statistically different from zero ($p < .05$).

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 2005 and 2007 Mathematics Assessments.

Table 14. Actual and second-period predicted percentages of students with disabilities in NAEP grade 8 mathematics assessments, starting point and change measures, using state-specific approach: By state, 2005 and 2007

State	Actual rates		Starting point measure—2005		Predicted	2005–07 Change measure[1]		Composite Index[2]
	2005	2007	Estimated	Std error	2007	Change	Std error	
Alabama	92.3	76.8	11.9 *	1.89	91.5	–14.7 *	4.41	(4,–1)
Alaska	84.2	63.1	4.3 *	1.89	81.2	–18.1 *	3.91	(3,–1)
Arizona	71.4	75.3	–7.7 *	3.89	71.1	4.2	5.68	(1,0)
Arkansas	80.2	81.7	4.4 *	2.20	78.8	2.9	3.45	(3,0)
California	84.4	83.0	4.6 *	1.55	85.9	–2.9	2.64	(3,0)
Colorado	83.9	87.2	5.0 *	1.96	82.4	4.8	2.97	(3,0)
Connecticut	83.7	90.7	1.6	1.37	83.2	7.5 *	2.64	(2,1)
Delaware	34.1	57.1	–39.8 *	2.07	35.3	21.8 *	3.28	(1,1)
District of Columbia	71.2	45.7	–1.4	2.32	72.0	–26.3 *	3.47	(2,–1)
Florida	85.6	83.5	5.9 *	2.23	86.3	–2.8	2.81	(4,0)
Georgia	82.0	50.2	1.2	2.42	86.0	–35.8 *	3.85	(2,–1)
Hawaii	86.1	90.8	5.6 *	1.73	88.0	2.8	2.23	(3,0)
Idaho	86.3	86.6	9.0 *	1.69	87.4	–0.8	2.28	(4,0)
Illinois	83.7	65.6	5.8 *	2.32	83.3	–17.7 *	4.59	(4,–1)
Indiana	76.5	63.8	–1.9	2.47	78.2	–14.4 *	4.46	(2,–1)
Iowa	83.8	84.2	6.0 *	2.01	84.2	0.0	3.44	(4,0)
Kansas	76.6	69.5	1.7	2.61	72.4	–2.9	4.13	(2,0)
Kentucky	72.7	51.5	–5.4 *	2.49	70.6	–19.1 *	4.79	(1,–1)
Louisiana	70.7	74.1	–6.0	4.35	66.3	7.8	5.69	(1,0)
Maine	75.2	71.9	–0.5	2.29	77.1	–5.2	3.66	(2,0)
Maryland	67.2	38.3	–9.1 *	3.71	63.6	–25.3 *	5.23	(1,–1)
Massachusetts	68.6	49.4	–9.3 *	2.65	64.1	–14.7 *	4.32	(1,–1)
Michigan	69.4	69.1	–8.2 *	2.94	65.5	3.6	4.47	(1,0)
Minnesota	85.5	83.4	7.8 *	1.82	84.7	–1.3	2.76	(4,0)
Mississippi	68.6	78.5	–13.5 *	3.91	76.1	2.4	4.09	(1,0)
Missouri	72.5	65.4	–2.6	2.73	72.2	–6.8	4.20	(2,0)
Montana	84.2	77.3	4.9 *	1.64	84.9	–7.6 *	3.07	(3,–1)
Nebraska	91.1	83.0	8.9 *	1.43	91.0	–8.0 *	2.47	(4,–1)
Nevada	82.4	73.8	0.5	2.55	82.6	–8.8	4.78	(2,0)
New Hampshire	87.9	83.3	8.1 *	1.59	87.3	–4.0	2.58	(4,0)
New Jersey	83.1	82.7	1.9	2.70	84.5	–1.8	3.75	(3,0)
New Mexico	87.1	83.5	7.5 *	1.65	80.7	2.8	3.59	(4,0)
New York	81.2	79.3	0.0	2.23	82.9	–3.6	3.47	(2,0)
North Carolina	85.8	86.7	4.7 *	1.90	86.7	0.0	1.99	(3,0)
North Dakota	74.0	58.1	–2.9	1.98	73.2	–15.1 *	3.10	(2,–1)
Ohio	59.9	53.6	–13.2 *	3.82	59.0	–5.4	5.00	(1,0)
Oklahoma	76.5	44.9	1.0	2.38	76.2	–31.3 *	4.21	(2,–1)
Oregon	82.1	78.0	7.4 *	1.66	85.0	–7.0 *	3.37	(4,–1)
Pennsylvania	80.8	77.9	3.2	2.73	81.0	–3.1	4.43	(3,0)
Rhode Island	85.1	88.0	3.1 *	1.55	85.8	2.2	2.37	(3,0)
South Carolina	59.2	60.2	–20.8 *	3.43	66.0	–5.8	5.19	(1,0)
South Dakota	82.9	78.3	5.8 *	1.45	82.4	–4.1	2.83	(3,0)
Tennessee	68.7	47.0	–4.6	2.86	69.1	–22.1 *	5.17	(2,–1)
Texas	61.0	58.5	–18.6 *	2.48	59.9	–1.4	3.52	(1,0)
Utah	82.4	77.1	6.2 *	1.68	81.8	–4.7	3.35	(4,0)
Vermont	79.2	78.0	1.8	1.61	75.8	2.2	2.49	(2,0)
Virginia	70.9	58.9	–10.2 *	2.81	69.1	–10.2 *	4.10	(1,–1)
Washington	83.0	73.1	6.9 *	2.24	80.8	–7.7 *	3.76	(4,–1)
West Virginia	83.0	88.9	2.0	1.74	85.4	3.5	2.18	(3,0)
Wisconsin	78.7	73.6	5.0 *	2.06	76.5	–2.9	4.21	(3,0)
Wyoming	89.4	84.7	9.3 *	1.84	87.0	–2.3	2.79	(4,0)

* Significantly different from zero ($p < .05$).

[1] In the state-specific approach, the c*hange measure* is, by construction, the difference between the 2007 actual and 2007 predicted rates.

[2] The *composite index* (*q*,*s*) is the quartile of the *starting point, q* (from 1, the lowest, to 4, the highest), and statistical significance of the *change* score, *s*, where *s* is –1 if the change is negative and statistically significant, 1 if positive and statistically significant, and 0 if changes are not statistically different from zero ($p < .05$).

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 2005 and 2007 Mathematics Assessments.

Table 15. Actual and second-period predicted percentages of students with disabilities in NAEP grade 4 reading assessments, starting point and change measures, using state-specific approach: By state, 2005 and 2007

State	Actual rates		Starting point measure—2005		Predicted	2005–07 Change measure[1]		Composite Index[2]
	2005	2007	Estimated	Std error	2007	Change	Std error	
Alabama	85.3	78.1	18.8 *	2.04	87.5	–9.4 *	2.94	(4,–1)
Alaska	84.5	80.9	16.4 *	2.63	84.8	–3.9	3.64	(4,0)
Arizona	66.7	74.9	–1.5	2.59	70.8	4.1	4.62	(2,0)
Arkansas	53.3	55.0	–9.5 *	3.24	55.5	–0.5	4.80	(1,0)
California	74.9	77.3	1.4	2.00	75.6	1.7	2.77	(3,0)
Colorado	78.6	79.9	12.9 *	2.25	76.7	3.2	2.82	(4,0)
Connecticut	78.5	85.1	7.7 *	2.52	79.3	5.8	3.30	(4,0)
Delaware	28.6	46.0	–34.0 *	2.16	27.5	18.5 *	2.72	(1,1)
District of Columbia	57.7	28.6	0.7	2.59	58.6	–30.0 *	3.39	(2,–1)
Florida	76.0	77.3	3.2	2.58	76.0	1.3	3.43	(3,0)
Georgia	59.7	43.4	–6.1 *	3.00	60.9	–17.5 *	4.22	(2,–1)
Hawaii	82.5	77.1	12.9 *	2.35	82.4	–5.3	3.79	(4,0)
Idaho	72.2	77.2	6.1 *	2.60	70.6	6.6	4.49	(3,0)
Illinois	65.3	67.5	–1.0	2.72	65.1	2.4	4.56	(2,0)
Indiana	75.2	77.6	5.4 *	2.74	77.0	0.6	4.37	(3,0)
Iowa	65.2	71.2	2.5	3.44	66.7	4.5	4.98	(3,0)
Kansas	77.4	62.4	13.0 *	2.48	76.5	–14.1 *	4.24	(4,–1)
Kentucky	45.4	53.1	–17.0 *	2.29	54.0	–0.9	3.48	(1,0)
Louisiana	41.2	79.3	–27.1 *	3.77	47.7	31.6 *	7.57	(1,1)
Maine	65.8	69.5	–2.9	2.57	62.0	7.5 *	3.47	(2,1)
Maryland	63.3	51.4	–6.0 *	2.48	62.6	–11.2 *	3.52	(2,–1)
Massachusetts	67.2	72.1	–2.1	2.21	64.5	7.6 *	3.37	(2,1)
Michigan	52.3	68.2	–12.5 *	3.39	50.5	17.7 *	4.58	(1,1)
Minnesota	84.5	77.1	12.4 *	1.78	84.2	–7.1 *	2.75	(4,–1)
Mississippi	66.9	77.6	–9.0 *	2.97	69.4	8.2 *	4.16	(1,1)
Missouri	55.8	78.9	–10.6 *	2.60	57.9	21.0 *	3.62	(1,1)
Montana	61.1	64.8	–0.1	2.90	59.3	5.5	4.65	(2,0)
Nebraska	75.2	72.1	2.6	2.65	75.0	–2.9	3.74	(3,0)
Nevada	60.4	70.1	–6.0	3.20	65.2	4.9	4.58	(2,0)
New Hampshire	83.0	79.8	12.6 *	2.26	82.6	–2.8	3.56	(4,0)
New Jersey	73.8	61.7	6.5	3.46	74.6	–12.9 *	5.62	(3,–1)
New Mexico	64.4	54.2	–2.0	3.70	64.1	–9.9	5.11	(2,0)
New York	74.5	72.5	3.8	2.44	76.2	–3.7	4.05	(3,0)
North Carolina	83.2	87.6	12.2 *	1.77	83.7	3.9	2.36	(4,0)
North Dakota	64.7	46.2	–4.8 *	1.95	61.0	–14.8 *	3.08	(2,–1)
Ohio	40.8	50.6	–16.8 *	2.78	45.0	5.6	4.67	(1,0)
Oklahoma	72.5	58.8	6.2 *	2.36	72.2	–13.4 *	4.44	(3,–1)
Oregon	69.4	74.3	0.4	2.34	70.0	4.3	3.49	(2,0)
Pennsylvania	73.3	72.0	7.4 *	2.68	73.3	–1.3	3.98	(3,0)
Rhode Island	88.0	82.9	16.2 *	1.63	86.7	–3.8	2.82	(4,0)
South Carolina	61.1	73.7	–10.5 *	2.42	62.7	11.0 *	3.92	(1,1)
South Dakota	71.4	66.2	3.8 *	1.77	72.6	–6.4 *	2.71	(3,–1)
Tennessee	38.4	37.0	–18.8 *	4.50	42.2	–5.2	5.12	(1,0)
Texas	58.5	51.7	–15.3 *	2.08	59.3	–7.6 *	3.25	(1,–1)
Utah	72.0	62.8	4.1	2.28	71.1	–8.3	4.57	(3,0)
Vermont	68.4	67.0	6.1 *	2.11	68.4	–1.4	2.93	(3,0)
Virginia	36.8	55.5	–26.3 *	2.60	39.2	16.3 *	4.24	(1,1)
Washington	76.8	72.5	11.4 *	2.98	78.0	–5.5	3.84	(4,0)
West Virginia	69.1	91.0	–2.4	2.32	70.2	20.8 *	2.82	(2,1)
Wisconsin	71.7	72.4	7.7 *	3.04	73.6	–1.2	4.00	(4,0)
Wyoming	90.4	78.6	21.1 *	1.45	89.5	–10.9 *	2.38	(4,–1)

* Significantly different from zero ($p < .05$).

[1] In the state-specific approach, the c*hange measure* is, by construction, the difference between the 2007 actual and 2007 predicted rates.

[2] The *composite index* (*q*,*s*) is the quartile of the *starting point, q* (from 1, the lowest, to 4, the highest), and statistical significance of the *chang*e score, *s*, where *s* is –1 if the change is negative and statistically significant, 1 if positive and statistically significant, and 0 if changes are not statistically different from zero ($p < .05$).

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 2005 and 2007 Reading Assessments.

Table 16. Actual and second-period predicted percentages of students with disabilities in NAEP grade 8 reading assessments, starting point and change measures, using state-specific approach: By state, 2005 and 2007

State	Actual rates		Starting point measure—2005		Predicted	2005–07 Change measure[1]		Composite Index[2]
	2005	2007	Estimated	Std error	2007	Change	Std error	
Alabama	88.0	74.2	14.4 *	2.30	87.4	–13.2 *	4.70	(4,–1)
Alaska	88.0	84.1	14.4 *	1.46	87.4	–3.3	2.41	(4,0)
Arizona	73.3	67.3	4.4	2.54	71.3	–4.0	5.81	(3,0)
Arkansas	62.0	62.1	–1.1	3.43	63.9	–1.8	5.41	(2,0)
California	80.1	80.3	7.6 *	1.75	82.9	–2.6	2.73	(3,0)
Colorado	76.7	78.1	7.0 *	2.78	78.8	–0.7	3.43	(3,0)
Connecticut	84.4	87.0	8.1 *	1.79	85.2	1.8	2.82	(3,0)
Delaware	33.3	62.6	–31.4 *	2.63	37.4	25.2 *	3.32	(1,1)
District of Columbia	62.1	33.6	–2.3	2.06	59.7	–26.1 *	3.09	(2,–1)
Florida	80.2	83.5	7.0 *	1.84	81.2	2.3	2.32	(3,0)
Georgia	62.4	45.0	–8.6 *	3.47	66.0	–21.0 *	4.66	(2,–1)
Hawaii	82.4	87.6	8.9 *	1.81	83.1	4.5	2.36	(4,0)
Idaho	81.5	74.9	9.7 *	2.18	77.2	–2.3	4.04	(4,0)
Illinois	72.4	72.4	4.5	3.25	70.3	2.1	4.52	(3,0)
Indiana	73.5	70.9	1.8	2.66	73.9	–3.0	4.25	(2,0)
Iowa	75.9	71.7	5.4 *	2.24	74.4	–2.7	3.88	(3,0)
Kansas	72.6	65.9	6.5 *	3.00	74.9	–9.0 *	4.02	(3,–1)
Kentucky	44.8	42.4	–18.9 *	2.90	48.5	–6.1	4.60	(1,0)
Louisiana	49.8	80.1	–19.6 *	3.86	52.9	27.2 *	4.98	(1,1)
Maine	65.9	67.9	–3.4	2.39	65.8	2.1	3.40	(2,0)
Maryland	69.1	47.8	0.8	3.06	69.2	–21.4 *	4.82	(2,–1)
Massachusetts	69.0	68.7	–2.6	2.51	70.4	–1.7	4.34	(2,0)
Michigan	56.1	63.4	–11.6 *	3.13	58.0	5.4	4.78	(1,0)
Minnesota	82.8	73.2	10.5 *	1.92	84.1	–10.9 *	3.08	(4,–1)
Mississippi	58.0	63.2	–16.9 *	4.03	64.0	–0.8	6.00	(1,0)
Missouri	49.5	76.1	–15.4 *	3.39	49.8	26.3 *	4.12	(1,1)
Montana	66.6	69.2	–1.2	2.46	67.5	1.7	3.90	(2,0)
Nebraska	77.2	73.8	8.7 *	1.84	80.0	–6.2	3.17	(4,0)
Nevada	74.7	69.5	1.8	2.38	72.7	–3.2	4.04	(2,0)
New Hampshire	87.7	80.7	12.9 *	1.29	88.2	–7.5 *	2.32	(4,–1)
New Jersey	78.5	64.7	4.6	2.92	78.7	–14.0 *	5.03	(3,–1)
New Mexico	68.7	60.5	–2.8	2.58	69.5	–9.0 *	3.74	(2,–1)
New York	65.2	66.8	–5.3	2.77	71.6	–4.8	4.69	(2,0)
North Carolina	80.9	83.2	8.4 *	2.27	81.8	1.4	2.93	(4,0)
North Dakota	55.9	38.7	–11.6 *	2.07	53.7	–15.0 *	3.27	(1,–1)
Ohio	50.8	50.9	–15.2 *	3.43	50.3	0.6	4.50	(1,0)
Oklahoma	74.2	59.3	7.0 *	2.64	76.3	–17.0 *	3.84	(3,–1)
Oregon	77.4	81.5	8.4 *	1.63	77.7	3.8	4.06	(4,0)
Pennsylvania	79.3	73.8	9.2 *	2.62	80.7	–6.9	3.90	(4,0)
Rhode Island	84.6	85.3	9.9 *	1.15	85.6	–0.3	1.82	(4,0)
South Carolina	52.2	57.2	–17.8 *	3.33	60.4	–3.2	5.45	(1,0)
South Dakota	73.0	50.6	8.6 *	2.15	74.4	–23.8 *	3.80	(4,–1)
Tennessee	43.4	40.7	–19.0 *	3.45	46.4	–5.7	5.49	(1,0)
Texas	63.1	56.6	–10.9 *	1.80	62.3	–5.7 *	2.72	(1,–1)
Utah	72.1	61.2	3.1	2.70	67.2	–6.0	4.46	(2,0)
Vermont	77.5	73.7	6.9 *	1.62	79.1	–5.4 *	2.33	(3,–1)
Virginia	54.9	55.9	–16.3 *	2.32	58.9	–3.0	4.04	(1,0)
Washington	72.5	67.5	5.7 *	2.75	73.6	–6.1	4.09	(3,0)
West Virginia	62.7	86.6	–6.1 *	2.91	67.4	19.2 *	3.68	(2,1)
Wisconsin	68.3	62.9	3.3	3.06	63.3	–0.4	4.82	(3,0)
Wyoming	81.6	76.7	9.4 *	2.08	80.0	–3.3	3.23	(4,0)

* Significantly different from zero ($p < .05$).

[1] In the state-specific approach, the c*hange measure* is, by construction, the difference between the 2007 actual and 2007 predicted rates.

[2] The *composite index* (*q*,*s*) is the quartile of the *starting point, q* (from 1, the lowest, to 4, the highest), and statistical significance of the *chang*e score, *s*, where *s* is –1 if the change is negative and statistically significant, 1 if positive and statistically significant, and 0 if changes are not statistically different from zero ($p < .05$).

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 2005 and 2007 Reading Assessments.

Table 17. Number of states in each composite index score category by estimation approach based on NAEP grades 4 and 8 mathematics assessments: 2005 and 2007

Starting Quartile	Grade 4						Grade 8					
	Nation-based			State-specific			Nation-based			State-specific		
	▼	=	▲	▼	=	▲	▼	=	▲	▼	=	▲
4 more inclusive	2	11	0	2	11	0	5	8	0	5	8	0
3	2	11	0	3	9	1	2	11	0	2	11	0
2	3	6	4	2	8	3	6	6	1	6	6	1
1 less inclusive	1	8	3	0	8	4	4	7	1	4	7	1
Overall	8	36	7	7	36	8	17	32	2	17	32	2

▲ Measure of change is positive and statistically significant indicating higher inclusion rates ($p < .05$).

= Measure of change is not statistically significant ($p < .05$).

▼ Measure of change is negative and statistically significant indicating lower inclusion rates ($p < .05$).

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 2005 and 2007 Mathematics Assessments.

Table 18. Number of states in each composite index score category by estimation approach based on NAEP grades 4 and 8 reading assessments: 2005 and 2007

Starting Quartile	Grade 4						Grade 8					
	Nation-based			State-specific			Nation-based			State-specific		
	▼	=	▲	▼	=	▲	▼	=	▲	▼	=	▲
4 more inclusive	8	5	0	4	9	0	7	6	0	4	9	0
3	6	7	0	3	10	0	7	6	0	4	9	0
2	4	8	1	4	6	3	6	6	1	4	8	1
1 less inclusive	0	5	7	1	4	7	1	8	3	2	7	3
Overall	18	25	8	12	29	10	21	26	4	14	33	4

▲ Measure of change is positive and statistically significant indicating higher inclusion rates ($p < .05$).

= Measure of change is not statistically significant ($p < .05$).

▼ Measure of change is negative and statistically significant indicating lower inclusion rates ($p < .05$).

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 2005 and 2007 Reading Assessments.

Table 19. States in each composite index score category by estimation approach based on NAEP grade 4 mathematics assessments: 2005 and 2007

Starting Quartile	Grade 4					
	Nation-based			State-specific		
	▼	=	▲	▼	=	▲
4 more inclusive	AR, ID	AL, AK, HI, IN, KS, NH, NM, UT, WA, WI, WY		ID, IN	AL, AK, AR, KS, NH, NM, SD, UT, WA, WI, WY	
3	IL, MO	CO, FL, GA, IA, MN, NE, NJ, NC, PA, SD, WV		GA, IL, MO	CO, FL, HI, MN, NE, NJ, NC, PA, WV	IA
2	ND, OK, TN	KY, LA, MA, ME, MT, RI	CT, NY, OR, VT	OK, TN	CT, KY, LA, ME, MA, MT, ND, RI	NY, OR, VT
1 less inclusive	MD	AZ, CA, DC, MI, NV, OH, TX, VA	DE, MS, SC		AZ, DC, MD, MI, NV, OH, TX, VA	CA, DE, MS, SC
Total	8	36	7	7	36	8

▲ Measure of change is positive and statistically significant indicating higher inclusion rates ($p < .05$).

= Measure of change is not statistically significant ($p < .05$).

▼ Measure of change is negative and statistically significant indicating lower inclusion rates ($p < .05$).

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 2005 and 2007 Mathematics Assessments.

Table 20. States in each composite index score category by estimation approach based on NAEP grade 8 mathematics assessments: 2005 and 2007

Starting Quartile	Grade 8					
	Nation-based			State-specific		
	▼	=	▲	▼	=	▲
4 more inclusive	AL, NE, NH, OR, WA	FL, ID, IA, MN, NM, SD, UT, WY		AL, IL, NE, OR, WA	FL, ID, IA, MN, NH, NM, UT, WY	
3	AK, IL	AR, CA, CO, HI, MT, NJ, NC, PA, RI, VT, WI		AK, MT	AR, CA, CO, HI, NJ, NC, PA, RI, SD, WV, WI	
2	DC, GA, IN, ND, OK, TN	KS, ME, MO, NV, NY, WV	CT	DC, GA, IN, ND, OK, TN	KS, ME, MO, NV, NY, VT	CT
1 less inclusive	KY, MD, MA, VA	AZ, LA, MI, MS, OH, SC, TX	DE	KY, MD, MA, VA	AZ, LA, MI, MS, OH, SC, TX	DE
Total	17	32	2	17	32	2

▲ Measure of change is positive and statistically significant indicating higher inclusion rates ($p <$. 05).

= Measure of change is not statistically significant ($p < .05$).

▼ Measure of change is negative and statistically significant indicating lower inclusion rates ($p < .05$).

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 2005 and 2007 Mathematics Assessments.

Table 21. States in each composite index score category by estimation approach based on NAEP grade 4 reading assessments: 2005 and 2007

Starting Quartile	Grade 4					
	Nation-based			State-specific		
	▼	=	▲	▼	=	▲
4 more inclusive	AL, KS, MN, NH, RI, WA, WI, WY	AK, CO, HI, NC, PA		AL, KS, MN, WY	AK, CO, CT, HI, NH, NC, RI, WA, WI	
3	DC, NJ, OK, SD, UT, VT	CT, FL, ID, IN, IA, NE, NY		NJ, OK, SD	CA, FL, ID, IN, IA, NE, NY, PA, UT, VT	
2	GA, MD, NM, ND	AZ, CA, IL, ME, MA, MT, NV, OR	WV	DC, GA, MD, ND	AZ, IL, MT, NV, NM, OR	ME, MA, WV
1 less inclusive		AR, KY, OH, TN, TX	DE, LA, MI, MS, MO, SC, VA	TX	AR, KY, OH, TN	DE, LA, MI, MS, MO, SC, VA
Total	18	25	8	12	29	10

▲ Measure of change is positive and statistically significant indicating higher inclusion rates ($p < .05$).

= Measure of change is not statistically significant ($p < .05$).

▼ Measure of change is negative and statistically significant indicating lower inclusion rates ($p < .05$).

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 2005 and 2007 Reading Assessments.

Table 22. States in each composite index score category by estimation approach based on NAEP grade 8 reading assessments: 2005 and 2007

Starting Quartile	Grade 8					
	Nation-based			State-specific		
	▼	=	▲	▼	=	▲
4 more inclusive	AL, ID, MN, NE, NH, PA, SD	AK, HI, NC, OR, RI, WY		AL, MN, NH, SD	AK, HI, ID, NE, NC, OR, PA, RI, WY	
3	AZ, IA, KS, NJ, OK, VT, WA	CA, CO, CT, FL, IL, WI		KS, NJ, OK, VT	AZ, CA, CO, CT, FL, IL, IA, WA, WI	
2	AR, DC, GA, MD, NM, UT	IN, ME, MA, MT, NV, NY	WV	DC, GA, MD, NM	AR, IN, ME, MA, MT, NV, NY, UT	WV
1 less inclusive	ND	KY, MI, MS, OH, SC, TN, TX, VA	DE, LA, MO	ND, TX	KY, MI, MS, OH, SC, TN, VA	DE, LA, MO
Total	21	26	4	14	33	4

▲ Measure of change is positive and statistically significant indicating higher inclusion rates ($p < .05$).

= Measure of change is not statistically significant ($p < .05$).

▼ Measure of change is negative and statistically significant indicating lower inclusion rates ($p < .05$).

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 2005 and 2007 Reading Assessments.

STATE-SPECIFIC RESULTS

Results for the state-specific approach are presented in tables 13 through 16.[26] Likelihood ratio tests were all significant at the .01 level indicating all models fit well, but pseudo R^2 statistics were low for a few states and high in others. Logit pseudo R^2 statistics ranges were .18 to .70 for mathematics grade 4, .13 to .60 for grade 8, .16 to .67 for reading grade 4, and .14 to .34 for reading grade 8. These mixed results indicate that the model may not explain a large amount of the variation in inclusion. The consequence for our analysis of poor fit is that the portion of the difference in actual inclusion rates across time explained by the controls will be smaller and the portion of the difference captured by our change measure will be larger. This will lead to larger magnitudes, negative and positive, in the change measure.

As shown in the first row of table 13, in 2005, Alabama included 89.6 percent of the state's SD students in 2005 and 88.3 percent in 2007. Alabama, as reported in the third column, had an inclusion rate that was 7.5 percentage points above the average of all states if all states had had the same characteristics as Alabama. This starting point measure was statistically different from zero at the 5 percent level. The fifth and sixth columns show Alabama's measure of change, –1.2, and its standard error, 2.62. This change was not statistically significant. Alabama had a starting point measure that placed it in the top quartile and a change score of 0; therefore, it had a composite index score of (4,0) as indicated in the last column.

The results in tables 13 through 16 generally follow patterns similar to those in tables 9 through 12, although there were slightly fewer statistically significant changes between 2005 and 2007 here. The numbers of states that were more inclusive of SDs in 2007 than in 2005 were again generally outnumbered by the numbers of states that were less inclusive:

- For grade 4 mathematics, 8 states were significantly more inclusive in 2007 than in 2005, and 7 states were less inclusive.
- For grade 8 mathematics, 2 states were significantly more inclusive, and 17 states were significantly less inclusive.
- For grade 4 reading, 10 states were significantly more inclusive, and 12 states were significantly less inclusive.
- For grade 8 reading, 4 states were significantly more inclusive, and 14 states were significantly less inclusive.

As in the nation-based analysis, states that were more inclusive in 2007 had starting point measures that placed them in the lowest two quartiles, as shown in tables 17 and 18. Many of the states that had negative significant change were located in upper quartiles but again were spread out among other quartiles. Tables 19 through 22 display the results with each state's abbreviation in the appropriate bin. As in tables 17 and 18, the rows in tables 19 through 22 indicate different starting points, and the columns indicate different directions of change. Upper rows are states with higher starting point indicating that they are relatively more inclusive. The left column signifies negative significant change, the right column signifies positive significant change, and the center column shows states with no significant change.

COMPARISON OF NATION-BASED AND STATE-SPECIFIC RESULTS

Looking at the raw change and starting point measures, the two approaches produce very similar results. The correlation between the nation-based and state-specific state-level change measures is greater than .95 for each subject and grade. The correlation between nation-based

26 Results for changes from 2003 to 2005 using the state-specific approach are in appendix tables B-5 through B-8.

and state-specific starting point measures is even higher at greater than .99 for each grade and subject.

Looking at the composite index scores, results for the two approaches are again generally the same but with some differences. When there are differences, they are most often due to differing statistical significance of the change measure. The nation-based approach finds more states becoming less inclusive, particularly in reading as shown in table 18, but they both find roughly the same number of states becoming more inclusive. Because the approaches for comparing the inclusion rates among states in the initial period are very similar, it is not surprising to find only a few discrepancies in the composite index score that are due to differences in the starting point quartile. As an example, comparing the results for mathematics grade 4 in table 9 for the nation-based approach and table 13 for the state-specific approach, 8 states differ in their composite index score because of differences in the significance of the change measure, but only 2 differ because of differences in the quartile of the index comparing states' inclusion rates.

SUMMARY

In response to concerns that rates of inclusion of students with disabilities on NAEP differ among states, we have developed two approaches using Oaxaca-Blinder decomposition techniques for measuring change in inclusion rates over time. These approaches measure change over time, holding constant both the proportion of students with different types and severities of disabilities and whether the student received an accommodation on the state assessment that is not allowed on NAEP. The approaches differ in how the student-level predicted probabilities are set. Under the nation-based approach, student-level predicted probabilities are set by a regression model that is estimated using observations from all states in the initial period. Under the state-specific approach, student-level predicted probabilities are set for each state separately using regression estimates for individual states in the initial period. For both approaches, we compare state inclusion rates in the initial period against each other to provide a *context* for the measured change. We expect states starting with higher relative inclusion rates to have less change than states starting out with lower relative inclusion rates. To make these two measures, one of the starting point and the other of change over time, easier to understand, a partition of the *starting point vs. change* space was introduced. The composite score index brings the partitions of these two measures together.

The measures developed in this study are limited by the validity of the variables used for identifying different types of students. To the extent that variables describing student characteristics are measured with error, our ability to control for changes in the distribution of students among these types is limited. Were we to have perfect measures, it is possible that we would see more captured by our controls and less captured in the change measure.

Between 2005 and 2007, in about one-third of the states there were significant changes in inclusion rates for the mathematics assessments, and in about one-half of the states there were significant changes in inclusion rates in the reading assessments. Overall, more states had, after adjusting for differences in SD populations, lower inclusion rates than higher inclusion rates of students with disabilities on NAEP in 2007 than 2005 except in grade 4 mathematics, where the numbers were about the same from 2005 to 2007. The nation-based approach identified more states as having a significant reduction in their inclusion rates than the state-specific approach did. Differences generally lay in differing significance levels of the estimated change measure. Differences in starting point quartile were relatively rare: in mathematics grade 4, two states switched quartiles; in mathematics grade 8, four states switched; in reading grade 4, four states switched; and in reading grade 8, zero states switched. Most states did not have statistically significant changes in their rate of inclusion. Most of the significant changes were in the direction of decreasing inclusion in 2007 compared with 2005. Those states that had increased their inclusion rates from 2005 to 2007 had relatively low initial inclusion rates in 2005.

REFERENCES

Blinder, A. S. (1973). Wage Discrimination: Reduced Form and Structural Variables. *Journal of Human Resources, 8*: 436–455.

Fairlie, R. W. (2003). *An Extension of the Blinder-Oaxaca Decomposition Technique to Logit and Probit Models* (Discussion Paper No. 873). New Haven, CT: Yale University Economic Growth Center.

Forgione Jr., P. D. (1999). *Issues Surrounding the Release of the 1998 NAEP Reading Report Card.* Testimony to the Committee on Education and the Workforce, U.S. House of Representatives, on May 27, 1999. Retrieved March 16, 2006 from http://www.house.gov/ed_workforce/hearings/106th/oi/naep52799/forgione.htm.

Jones, M. P. (1996, March). Indicator and Stratification Methods for Missing Explanatory Variables in Multiple Linear Regression. *Journal of the American Statistical Association, 91:* 222–230.

McLaughlin, D. H. (2000). *Protecting State NAEP Trends from Changes in SD/LEP Inclusion Rates* (Report to the National Institute of Statistical Sciences). Palo Alto, CA: American Institutes for Research.

McLaughlin, D. H. (2001). Exclusions and Accommodations Affect State NAEP Gain Statistics: Mathematics, 1996 to 2000 (appendix to chapter 4 in the NAEP Validity Studies Report on Research Priorities). Palo Alto, CA: American Institutes for Research.

McLaughlin, D. H. (2003). Full Population Estimates of Reading Gains Between 1998 and 2002 (Report to NCES supporting inclusion of full population estimates in the report of the 2002 NAEP reading assessment). Palo Alto, CA: American Institutes for Research.

Nagelkerke, N. J. D. (1991). A Note on a General Definition of the Coefficient of Determination. *Biometrika*, *78*: 691-692.

National Assessment of Educational Progress Authorization Act of 2002. Retrieved April 1, 2008, from http://www.nagb.org/about/plaw.html.

Oaxaca, R. (1973). Male-female Wage Differentials in Urban Labor Markets. *International Economic Review, 14:* 693–709.

Reschly, D. J. (1996). Identification and Assessment of Students with Disabilities. *The Future of Children*, *6*: 40-53.

United States Department of Education. (2002). *No Child Left Behind: A Desktop Reference.* Washington, DC: Office of Elementary and Secondary Education.

United States Government Accountability Office. (2005). *No Child Left Behind Act: Most Students with Disabilities Participated in Statewide Assessments, but Inclusion Options Could Be Improved* (GAO-05-618). Washington, DC: Author.

Vach, W., and Blettner, M. (1991). Biased Estimation of the Odds Ratio in Case-Control Studies Due to the Use of Ad Hoc Methods of Correcting for Missing Values for Confounding Variables. *American Journal of Epidemiology, 134:* 895–907.