NBER WORKING PAPER SERIES

DOES SPECIAL EDUCATION RAISE ACADEMIC ACHIEVEMENT FOR STUDENTS WITH DISABILITIES?

Eric A. Hanushek
John F. Kain
Steven G. Rivkin

Working Paper 6690
http://www.nber.org/papers/w6690

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 1998

Does Special Education Raise Academic Achievement for Students with Disabilities?
Eric A. Hanushek, John F. Kain, and Steven G. Rivkin
NBER Working Paper No. 6690
August 1998
JEL No. I2, H4

ABSTRACT

While special education has become a hotly debated issue of school policy, most of the discussion has centered on the aggregate costs of providing mandated programs for disabled children. Little attention has been paid to the effectiveness of such programs or possible interactions with the provision of regular education. This study, building on the unique data of the Harvard/UTD Texas Schools Project, provides direct evidence on the effectiveness of special education programs. The average special education program boosts mathematics and reading achievement of special education students, particularly those classified as learning disabled, while not detracting from regular education students. These results are estimated quite precisely from models of fixed effects in achievement gains, and they are robust to a series of specification tests. At this stage, it is not possible to judge whether the program benefits are sufficiently large to justify the added spending involved.

Eric A. Hanushek
Wallis Institute of Political Economy
University of Rochester
Rochester, NY 14627
and NBER
hanu@troi.cc.rochester.edu

John F. Kain
Green Ctr. for the Study of Science and Society
University of Texas at Dallas
P.O. Box 830688
Richardson, TX 75083-0688
jkain@utdallas.edu

Steven G. Rivkin
Amherst College
Department of Economics
Amherst, MA 01002
rivkin@acesgr.econ.amherst.edu

Does Special Education Raise Academic Achievement for Students with Disabilities?

by Eric A. Hanushek, John F. Kain, and Steven G. Rivkin*

One of the most discussed but least analyzed issues in education today is special education. A disproportionate amount of school funding goes to the education of disabled children—perhaps as much as one-fifth of total current spending for slightly more than 10 percent of students. Yet, extraordinarily little evidence has accumulated about the effectiveness of special education programs in raising achievement. Nor is much known about the impact on regular education students of changes in the composition of classrooms brought about by expansions and modifications to special education programs.[1]

The main explanation for the lack of empirical analysis is that differences between special education and nonspecial education students inhibit the study of special education. A comparison of special education and nonspecial education students does not provide a valid measure of program effectiveness, because special education students by definition differ in some respect, implying that achievement differences confound program effects with other factors. Similarly, the correlation between achievement for nonspecial education students and the percentage of the student body classified as special education does not provide an unbiased measure of the impact on regular education

---

*University of Rochester and National Bureau of Economic Research, University of Texas at Dallas, and Amherst College, respectively. The analysis in this paper has been supported by grants from the Smith Richardson Foundation and the William H. Donner Foundation. Much of the original data development was funded by grants from the Spencer Foundation. An earlier version of this paper was prepared for the conference on Achieving Universal Literacy held at the Green Center of the University of Texas at Dallas in April 1998. Julie Cullen not only provided many useful comments but also gave us access to the fiscal instruments she had developed for special education in Texas. We also thank seminar participants at the University of California, Santa Cruz, and at the Institute for Research on Poverty at the University of Wisconsin, Madison, for many helpful comments.

[1]Recent work on the question of whether special education spending detracts from spending on regular classrooms provides evidence on possible fiscal effects of increased spending on special education. Cullen (1997) examines the impact of special education programs on financing and performance for regular education students. Lankford and Wykoff (1996) document changes in public school expenditures on regular and special education in New York public schools during the 1980s.

students, because differences among schools in special education classification rates are likely to be correlated with other factors that affect achievement. Moreover, any expansion or contraction of special education programs alters the composition of the student body and consequently the average academic performance of nonspecial education students.

This analysis attempts to overcome these methodological problems by exploiting longitudinal information on individual students that is found in the extraordinarily rich data set constructed under the Harvard/UTD Texas Schools Project. The large number of special education students in this data set, which follows several entire cohorts of Texas elementary school students for a number of grades, permits detailed investigations of the effects of special education placement on student achievement for students with different types of disabilities. Comparisons of academic performance before and after placement into special education provide much better evidence of programmatic effects than existing cross-sectional data. Rather than identifying the impact of special education by achievement differences between special and nonspecial education students, individual fixed effects models identify the impact by measuring any change in achievement resulting from a change in special education status.

Despite controlling for all time-invariant unobserved differences among students, a potential problem with the fixed effect approach employed here is the assumption that changes in special education status are not accompanied by other changes that affect achievement. If a deterioration in skills accompanies classification as disabled, fixed effects models would tend to underestimate the impact of special education. Conversely, if a transitory downturn in prior year achievement raises the probability of classification as special education, these models would tend to overestimate the impact of special education by attributing the recovery from a temporary negative shock to the change in programs. We address the issue of endogeneity bias in a number of ways, including the use of instrumental variables based upon differences across districts in the fiscal incentives to classify students as disabled induced by state finance law.

Following the analysis of the average impact of special education programs, we investigate the possibility that specific types of special education programs produce systematically different outcomes. Of particular interest is the impact of mainstreaming on achievement. The state of Texas emphasizes the importance of educating students in the least restrictive environment and has reinforced this policy through the use of fiscal incentives. Straightforward extensions of the basic estimation strategy for average special education effects are employed to investigate achievement differences between mainstreamed and nonmainstreamed special education students. The state's decision to increase substantially the financial incentive to mainstream students provides a potentially exogenous source of identification, particularly because the magnitude of the change varied by district wealth.

The increase in mainstreaming and special education programs more generally may also affect nonspecial education students by altering the student composition and resources in regular classrooms. While evidence suggests that special education programs may have crowded out regular education spending in New York and Texas, there is little evidence on the impact of fiscal crowding out on achievement[2] and virtually no evidence on the effects of changes in classroom composition on nonspecial education students. We can directly examine these questions by considering the intertemporal pattern of achievement gains for non-special education students as the fraction classified as disabled varies over time.

The primary results are straightforward. Special education programs on average boost the achievement of students provided this special treatment. This fundamental result, which emerges once individual differences are adequately considered, is robust to alternative estimation approaches that deal with issues of endogenous placement into special education. Balanced against this, we find that special education students do not detract from the education of regular education students. Rather achievement

[2]Cullen (1997) finds evidence that fiscal crowding out harms nonspecial education students using aggregate school and district level data.

gains for students who do not receive special education are positively related to the percentage of students classified as both physically and learning disabled, and there is little or no evidence that mainstreaming systematically harms nonspecial education students. Whether it is the additional revenue obtained from placing more students in special education or other changes in the regular classroom environment that accounts for this positive relationship is unclear and requires further investigation.

The results here apply just to achievement outcomes of special education, and just to the special education population taking the standardized Texas tests. Moreover, no consideration is given to the costs of special education programs. Nonetheless, the evidence provides a convincing case that the typical special education program provides the intended benefits without harming achievement for the nonspecial education population.

Background

The Individuals with Disabilities Education Act, or IDEA, translated concerns about the education of children with both physical and mental disabilities into federal law with its enactment in 1975.[3] This Act prescribed a series of diagnostics, counseling activities, and services to be provided for disabled students. While the data are sketchy, it appears that a large number of children previously excluded were subsequently brought into the public schools. Moreover, they were given legal rights to an education appropriate for them (see Singer and Butler 1987). To implement this and subsequent laws and regulations, school systems expanded staff and programs, developing entirely new administrative structures in many cases. The general thrust has been to provide regular classroom instruction where possible ("mainstreaming") along with specialized instruction to deal with specific

---

[3]This Act, P.L. 94-142, was originally the Education for All Handicapped Children Act and was re-titled IDEA in 1990. It is commonly identified as having direct and significant effects on the cost and methods of delivery of local education. See discussion and evaluation in Hartman [1980], Singer and Butler [1987], and Monk [1990].

needs. The existence of partial categorical funding from the state and federal governments and of intensive instruction for individual students creates both incentives for school systems to expand the population of identified special education students and incentives for parents to seek admission of their children into special education programs (see Hartman (1980), Monk (1990), Sack (1998)).[4] The result has been growth in the number of special education students even as the total student population has fallen.

Figure 1 shows the aggregate changes between 1977 and 1994 in the population identified as disabled.[5] Despite the fact that overall public school enrollment remains roughly constant over this period, the number of students classified as disabled increases from 3.7 million in 1977 to 5.3 million in 1994, causing the percentage of students classified as disabled to increase from 8.3 to 12.2 percent. Figure 1 shows that virtually all of the growth comes from increases in students classified as having specified learning disabilities.[6] Students with learning disabilities grow from 22 percent to 46 percent of all disabled students over this period. This category encompasses a continuum of learning conditions where it is difficult to describe and apply a precise cutoff in evaluation and assessment. This discretion also leads to considerable variation in classification across states, districts, and time(Reschly 1996,

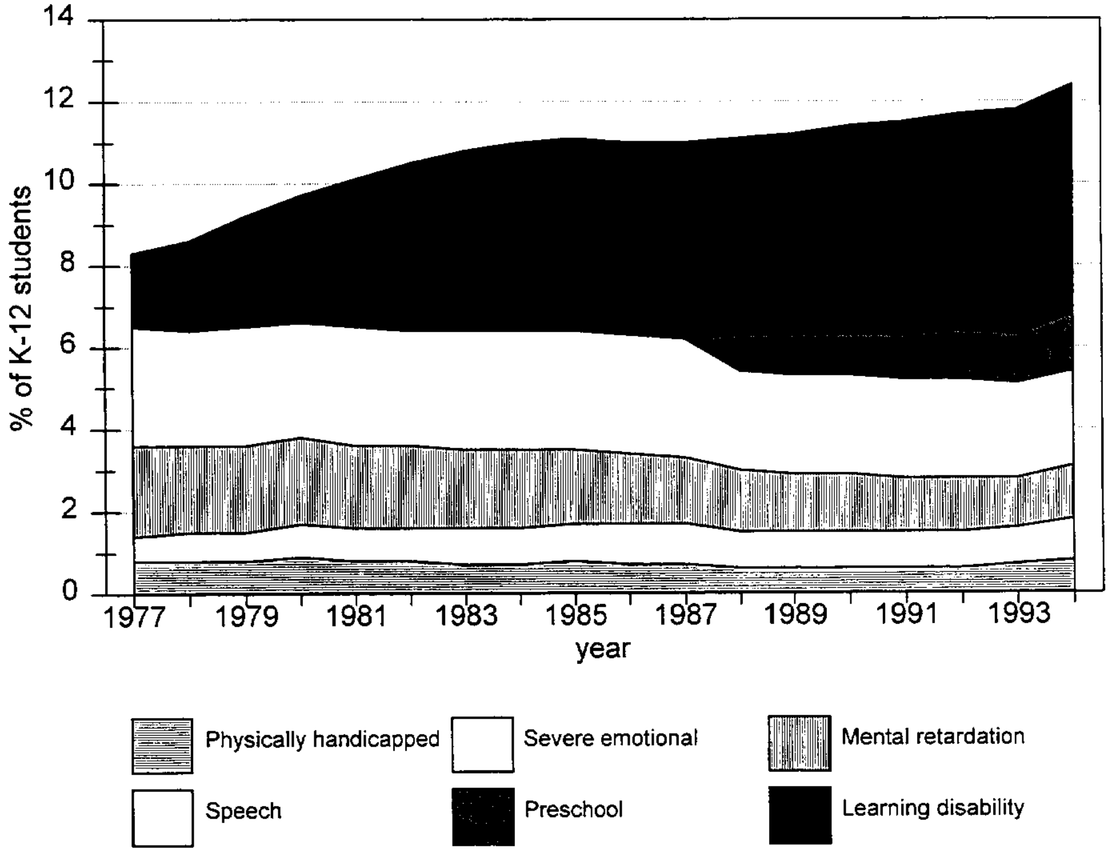
---

[4]The financing of special education differs significantly across states and localities. In overall terms, about 8 percent of special education funding is federal, some 56 percent comes directly from states, and the remainder is local. These shares are approximately equivalent to the shares of total elementary and secondary spending. Moreover, as with total funding of public schools, wide variation in the state funding formulae exist—leading to the possibility that the different incentives for classifying students have an impact on the operations of special education programs (Parrish and Chambers 1996).

[5]Data on special education comes from annual reports required as part of the Individuals With Disabilities Education Act of 1976. Prior to this Act, no consistent data on handicapped students or their schooling are available.

[6]Note that students age 3-5 in preschool programs appear to increase in 1988. This jump is an accounting artifact, deriving from removal of a prior requirement that states had to classify eligible preschool students by specific disability. Thus, while these students were spread across over categories before 1987-88, in that year and after preschool disabled students were reported separately.

Figure 1
% Special Education, by disability



Lewit and Baker, 1996). The more clearly defined physical disabilities represent less than ten percent of special education students.

The expansion of special education has raised concerns about adverse impacts on resources and school quality for non-special education students. While it is sometimes stated that special education accounted for the entire growth in school spending in recent years, the overall cost implications of the growth in special education can be put into perspective by viewing the growth in educational spending over the decade of the 1980s. These calculations, described in detail in Hanushek and Rivkin (1997), use aggregate data on enrollment, staff, and the cost of special education services to investigate how the expansion of special education influenced the overall growth in per student spending. The evidence suggests that special education accounted for roughly 20 percent of the increase in education spending, slightly less than double the share of special education students.[7] Thus, though the special education sector had a disproportionate effect on spending, it certainly did not account for the majority of the U.S. average increases in real per-student expenditures as some have suggested.

The fiscal impact of special education does rise in times of fiscal stringency, because the legal status of such spending dictates that it takes precedent over regular education spending. The decade of the 1990s has witnessed just such stringency in fiscal conditions, and recent data provide some indication of the future course of school expenditure. Perhaps the most significant fact has been that the growth rate in expenditure per student appears to have fallen precipitously in the early 1990s. While real spending per pupil grew at a 3.75 percent real annual rate in the 1980s, there was essentially

[7]A variety of caveats and cautions are also necessary. The calculations summarized in the text concentrate just on the decade of the 1980s. The growth during the 1970s in expenditure that is related to special education is clearly larger. Before the 1975 legislation, many students in need of special services apparently did not even attend school. Nevertheless, because of a lack of reporting requirements and data collection, it is not possible to get any overall estimates of the growth in expenditure in the 1970s that resulted from special education. There are wide variations in the costs of different handicapping conditions which will affect these calculations (see Chaikind, Danielson and Brauen (1993) for a discussion of cost differences), although the largest recent growth in students has come in less expensive categories such as less severe learning disabilities.

no growth from 1990 through 1996 (National Center for Education Statistics (1996)). One immediate impact of this slowed growth in spending is that special education spending—which appears to have a different dynamic—becomes relatively more important. Because of the mandated status of much special education spending, expansion of special education in either scope or intensity takes a larger share of any new money when there is lessened total budgetary growth.[8] With the continued rise in the special education classification rate, it is likely that special education will become more, not less, of a policy issue, making it even more important to identify program benefits and costs.[9]

The Texas Schools Microdata Panel

The cornerstone of this research is the analysis of a unique microdata set of school operations constructed by the Harvard/UTD Texas Schools Project, a project conceived of and directed by John Kain. The Texas Schools Microdata Panel, or TSMP, currently contains extensive data for five entire cohorts of Texas students as they age through school. The TSMP tracks elementary students as they progress through grades; it measures student performance each spring; and it contains detailed information about their school services. For each cohort there are more than 200,000 students in more than 3,000 public schools. The substantial numbers of students from each school and the ample numbers of students who change special education status are especially important for the methodology pursued here.

---

[8]Expanded special education could also lead to actual reductions in money available for regular education, although there is little evidence that this has happened very frequently.

[9]Such increased relative importance of special education is just the finding of Lankford and Wyckoff (1996) in their analysis of budgetary changes for New York State in the early 1990s. In their analysis, as overall growth in budgets slowed, special education consumes a greater than proportionate share of increases. The extreme in New York State is New York City, where the fiscal absorption of special education is magnified both by rapidly growing spending per special education student and slow growth in the district's overall spending per student. This channeling of funds toward special education could add to voter's apparent discontent with spending growth in the 1990s.

The student data contain a limited number of student, family and program characteristics including race, ethnicity, gender, eligibility for a free or reduced price lunch and special education status, but the panel feature can be exploited to account implicitly for time invariant individual and school effects on achievement. Students who switch public schools within the state of Texas can be followed just as students who remain in the same school or district.

Beginning in 1993, the Texas Assessment of Academic Skills (TAAS) was administered each year to eligible students enrolled in grades three through eight.[10] The criteria referenced tests evaluate student mastery of grade-specific subject matter. Unique IDs link the student records with the test data. We use test results for reading and mathematics, subjects that are examined in all grades. Reading and math tests each contain approximately 50 questions. Because the number of questions and average percent right varies across time and grades, we transform all test results into standardized scores with a mean of zero and variance equal to one. The regression results are robust to a number of transformations including the raw percentage correct.

The normalization of the test scores implies that the estimated models give information about relative gains across students. It is not possible to observe any overall increases or decreases in the level of student performance.

The student IDs also link the student records with a separate special education module. These data contain information on disability type and academic setting. Special education students are served in a number of settings, ranging from mainstreaming (assistance while in the regular classroom) to separate schools, though the majority of students are served in resource rooms at the regular campus.

[10]Many special education students are exempted from the tests, as are other students for whom the test would not be educationally appropriate. In each year roughly 15 percent of students do not take the tests, either because of an exemption or because of repeated absences on testing days (see Table 3, below). These matters are explicitly considered in the analysis that follows.

The Texas Special Education Population

This analysis focuses on student achievement in elementary schools. Table 1 describes the distribution of Texas public school students by disability type for 4th grade students in 1994 and 1995 and 5th grade students in 1994, 1995 and 1996. More than 15 percent of students are classified as disabled in each grade. Though the overall percentage of students who receive special education services is fairly stable, the composition of those served by special education changes markedly between grades four and five. In 4th grade 2.7 percent of all students (21 percent of special education students) receive therapy for speech impairments, but by 5th grade only 1.8 percent of students are classified as speech impaired. Conversely, the percentage of students classified as learning disabled rises from 8 percent in 4th grade to more than 9 percent (two-thirds of special education students) in 5th grade. This disability category— for which schools exert the most discretion in classification decisions— becomes increasingly important as students age. These two categories plus students classified as emotionally disturbed account for more than 85 percent of all students classified as disabled in grades four and five.

The changes over time in the distribution of disabilities are reflected in the transitions into and out of special education between grades four and five. As seen in Table 2, for many students special education is not a career but a set of varying programs. Over 10 percent of students classified as disabled in 4th grade do not receive special education in the following school year, while 16 percent of students who receive special education in 5th grade do not receive special education in the previous year. As expected from the grade patterns in Table 1, the transitions vary dramatically by disability type. A much higher percentage of students classified as learning disabled or emotionally disturbed enter special education in the 5th grade than exit following the 4th grade, while just the opposite is true for students classified as speech impaired. The subsequent analysis will exploit the transitions into and out of special education in order to identify the programmatic impacts on achievement.

Table 1. Distribution of 4th & 5th Grade Students by Disability

	Grade 4[a]		Grade 5[b]	
	% of all students	% of all identified disabilities	% of all students	% of all identified disabilities
Not classified as special education	84.6	n.a.	84.3	n.a.
Learning Disabled	8.0	60.8	9.1	65.9
Speech Impairment	2.7	20.6	1.8	12.9
Emotionally disturbed	0.8	6.4	1.0	7.6
Mentally retardation	0.7	5.1	0.8	5.9
Other physical impairment	0.5	4.0	0.6	4.5
Orthopedic impairment	0.2	1.2	0.2	1.2
Auditory impairment	0.1	1.0	0.1	1.0
Visual impairment	0.1	0.5	0.1	0.5
Autism	0.1	0.5	0.1	0.5
Deaf blind	0.0	0.0	0.0	0.0
Traumatic brain injury	0.0	0.0	0.0	0.0
Unknown disability	2.3	n.a.	1.9	n.a.
All disabilities	15.4	100.0	15.7	100.0
Observations	579,716		856,980	

Notes: a. 1994 and 1995 school years.
b. 1994, 1995, and 1996 school years.

n.a. — not applicable

Table 2. Transition Rates into and out of Special Education for 4th and 5th Grade Students, for selected disability types

	% of Special Education Students in 5th Grade not in Special Education in 4th Grade	% of Special Education Students in 4th Grade not in Special Education in 5th Grade
All Disabilities	15.9%	10.6%
Learning Disabled	9.0%	4.4%
Emotionally Disturbed	12.4%	3.9%
Speech Impaired	11.8%	37.5%

Special education is designed to meet a variety of student needs in addition to cognitive achievement, and the objectives clearly depend on the type of handicapping condition of the student. Nonetheless, an important element of the general set of programs is providing extra services that would enable handicapped students to compete with other students. Indeed, the National Academy panel studying special education concentrates much of its attention on the concept that special education students should be included in the standards for the entire school system (McDonnell, McLaughlin, and Morison 1997).[11]

Our analysis concentrates entirely on cognitive test outcomes of performance for both special education and regular education students, and thus the nature of test taking is relevant. Unfortunately, Texas public schools do not test the majority of students classified as disabled. Table 3 displays rates of test taking for 4th and 5th grade students by disability. More than 80 percent of students without identified disabilities take the exams, but only about 30 percent of those with disabilities do so. Moreover, there is substantial variation by type of disability. Less than 30 percent of learning disabled and emotionally disturbed children take math and reading tests, while roughly three-quarters of the speech impaired students complete the tests. Special education students are excused from the test if the Individualized Education Program (IEP) devised for each child reports that these tests are not an appropriate measurement instrument given the student's current situation. Undoubtedly there is substantial variation across schools in the willingness to excuse students from the tests, and a portion of this may involve strategic considerations by school personnel.

The selective nature of test taking introduces two questions. First, if schools employ systematic patterns of selective test administration, the results for the tested population could be biased. Below we

---

[11]See also Olson and Goldstein (1997) on initiatives to provide more testing of special education students in the National Assessment of Educational Progress (NAEP) and other U.S. government sponsored educational data bases.

Table 3. Percentage of 4th & 5th Grade Students With Valid Test Data by Disability

	Grade 4		Grade 5	
	Math	Reading	Math	Reading
Not classified as special education	81.9	81.7	84.9	84.7
Speech Impairment	75.6	74.9	77.5	77.0
Visual impairment	38.0	37.1	38.9	39.3
Other physical impairment	27.8	27.6	31.4	31.9
Auditory impairment	27.0	23.3	27.2	23.8
Orthopedic impairment	25.6	26.5	27.4	29.1
Learning Disabled	24.1	17.7	26.7	21.7
Emotionally disturbed	22.4	21.3	27.0	26.5
Autism	5.6	6.2	6.0	6.2
Mentally retardation	0.2	0.3	0.3	0.5
Deaf blind	0.0	0.0	0.7	0.0
Traumatic brain injury	0.0	0.0	18.0	12.0
Unknown disability	24.3	21.6	27.7	26.5
All disabilities	30.6	26.6	30.1	27.0

Note: Bilingual students are excluded.

examine the sensitivity of the results to school test taking criteria, particularly for the students who transition into special education. Second, some question arises as to whether the results obtained from the tested population are generalizable to all students who receive special education. The currently available data are insufficient to address this latter issue.

Empirical Model

The value-added framework, which is today the "baseline model" for the examination of student performance, provides the starting point for the empirical analysis of the effects of special education on achievement. The value added model conditions current achievement on a prior measure of achievement and on intervening inputs. This formulation, which we apply in a simple difference form for standardized achievement scores (i.e., outcomes are measured as the difference in scores between grade t and t-1), eliminates any fixed individual differences in the level of achievement. This specification has been used extensively because it effectively accounts for the entire history of school and family inputs that affect the level of achievement in grade t-1. It also handles variations in ability to the extent that they affect levels of performance. It does not, however, deal with any conditions that might affect the rate of learning gain — a serious matter in the consideration of various disabilities which might make the acquisition of new knowledge more difficult. Thus, for this analysis we take the model one step further.

Equation (1) describes the standard value-added formulation: Test score gain (ΔA_{igst}) for student i in grade g in school s in year t is modeled as a function of special education status (SP), vectors of family characteristics (X), school demographic characteristics (D), and community type dummy variables (C) and three error components: a time invariant individual component (γ_i), a school quality component that varies across grades (δ_{gst}),[12] and a random error (ε_{igst}). The family characteristics include information on race, ethnicity, gender, and whether the student is eligible to receive a free or reduced price lunch.

$$(1) \qquad \Delta A_{igst} = SP_{igst}\lambda + X_i\beta + D_{gst}\theta + C_s\eta + \gamma_i + \delta_{gst} + \varepsilon_{igst}$$

Equation (1) can be estimated by OLS using cross-sectional data on student achievement, in which case the coefficient λ captures the average difference in test score gains between special education and nonspecial education students, controlling for observed family, school and community characteristics. Interpreting λ as the causal impact of special education requires that none of the error components are correlated with both the probability of classification as disabled and the gain in achievement. Because selection into special education is related to unobservable school and student characteristics, this assumption is likely to be violated, even though the use of a value added framework undoubtedly reduces the endogeneity bias by accounting for fixed differences in the level of achievement.

Rather than using regular education students as a control group, the panel data can be used to identify special education effects by comparing achievement gains while receiving special education to gains while not in the program. In other words, the comparison group for the estimation of program effects is composed of special education students during grades in which they are not classified as special education. The individual fixed effect estimator can be written as deviations (symbolized by the dot) from each student's mean of all variables, as in Equation (2). In this formulation, all time

---

[12]Rivkin, Hanushek and Kain (1997) document substantial differences in school quality between grades.

invariant individual, family and community factors that might contaminate the estimates of special education program effects are eliminated. In addition, the samples are restricted in some specifications to students who do not switch schools between 4th and 5th grades, effectively ruling out the possibility that changes in school quality that accompany a change in special education status biases the coefficients.

$$\Delta A_{igst} = SP_{gst}\lambda + D_{gst}\theta + \delta_{gst} + \varepsilon_{igst} \qquad (2)$$

The advantages of eliminating fixed individual effects come at a cost, because the effects of special education are estimated entirely on the basis of students who transition into or out of special education during the periods of observation. That is, anybody who is always in regular education or always in special education will have the stable component of school programs captured in the individual fixed effect. Fortunately, the very large samples include a substantial number of students who transition into or out of special education.

The removal of individual fixed effects and school movement almost certainly reduces contamination due to the nonrandom selection into special education. Nevertheless, the possibility remains that the fixed effect estimate of λ confounds the true effects of special education with unobservable differences that change over time. The most significant potential problem involves various aspects of endogenous selection into special education. In the empirical section we examine the robustness of the results using several different approaches to address these problems.

Analysis of Special Education Students

The estimation begins with analysis of the overall effects of special education programs along with variations across type of disability and program setting. We examine test score gains in 4th and

5th grades.[13] The focus on 4th and 5th grade reflects the large samples available for these grades and the fact that a majority of schools in Texas teach these grades at the same campus, limiting the added complications of school movement.

The 4th grade sample includes two cohorts, while the 5th grade sample includes three cohorts of students. The estimation samples contain from 254,294 students (fixed effects models restricted to students who do not switch schools) to 601,526 students (three 5th grade cohorts). The basic estimation begins with all students in the sample years for whom there is valid test data and information about special education status. Concerns about measurement error, which is amplified in the fixed effect form of estimation, led us to exclude the bottom one percent of test scores (roughly students who scored lower than random guessing) and the top and bottom one percent of test score gains from the regressions. The coefficients, however, remain largely unchanged by these deletions.

Following the basic estimation, we turn to a variety of specification checks designed to evaluate the importance of any potential biases introduced by the endogenous selection of students into and out of special education.

Basic Results

Special education program effects are estimated for all special education students combined and separately for learning disabled, emotionally disturbed and speech impaired students.[14] These three categories were chosen for several reasons. First, they are the three largest disability categories; the many other categories have extremely small sample sizes and very few students who transition in and out, making detection of any effects very difficult. Second, these are the disabilities over which the

[13]Gains are measured by the difference between the current and previous year test scores. Test score gain in 4th grade equals 4th grade test score minus 3rd grade test score and 5th grade gain equals 5th grade score minus 4th grade score.

[14]When separate specifications are estimated for each disability type, all students in other disability categories are excluded from the analytical samples.

schools exert the largest degree of discretion, and therefore from a policy point of view a decision to expand or contract special education largely refers to a decision to expand or contract these categories of disabilities. Third, we expect special education to have its largest achievement impact on learning disabled students and a much smaller impact on students classified as speech impaired. Therefore, comparisons of program effects across disability categories provide some information on the validity of the estimation strategies.

The standard for comparison is ordinary least squares estimates of special education program effects based upon the simple cross-sectional model of equation (1), which compares special education students to regular education students. OLS results for math and reading test scores by grade and disability type are presented in Table 4. In addition to the indicator for special education program participation, each regression includes dummy variables for cohort, whether the student is Black, Hispanic or Asian, eligible for a subsidized lunch, and community type (categorized as suburban, large urban, small urban and rural) along with the percentages of students in the school who are Black, Hispanic, and eligible for a free or reduced price lunch.

The OLS estimates show no consistent pattern of achievement gains across grades, tests and disability types. Controlling for observed differences, special education students in 4th grade have significantly lower math score gains but not significantly lower reading score gains. (Note that there are substantial differences in the average level of achievement between regular and special education students, even if the differences in gains are small). Separate estimates for the learning disabled and emotionally disturbed retain this pattern, which is not surprising in the case of the learning disabled because a substantial percentage of special education students are classified as learning disabled. In contrast, the 5th grade math and reading test score gains are not significantly different for special education students, with the exception of students classified as learning disabled whose gains exceed those of the nonspecial education students. This pattern may reflect the different selection mechanisms

Table 4. Effects of Special Education on 4th & 5th Grade Test Score Gains, by Type of Disability
(Huber-White adjusted t statistics in brackets)

	Math		Reading	
	4th	5th	4th	5th
All disabilities	-0.06	0.00	0.01	0.01
	[-9.43]	[0.16]	[0.97]	[2.19]
Learning disabled	-0.10	0.00	0.04	0.02
	[-10.8]	[0.05]	[4.48]	[2.95]
Emotionally disturbed	-0.08	-0.02	-0.06	0.02
	[-2.59]	[-1.12]	[-2.10]	[1.28]
Speech impaired	0.00	0.00	0.01	0.00
	[0.33]	[0.15]	[0.98]	[0.57]

Note: Each coefficient comes from a separate ordinary least squares regression that includes indicator variables for black, Hispanic, female, eligibility for free or reduced price lunch, cohort, and community type (three separate) plus % black, % Hispanic, and % eligible for subsidized lunch for the school. The estimates for each disability type come from separate regressions comparing just the identified disability to all regular education students.

at work in 4th and 5th grade, or actual differences in program effectiveness. As we noted earlier, these OLS coefficients almost surely confound the true program effect with individual and school influences linked with selection into the special education program.

To isolate the true program effect, we estimate models with individual fixed effects for all special education students (Table 5) and by disability type (Table 6). Samples are pooled across grades. Because the impact of any characteristics that do not change between 4th and 5th grade is captured by the fixed effect, the special education coefficient is identified by students whose special education status changes between the two grades. The first and third columns report regressions from samples that include all students (including students who do not change programs).[15] The second and fourth columns restrict the sample to students who remain in the same school for both grades in order to ensure that special education effects are not confused with any systematic differences across schools. The impact of special education is also permitted to differ by the timing of the transition between special and non-special education. We expect the special education effect to be larger for students who enter special education in grade g than for students who exit special education following grade g-1. Those who exit may have gained the skills needed to perform in regular classrooms, or they may not have benefitted from the special program. In either case, the special education/non-special education differential in these students' gains would understate the average program effect.

The fixed effect estimates show unequivocally that special education increases test score gains for math and reading. The coefficients are quite similar when the sample is restricted to students who remain in the same school for both 4th and 5th grade, suggesting that confounding factors related to the act of switching schools introduce little if any bias. The coefficients are all significant at conventional levels, though the effects on mathematics achievement are roughly twice as large as those for reading

---

[15]Students whose special education status does not change are included in order to more precisely estimate cohort differences in test scores and the effects of school demographics.

Table 5. Estimated Effects of Special Education on 4th & 5th Grade Test Score Gains, Controlling for Individual Fixed Effects and School Mobility
(Huber-White adjusted t statistics in brackets, number of transitions in parentheses)

	Math		Reading	
Total special education	0.07 [6.11] (9,503)	0.09 [6.33] (6,871)	0.03 [2.14] (9,167)	0.04 [2.65] (4,072)
Entrants into special education	0.17 [8.72] (4,439)	0.17 [7.66] (3,308)	0.07 [3.42] (4,205)	0.08 [3.35] (3,127)
Exiters from special education	0.00 [0.27] (5,064)	0.02 [1.17] (3,563)	0.01 [0.48] (4,962)	0.01 [0.31] (3,517)
Individual fixed effects	yes	yes	yes	yes
School nonmovers only	no	yes	no	yes

Note: Each estimated coefficient comes from an individual fixed effect regression that includes an indicator variable for cohort plus % black, % Hispanic, and % subsidized lunch for the school. The estimates for entrants and exiters come from a single regression that allows for separate effects by transition timing. The specifications estimated for nonmovers come from samples that include only students who stay at the same campus for 4th and 5th grades.

Table 6. Estimates of Special Education on 4th & 5th Grade Math and Reading Test Score Gains by Disability Type, Controlling for Individual Fixed Effects and School Mobility
(Huber-White adjusted t statistics in brackets, number of transitions in parentheses)

	Math			Reading		
	All	Entrants	Exiters	All	Entrants	Exiters
Learning Disabled	0.14	0.18	0.06	0.04	0.04	0.02
	[4.35]	[4.53]	[1.30]	[0.96]	[0.91]	[0.39]
	(1,433)	(885)	(548)	(1,276)	(783)	(493)
Emotionally Disturbed	0.12	0.13	0.06	0.04	0.04	0.03
	[1.36]	[1.36]	[0.29]	[0.46]	[0.43]	[0.16]
	(216)	(176)	(40)	(211)	(173)	(38)
Speech Impaired	0.01	0.02	0.01	0.03	0.00	0.04
	[0.62]	[0.46]	[0.46]	[1.78]	[0.00]	[1.97]
	(3,418)	(525)	(2,893)	(3,412)	(517)	(2,895)
Individual fixed effects	yes	yes	yes	yes	yes	yes
School nonmovers only	yes	yes	yes	yes	yes	yes

Note: Separate regressions are estimated for each disability type. Each estimated coefficient comes from an individual fixed effect regression that includes an indicator variable for cohort plus % black, % Hispanic, and % subsidized lunch for the school. The samples include only students who stay at the same campus for 4th and 5th grades.

(an effect of .09 standard deviations versus .04). The differential impact on math is consistent with studies of overall school quality using this same body of data (cf. Rivkin, Hanushek, and Kain 1998).[16]

Table 5 also shows that allowing for separate effects by transition timing produces the expected result: Students who enter special education in 5th grade appear to benefit far more than those who exit after 4th grade. The coefficient for entrants increases to 0.17 in math and 0.08 in reading, and all coefficients are significant at the 1 percent level. In contrast, none of the special education effects for those who exit are statistically significant at any conventional level, consistent with both positive and zero effects of special education for those students, an issue to which we return below. (Note: exits are coded such that a positive coefficient for exits indicates that special education is beneficial).

An alternative explanation for the differential between entrants and exits is that special education has a much smaller effect on speech impaired students (the majority of those who exit) than on the learning disabled (the largest share of entrants). Table 6 reports the estimates from parallel models for students classified as learning disabled, emotionally disturbed and speech impaired respectively, with the restriction that all students with other disabilities in either 4th or 5th grade are excluded from each sample. Again the results are much stronger for math achievement gains, while the coefficients from the reading gains specifications are very noisily measured. For those classified as learning disabled, the pattern of special education effects on math achievement is almost identical to that for all special education students, including the much larger estimated effects for entrants. Quite similar estimates are also found for students classified as emotionally disturbed, though the fact that only 216 such students transition into or out of special education likely accounts for the large standard errors.

---

[16]The larger impact of schools on math than reading for regular education could be explained by parents' having greater ability to help children in reading. It is less obvious that schools should have a larger impact in math for special education students, since anecdotal evidence suggests that reading problems are often central to evaluation and classification for special education. Reading programs also appear plentiful in special education.

In contrast, the effects of special education on math achievement are very small and statistically insignificant for students classified as speech impaired. There is little reason a priori to believe that special education should exert a substantial impact on math achievement for students with speech impairments. Thus, the finding that special education raises math achievement for students classified as learning disabled but not for those classified as speech impaired provides support for the belief that the positive association between achievement and special education captures a causal relationship.

Specification Checks

Procedures for placing students into and out of special education pose the most significant potential interpretative questions about the previous results. Consider the probability that a non-special education student in grade g-1 enters special education in grade g:[17]

$$SP^{*}_{gst} = \Delta A_{igst-1}\pi + D_{gst}\psi + \zeta_{gst} + \upsilon_{igst} \quad (3)$$

where SP* indexes the propensity to enter the special education program (with entrance into special education if SP* > 0 and remaining in regular education otherwise). The propensity to enter special education is written as a function of the prior year's achievement gain, student demographics, a school effect that may vary by grade and time (ζ_{gst}) and a random error which captures any transitive changes in student performance or behavior.[18]

The fixed effects estimation of the achievement models removes any time invariant elements of achievement gains, demographics, or unmeasured influences. Therefore, only time varying influences

[17]An analysis of the decision to exit special education is identical.

[18]Equation 3 can easily be written to include both prior gains and the level of achievement. In fact, in the empirical work below we look at such a formulation.

on selection propensities that are also related to student achievement gains can contaminate the estimates. Perhaps the most significant concern would be that students are placed into special education based in part on temporarily poor performance, due to a variety of things ranging, say, from having a bad test day to new problems at home (i.e., $\pi < 0$ in eq. 3 and the error in A_{t-1} is also negative). We estimate the effects of special education programs for entrants in grade g by contrasting their achievement gains in that grade with their achievement gains prior to entering special education (i.e., with performance in grade g-1). Any recovery in grade g from the temporary downturn in grade g-1 will be manifested as a larger than usual achievement gain in grade g. A temporarily low score on 4^{th} grade tests could lead, for example, to placement in special education and to a high gain in achievement in the 5^{th} grade relative to the 4^{th} grade—simply because a low 4^{th} grade score both reduces the 4^{th} grade gain and increases the 5^{th} grade gain. If the temporary achievement problem is self-correcting (because of a better test taking day, because of a correction in outside problems with family or friends, or whatever), the previous fixed effects gain regressions might make it appear that the special education program led to the improved gains when in fact it did not.[19]

To consider this possibility, we first examine the link between entry into special education and prior achievement. Appendix Table A1 reports coefficients for 3^{rd} and 4^{th} grade test scores in linear probability models of entry into special education in 5^{th} grade. These estimates employ school fixed effects and are provided separately for all disabilities and for learning disabled.[20] While these models incorporate measured demographic differences among students, they do not incorporate individual fixed

---

[19]Biases in the opposite direction are also possible. If temporary factors during grade g – such as increasingly disruptive behavior in class or a reduced ability to concentrate – both reduce current achievement and increase the probability of being put into special education during the year, a downward bias of the estimate of λ would result. We are most concerned about upward biases that would erroneously make special education programs appear effective and thus concentrate our attention on those possibilities.

[20]In order to eliminate any between school differences in special education selection, samples are restricted to students remaining at the same campus in grades g and g-1. The linear probability models include the student characteristics, community type dummy variables and school demographic characteristics as regressors.

effects or the long-run patterns of performance. The results show that both lower 3rd and 4th grade performance raise the probability of entering special education in 5th grade. For math, 4th grade performance has a stronger influence on special education placement than 3rd grade performance, but the opposite is the case for reading. Overall, the results indicate that students who perform poorly for a number of years are more likely to enter special education than those with a single low test score. Nonetheless, these results are also consistent with the possibility that a transitory decline in performance that reduces 4th grade test scores raises the probability of classification, confounding program effects with the recovery from a temporary shock.

To consider potential placement biases directly, we introduce achievement gain information for the 6th grade for the one cohort with observed gains in 4th through 6th grades. Specifically, we throw out information on 5th grade gains and contrast 6th grade gains with 4th grade gains— all within a fixed effects framework.[21] Importantly, we exclude all transitions into or out of special education except for students who are not classified as special education in 4th and 5th grade but enter special education in 6th grade. This roughly cuts in half any bias caused by a temporary downturn in achievement that leads to classification: It remains true that the 6th grade gain would be overstated by a temporarily low 5th grade score, but the 4th grade gain (the comparison year) would not be affected. If such bias is driving the results, we would expect the estimates of special education program effects to be roughly half as large as those reported in the previous tables.

The interrupted panel estimation provides no support for the hypothesis that temporary score changes operate through selection to yield the positive special education effects. Table 7 displays the coefficients for all special education students and for students classified as learning disabled. The

[21]Because only a single cohort can be used and many students switch schools in 6th grade, we include school switchers. This choice is consistent with the minimal differences in results reported in Table 5. Specifications that excluded switchers generated similar, but much less precisely estimated effects.

Table 7. Interrupted Panel Estimates of Special Education Effects: 4th & 6th Grade Test Score Gains

(Huber-White adjusted t statistics in brackets, number of transitions in parentheses)

	All Disabilities		Learning Disabled	
	Math	Reading	Math	Reading
Entrant into special education in 6th (Not in special education in 4th and 5th)	0.18 [5.46] (1,213)	0.16 [4.67] (1,204)	0.32 [5.40] (311)	0.18 [2.70] (315)
Individual fixed effects	yes	yes	yes	yes
School nonmovers only	no	no	no	no

Note: Each coefficient comes from an individual fixed effect regression that includes an indicator variable for cohort plus % black, % Hispanic, and % subsidized lunch for the school. The estimates for learning disabled come from separate regressions excluding all students ever classified with other disabilities. Coefficients come from samples that exclude all students who change special education status unless they enter in grade 6.

coefficient from the math achievement regression that includes all special education students is virtually identical to the estimate in Table 5, while the coefficient from the math achievement regression that excludes all disability types other than learning disabled is slightly larger than that reported in Table 6. In addition, the reading coefficients are both slightly larger than those produced by the sample of students in grades four and five. Consistent with the selection regressions, there is little or no evidence that the prior estimates confounded the effects of special education with student recovery from temporary negative shocks. If anything, the negative relationship between achievement two years prior to entry and the probability of entry may introduce a downward bias.

A second variant of selection effects is the possibility that schools manipulate which students take the tests. Specifically, since test taking is relatively low in special education, schools might use testing strategically. With increased attention to testing and accountability, schools could actively intervene in the selection of students who take the tests, excluding those they expect to perform badly. Because, however, special education effects estimated here are identified by the individual student difference in test score gains inside and outside of special education, bias is introduced only by very special kinds of selection. Schools would have to exclude systematically students whom they expected to gain the least from special education in comparison to their gains in regular education. While such manipulation is possible, it seems unlikely that many schools would focus on this select group of students, particularly since the state does not monitor achievement gains for special education programs. Nevertheless, we do examine the possibility that nonrandom test taking contaminates the results.

To assess the importance of test selection, we repeat the fixed effects regressions for the 4^{th} and 5^{th} gains but include only schools for which 100 percent of the students who took tests while in regular education programs also took tests in the special education programs. Slightly more than half of the transitioning students are excluded from the math achievement regressions, and roughly 40 percent are excluded from the reading achievement regressions. Nonetheless, as seen in Appendix Table A2, these results yield the same conclusions as those based on the full sample: Special education programs improve student performance. Even though the samples are reduced, the point estimates for the full and the restricted estimates are qualitatively the same.

A third specific source of bias that we examine is the possibility that differences in school quality across grades confound the estimated program effects. This bias could come through direct selection relationships with quality aspects of schools or through other grade and school factors.[22] Rather than identifying program effects solely by comparisons of gains in special education with gains out of special education, we now compare the school average 5th minus 4th grade gain differential for students who enter special education in 5th grade with the school average 5th minus 4th grade differential for students not classified as disabled in either grade. In other words, in estimating the effects of special education we control for patterns of between-grade differences in a school's quality by using information from performance of regular education students in the school.[23] The results reported in Appendix Table A3 again show no evidence of bias: Estimates of the impact of special education programs are virtually identical to the fixed effect estimates reported in Tables 5 and 6 that did not control for grade differences in quality.

Table 7 and Appendix Tables A1-A3 provide little evidence that the estimated special education effects are driven by temporary shocks to achievement, manipulative test taking, or changes in school

---

[22]If, say, better teachers are more likely to seek special assistance for their students and students spend at least part of the day with the regular classroom teacher, an upward bias will be introduced by a positive correlation between the transition probability and unobserved school quality. Conversely, if less skilled teachers are more likely to encourage classification of students, a downward biased might be introduced. These appear as effects of ζ_{gst} in equation 3, which would also enter the achievement models.

[23]Specifically, we include the two school average gain differentials as two observations for each school and cohort in a school fixed effects regression of the difference in gains on an indicators for cohort and whether the gain differential refers to students who entered special education in 5th grade. These coefficients can be thought of as difference-in-difference-in-difference estimates of special education effects.

quality. It remains a possibility nonetheless that still other, unknown factors might contaminate the estimates. If we could identify valid instrumental variables that were related to the probability of special education classification but otherwise unrelated to achievement gains, we could investigate the existence of such contamination. While it is generally difficult to identify valid instruments in situations of joint decision making such as schools, changes in Texas school financing formulae raised the possibility that such instruments could be found. As Cullen (1997) describes in detail, Texas altered the formula determining state educational aid a number of times during the 1990s, including altering the additional dollars received for classifying a student as disabled. More importantly for our purposes, the magnitude of the change in state revenue depended upon district wealth, meaning that there was substantial between district variation in the change over time in the fiscal incentive to classify students as disabled. Following Cullen, we use the predicted change in state aid from classifying an additional student as disabled as an instrument for special education, assuming that within district changes over time in fiscal incentives are unlikely to be directly related to achievement, particular controlling for student fixed effects.[24]

The first stage of the instrumental estimates reported in Appendix Table A4 show that the probability of receiving special education is positively related to the predicted revenue gain. In contrast to Cullen's work which also found fiscal effects on placement with district level data, however, the coefficient for the sample of math test-takers is only marginally significant at the 5% level and the coefficient for the sample of reading test takers is not significant at any conventional level.

---

[24]As Cullen points out, it is important to use the predicted change in state aid because the actual change in aid depends upon the type of disability, instructional setting, district tax effort and current enrollment in special education, all which could be correlated with school quality. We did experiment with the weights applied to the various disability types in determining the predicted revenue increase, but the results were not sensitive to the choice of weights. It is important to note that even the predicted revenue change could be related to other determinants of achievement if other changes in state educational policy during this period varied systematically by district size and wealth.

Consequently, it comes as no surprise that the instrumental variable estimates are quite noisy and uninformative for both math and reading. (For example, the IV point estimate in the preferred fixed effect formulation suggests that special education boosts reading achievement by an incredible eight standard deviations).

The specification tests as a whole provide no evidence that the estimated program effects confound the true special education effects with other factors. While it remains possible as in virtually all empirical work that the estimates are contaminated by other influences, we have no reason to believe that the robust results here reflect anything but the causal impacts of special education programs.

Targeting and Dynamics

The pattern of estimates raises several questions concerning the ability or willingness of schools and families to target services at students who receive the greatest benefits. Larger effects for entrants than for students who exit special education is consistent with the notion that schools target services where they are most effective, but it is certainly far from definitive evidence. We now use the cohort with test score gains for 4th, 5th and 6th grade to investigate further the nature of benefits of special education programs.

The first two rows of Table 8 provide information on the question of why the test score gains of those who exit special education are similar to their gains while in the program. The coefficients in the first row are identified by comparisons of 5^{th} and 4^{th} grade gains for students who enter special education in 5th grade but do not remain for the 6th grade. (In this and the remaining entries of the table, all transitions into or out of special education other that the specified transition are excluded from the sample). In the second row, the same group of students is considered but their performance in regular education (i.e., 4^{th} and 6^{th} grades) is compared instead of their 4^{th} and 5^{th} grade performance. If special education helped these students so much that they no longer needed such assistance, both the

Table 8. Estimates of Special Education Effects for Differing Patterns of Program Participation
(Huber-White adjusted t statistics in brackets, number of transitions in parentheses)

	Grades included	All disabilities		Learning disabled	
		Math	Reading	Math	Reading
Entrant into spec. ed. in 5th grade (not in special education in 6^{th})	4 & 5	-0.04 [-0.74] (410)	-0.01 [-0.15] (409)	0.21 [1.16] (37)	0.16 [0.82] (39)
Entrant into spec. ed. in 5th grade (not in special education in 6^{th})	4 & 6	0.08 [1.49] (410)	0.07 [1.25] (409)	-0.02 [-0.13] (37)	0.08 [0.45] (39)
Entrant into spec. ed. in 5th grade (in special education in 6^{th})	4 & 5	0.23 [5.94] (1,031)	0.07 [1.62] (951)	0.17 [2.51] (313)	0.15 [1.97] (266)
Entrant into spec. ed. in 5th grade (in special education in 6^{th})	4 & 6	0.24 [7.05] (1,031)	0.15 [4.12] (951)	0.17 [2.71] (313)	0.20 [2.78] (266)
Individual fixed effects		yes	yes	yes	yes
School nonmovers only		no	no	no	no

Note: Each coefficient comes from an individual fixed effect regression that includes an indicator variable for cohort plus % black, % Hispanic, and % subsidized lunch for the school. The estimates for learning disabled come from separate regressions excluding all students ever classified with other disabilities. Coefficients come from samples that exclude all students who change special education status unless they enter in grade 5 and meet the specified grade 6 criterion.

5th and 6th grade gains should be significantly larger than the 4th grade gains. On the other hand, if these students were dropped from special education because it had little or no positive effect, one would expect little or no difference between their performance prior to and following the special intervention. By examining both the 5th and 6th grade gains, we also investigate the possibility that the benefits of special education may not be realized during the year of treatment.

The estimates provide little support for the notion that achievement rises for students who spend only one year in special education. None of the coefficients in the top two rows are statistically significant, though the small number of transitions among the learning disabled undoubtedly reduces the precision of the estimates. From this evidence, it appears that those who exit special education were not benefitting from the intervention, at least in terms of higher academic achievement.

The final two rows in Table 8 provide a preliminary look at the dynamics of special education effects for students who remain in the program for at least two years. The coefficients in the third row are identified by comparisons of 5th and 4th grade gains for students who enter special education in 5th grade and remain in special education in 6th grade; the coefficients in fourth row are identified by comparisons of 6th and 4th grade gains for these same students. If there are diminishing returns to special education, we would expect the 6th minus 4th grade differential to be smaller than the 5th minus 4th grade differential.

The results show otherwise: the math coefficients are virtually identical for both the learning disabled and all disability specifications, while the reading coefficients are if anything slightly higher in the bottom row. There is no evidence of diminishing returns, indicating that students who remain in special education continue to receive substantial benefits. In combination with the findings for students who exit special education, this evidence suggests that schools target services toward students who receive the highest benefits.

Program Setting

Much of the programmatic debate about special education has focused on the issue of mainstreaming. The original federal legislation called for providing special education within the least restrictive environment (see, for example, Martin, Martin, and Terman 1996). It also called for providing an education appropriate to each child. These goals could clearly conflict, but there has been steady pressure to "mainstream" special education students by including them in the regular classroom setting to every extent possible. In Texas, the pressure to mainstream has been incorporated into school finance legislation, and the revenue gain from having an additional mainstreamed special education student rose dramatically in 1995. At the same time, the use of mainstreaming appears to be an element of conflict with parents of students in regular education who are worried that special education students may detract from the education of their children.

While the objectives of mainstreaming go far beyond achievement gains, its impact on achievement is nevertheless important. Table 9 reports OLS and fixed effect estimates of special education by mainstream status for both all special education students and those classified as learning disabled. The fixed effects coefficients are identified by students who switch program type or special education status.

The fixed effect estimates again provide the best information about the effects of special education. For math, where the impacts of special education programs appear generally larger, the results show little difference in special education effects by treatment setting, with point estimates very close to the estimates that did not differentiate by setting. For reading, the estimates for mainstreamed students are insignificant, while achievement effects for nonmainstreamed students are significantly positive and close in magnitude to the estimates ignoring setting. Less than 15 percent of special education students are mainstreamed even following the increase in fiscal incentives, so the sample

Table 9. Effects of Special Education on 4th & 5th Grade Test Score Gains, by Setting
(Huber-White adjusted t statistics in brackets)

	Math			Reading		
	4th	5th	4th & 5th	4th	5th	4th & 5th
All Disabilities						
mainstreamed	-0.04	0.01	0.10	-0.01	0	-0.04
	[-2.08]	[0.63]	[3.55]	[-0.33]	[0.37]	[-1.16]
not mainstreamed	-0.06	0.00	0.09	0.01	0.01	0.05
	[-9.39]	[0.01]	[6.06]	[1.06]	[2.34]	[3.08]
Learning Disabilities						
mainstreamed	-0.06	0.021	0.160	0.02	0.02	-0.03
	[-2.60]	[1.70]	[4.03]	[0.86]	[1.16]	[-0.80]
not mainstreamed	-0.11	0.00	0.13	0.05	0.02	0.07
	[-10.8]	[0.66]	[3.84]	[4.56]	[2.80]	[1.72]
Individual fixed effects	no	no	yes	no	no	yes
School nonmovers only	yes	yes	yes	yes	yes	yes

Note: Each coefficient comes from an individual fixed effect regression that includes an indicator variable for cohort plus % black, % Hispanic, and % subsidized lunch for the school. The estimates for learning disabled come from separate regressions excluding all students ever classified with other disabilities. The estimates by setting come from a single regression that allows for separate effects for students mainstreamed and not mainstreamed. The samples include only students who stay at the same campus for 4th and 5th grades.

sizes of students who transition into or out of mainstreamed special education programs are fairly small.

One difficulty in interpreting these results is that students are not randomly selected into program settings, so this is not the type of clean experiment that would be produced if students were randomly assigned to different settings. Student fixed effects remove time invariant characteristics, but any changes in student well-being or knowledge gained from observing students in different settings could potentially contaminate the estimated impact of mainstreaming. A possible solution is the use of instrumental variables, but the weak first stage explanatory power of fiscal incentives for mainstreaming again made instrumental variable estimates uninformative.

A second problem for interpretation is that information on setting comes from student records, but the definitions of what is and is not mainstreaming are likely to be problematic. For example, students treated in resource rooms may spend a majority of the day in regular classes, so there is no clear distinction between mainstreaming and nonmainstreaming in terms of exposure to non-special education class activities. In the extreme, the fiscal incentive to mainstream may lead schools to relabel settings without altering substantially the nature of the intervention. Without being able to control effectively for pupil placement, the results are consistent either with setting not making a difference for student performance or with school officials being very effective at designing programs of study for each individual student.

The Effects of Special Education on Regular Education Students

The remaining component of this analysis is to consider the effects of special education programs on regular education students. The most systematic investigation of this is Cullen (1997). She provides a detailed analysis of how the expansion of special education affects the funds available for regular education students and their achievement in Texas using instrumental variable techniques.

Her results show that increases in special education reduce regular education funding and achievement, consistent with the beliefs of many parents and educators concerned about the recent expansion of special education.

In contrast to Cullen who focuses solely on the fiscal impact of special education, we ask a more general question: Do changes in the proportion of students classified as disabled affect the achievement of nonspecial education students? Such changes may affect regular education students in a myriad of ways including changes in the composition of classes, in the emphasis or focus of teachers, or in available resources. These estimates, which expand on Rivkin, Hanushek, and Kain (1998) and which follow in the general structure of the previous analysis, consider how the proportion of students classified as special education in a given school and grade affect achievement gains for regular education students. As before, we estimate these models with a series of explicit controls for individual student, student body composition, and community type factors, with student fixed effects, and with instrumental variables based on fiscal incentives. We also include information on teacher experience and average class size of regular classrooms to control for other changes in school characteristics that might coincide with changes in special education enrollment. Since special education enrollment might affect regular class sizes and teacher experience, we also estimated specifications that exclude these school characteristics. The sample used in the fixed effect specification is restricted to students not classified as special education in either grade.[25]

For reasons developed previously, we believe that the most reliable estimates come from the models that include student fixed effects.[26] The basic fixed effect estimates in Table 10 results show

---

[25]A small number of schools are excluded from the analysis if data on school characteristics is missing or measured with error. See Rivkin, Hanushek and Kain (1998) for a description of which schools are excluded.

[26]The effects of special education are identified here by differences across grades in the proportion of students classified as special education. The OLS estimates (not reported) showed no consistent pattern.

Table 10. Estimated Effects of the Proportion Special Education Students on 4th and 5th Grade Test Score Gains for Students Not Receiving Special Education
(Huber-White adjusted t statistics in brackets)

	Math	Reading
Total special education		
Proportion special education	0.44	0.36
	[2.97]	[3.25]
Proportions by setting		
Mainstreamed	0.40	0.39
	[1.37]	[1.86]
Not mainstreamed	0.44	0.36
	[2.97]	[3.24]
Proportions by disability type		
Learning disabilities	0.32	0.30
	[1.67]	[2.07]
Other disabilities	0.46	0.44
	[2.62]	[3.32]
Speech impaired	-0.10	-0.13
	[-0.30]	[-0.51]
Individual fixed effects	yes	yes
School nonmovers only	yes	yes

Note: Each coefficient comes from an individual fixed effect regression of test score gains for regular education students (in the appropriate subject) on the proportion of students in the grade in the specific categories and on an indicator variable for cohort plus % black, % Hispanic, and % eligible for subsidized lunch for the school. Three separate regressions are estimated for each subject: the first aggregates all special education students into a single group; the second distinguishes among special education students by setting; and the third divides special education students by disability type.

that an increase in the proportion of students classified as disabled *raises* achievement for students not classified as disabled. The estimated parameters indicate that a 10 percentage point increase in the percentage of students classified as disabled increases achievement roughly 0.04 standard deviations for both reading and math. (A change of one special education student in a class of 20 students would be a five percent change. The standard deviation in our sample of the percentage classified as disabled is 6 percent). While similar selection mechanisms to those discussed above could also contaminate these estimates, the prior specification tests failed to find evidence of contamination in the analysis of special education students. Further, movements of special education students on the margin constitute only a small percentage of regular education students. Combined, these strongly suggest that the regular education estimates are not driven by factors other than those directly related to the special education classification rate.

As noted above, a number of factors either singly or as a group could account for the link between achievement and the special education classification rate. We now take several steps to try to gain some understanding of the transmission mechanism. First, we exclude the teacher experience and average class size variables from the regressions. If special education affected achievement through fiscal influences on the quantity of resources devoted to regular education, the exclusion of the teacher experience and class size variables should alter the coefficient on percent classified as disabled. The results (not reported) show that the exclusion of these variables has virtually no effect on the coefficients, which provides preliminary evidence that resources are not driving the link between achievement and special education classification rates.[27]

---

[27]These estimates are identified by the relationship between within school changes in achievement and special education classification rates. It is quite possible that expansion of special education affects the quantity of resources devoted to regular education throughout the state, and such statewide impacts would not be uncovered in our analysis.

We repeated this experiment for the specifications that divide students by disability type and obtained similar results. We also found little or no evidence that changes in the pupil/aide ratio affect achievement.

We next separate special education students into two categories: 1) the proportion of all students in mainstreamed special education programs; and 2) the proportion of all students in nonmainstreamed special education programs. Table 10 shows that the setting makes no difference to the achievement gains of regular education students, as the coefficients are virtually identical. This suggests that the physical inclusion of special education students does not appear to harm regular education students and that something other than the setting is the important factor.

Finally, we separate special education students into three categories: those classified as either learning disabled or emotionally disturbed; those classified as speech impaired; and the remaining disability types including mental retardation and a variety of physical disabilities. This serves as an informal specification check in addition to providing information on the causal pathways, because it is unlikely that students attending roughly one half hour per week of speech therapy should have a major impact on either the learning environment or the resources devoted to regular classrooms. The fixed effect estimates in Table 10 confirm this hypothesis, as the coefficient for proportion speech impaired is not significant at any conventional level for either reading or math achievement despite the fact that changes in the proportion speech impaired account for much of the change in classification rates between 4th and 5th grade. In contrast, the proportions with physical disabilities and learning or emotional difficulties are both positively related to achievement, though the estimates are somewhat larger and more precisely estimated for the proportion physically disabled. Because programs and services differ significantly across these disabilities, the similar magnitudes of the coefficients for these two categories do not point to a specific interpretation of the relationship to the achievement of regular education students.

The pattern of estimates in Tables 10 suggests a causal link between achievement for nonspecial education students and the special education classification rate, but the precise underlying causes are not identified. Special education programs likely reduce class size[28] for the regular education students — either because of pull-outs or because of other resources devoted to the special education students that allow for substitution of the regular teachers time, and they may also permit teachers to increase the pace or difficulty of class.[29] Limited evidence suggests that pure resource effects play a secondary role to changes in the classroom environment. Notice, however, that we use grade specific classification rates, while categorical aid and fiscal transfers are determined at the district level. Thus, any overall resource impacts are likely to be weakly related to changes in classroom resources for a specific grade and school, and it is not surprising that we do not identify any resource driven effects.

Conclusions

For good reason, previous discussions of special education have concentrated on issues of costs. It is well documented that providing high quality schooling for students with various disabilities is more expensive than that for regular education. Yet the focus on costs has often obscured the fact that there is educational purpose in the programmatic designs of special education, and the benefits to special education students may well justify the costs.

This paper concentrates on identifying the effects of special education programs on achievement. The unique data for entire cohorts of Texas elementary students permit detailed investigations of how special education impacts both special education students and nonspecial education students. Specifically, the repeated performance measures allow us to identify program

[28]The class size measure comes from average the number of students in the classroom as reported by just the regular education teachers in a grade. These averages should include the special education students assigned to the class, although they are probably subject to error.

[29]The estimates for regular education performance were also approached from an instrumental variables strategy, but the instruments provided little explanatory power in the first stage predictions of special education proportions and generated quite noisy IV estimates.

effects by contrasting the achievement gains of students who experience both special education and regular education. These estimates, which fully allow for any persistent individual handicapping conditions and individual ability differences, indicate that special education programs on average have a significantly beneficial effect on performance. One year in special education programs boosts the average math score by at least 0.09 standard deviations and the average reading score by at least 0.04 standard deviations over what would be expected in regular education classes.[30] There is further evidence that schools target services toward those who benefit most.

The special education effects are estimated for single years in the program. One set of estimates for both 5th and 6th grade participation in special education indicates that the positive effects do cumulate. Nonetheless, our current estimates do not provide sufficient evidence about the time path of achievement that can be expected from longer participation in special education programs.

At the same time, similar estimation of achievement growth by students in regular education provides no evidence that higher rates of special education classification detract from their performance. Achievement appears to be positively related to special education classification rates. This evidence leads to a much more benign view of special education than is typically found, though it is important to recognize that our analysis ignores any negative impacts of special education common to all schools in Texas, such as reduced state aid for regular education. Nevertheless, the evidence we have suggests that special education programs on average (as currently operated and funded) benefit both special education and nonspecial education students.

The overall effects of special education are also quite robust to heterogeneity by disability type and by setting of the special education programs. By replicating the achievement analysis of the

---

[30]All of our estimates concentrate on gains in relative student performance. There are substantial differences in average levels of performance between regular and special education students. Additionally, since test scores are normalized to have mean zero and variance one in each year, any overall performance gains or losses across the state cannot be identified.

effectiveness of special education programs for just those identified as learning disabled, we could isolate a large but more homogeneous group of students and (presumably) programmatic features. This analysis yielded qualitatively similar results as that for all students. On the other hand, the achievement effects of special education programs for speech impairments were, as expected, minor and insignificant for both special education students and for regular education students.

A second aspect of disaggregating the results involved a preliminary look at instructional setting. Each of the prior analyses was redone to allow for differences between mainstreamed and nonmainstreamed students. This analysis found very similar effects across instructional setting—quite different from much of the general discussions of the educational effects of mainstreaming on both special and regular education students. Most important, there is no evidence that mainstreaming adversely affects nonspecial education students.

Finally, an elaborate series of specification analyses was conducted. Most critically, we allow for temporary achievement effects that might influence both placement in special education and specific patterns of achievement gains. We also allow for potential strategic behavior by school officials in selecting students who were eligible for taking the state achievement tests. Neither influences the estimates.

On the methodological side, this analysis demonstrates the value of repeated panel data on large sample sizes. All prior analyses have been unable to follow gains of individual students across time in ways that would permit examination of programmatic effects. Furthermore, because movements into and out of special education— particularly when disaggregated by disability— remain relatively rare phenomena, very large samples are required to obtain clean estimates of program effects.

This analysis concentrates on average program effects, only minimally disaggregated by disability type and setting. Our other work (Rivkin, Hanushek, and Kain 1998), however, shows dramatic differences in achievement across teachers of regular education students. It is reasonable to believe this holds for special education programs and teachers also. At the very least, many special education students spend significant time with the regular education teachers in regular classrooms, and this would be expected to have powerful effects on the achievement of special education students. In many ways, it is remarkable that the estimated average effects are as strong as they are. More analysis is nonetheless needed to investigate the heterogeneity of performance across teachers, programs, and schools.

None of this analysis has considered costs, though without doubt, special education involves added costs. In addition, the analysis has also concentrated exclusively on issues of academic achievement. Special education programs typically have many goals in addition to raising achievement. Finally, only a third of the special education students take the regular tests. These issues suggest that further analysis is required to understand both the generalizability of these results to the entire special education population and the larger impact of these programs outside of the achievement realm. There are positive effects of special education (on achievement), but much work remains to be done before we can draw comprehensive conclusions about the costs and benefits of special education.

Selected References

Chaikind, Stephen, Louis C. Danielson, and Marsha L. Brauen. "What do we know about the costs of special education? A selected review." *Journal of Special Education* 26, no. 4 (1993): 344-70.

Cullen, Julie Berry. "Essays on special education finance and interogvernmental relations." Ph.D. Dissertation, Massachusetts Insitute of Technology, 1997.

Ferguson, Ronald. "Paying for public education: New evidence on how and why money matters." *Harvard Journal on Legislation* 28, no. 2 (Summer 1991): 465-98.

Hanushek, Eric A., and Steven G. Rivkin. "Understanding the twentieth-century growth in U.S. school spending." *Journal of Human Resources* 32, no. 1 (Winter 1997): 35-68.

Hanushek, Eric A., and with others. *Making schools work: Improving performance and controlling costs*. Washington, DC: Brookings Institution, 1994.

Hartman, William T. "Policy effects of special education funding formulas." *Journal of Education Finance* 6 (Fall 1980): 135-59.

Kain, John F., and Kraig Singleton. "Equality of Educational Opportunity revisited." *New England Economic Review* (May/June 1996): 87-111.

Lankford, Hamilton, and James Wyckoff. "Where has the money gone? An analysis of school district spending in New York, 1979-80 to 1991-92." In *Putting Children First*. Vol. II. Reports Prepared for the New York State Special Commission on Educational Structure, Policies and Practices. Albany, NY: New York State Commission on Educational Structure, Policies and Practices, 1993.

---. "The allocation of resources to special education and regular instruction." In *Holding schools accountable: Performance-based reform in education*, edited by Helen F. Ladd, 221-57. Washington, DC: Brookings, 1996.

Lewit, Eugene M., and Linda Schuurmann Baker. "Child indicators: Children in special education." *The Future of Children* 6, no. 1 (Spring 1996): 139-51.

Martin, Edwin W., Reed Martin, and Donna L. Terman. "The legislative and litigation history of special education." *The Future of Children* 6, no. 1 (Spring 1996): 25-39.

McDonnell, Lorraine M., Margaret J. McLaughlin, and Patricia Morison, eds. *Educating one and all: Students with disabilities and standards-based reform.* Washington, DC: National Academy Press, 1997.

Monk, David H. *Educational finance: An economic approach.* New York: McGraw-Hill, 1990.

Olson, John F., and Arnold A. Goldstein. *The inclusion of students with disabilities and limited English proficient students in large-scale assessments: A summary of recent progress.* Washington, DC: National Center for Educaiton Statistics, 1997.

Parrish, Thomas B., and Jay G. Chambers. "Financing special education." *The Future of Children* 6, no. 1 (Spring 1996): 121-38.

Reschly, Daniel J. "Identification and assessment of students with disabilities." *The Future of Children* 6, no. 1 (Spring 1996): 40-53.

Rivkin, Steven G., Eric A. Hanushek, and John F. Kain. Teachers, schools, and academic achievement. Paper presented at the annual meetings of the Econometric Society, Chicago, January 1998.

Rothstein, Richard, and Karen Hawley Miles. *Where's the money gone? Changes in the level and composition of education spending.* Washington, DC: Economic Policy Institute, 1995.

Sack, Joetta L. "Spec. ed. designation varies widely across country." *Education Week* 17 (24 June 1998): 1.

Singer, Judith D., and John A. Butler. "The education for all handicapped children act: Schools as agents of social reform." *Harvard Educational Review* 57, no. 2 (May 1987): 125-52.

Wagner, Mary M., and Jose Blackorby. "Transition from high school to work or college: How special education students fare." *The Future of Children* 6, no. 1 (Spring 1996): 103-20.

Appendix Table A1. Estimated relationship between prior test scores and the probability of entering special education in 5th Grade (Huber-White adjusted t statistics in brackets)

	All Disabilities			Learning Disabled		
3rd grade math	-0.0037		-0.001	-0.0014		-0.0001
	[-8.22]		[-2.23]	[-5.73]		[-0.61]
4th grade math	-0.012		-0.0066	-0.004		-0.0019
	[-22.3]		[-12.3]	[-14.1]		[-7.00]
3rd grade reading		-0.0078	-0.0062		-0.0031	-0.0026
		[-14.8]	[-11.3]		[-10.1]	[-8.31]
4th grade reading		-0.0078	-0.0034		-0.0023	-0.001
		[-15.9]	[-6.60]		[-8.62]	[-3.58]
School fixed effects	yes	yes	yes	yes	yes	yes
School nonmovers only	yes	yes	yes	yes	yes	yes

Note: Each coefficient comes from a separate school fixed effect regression. The dependent variable equals one if the student was classified as special education in the 5th grade but not in the 4th grade and equals zero otherwise. In addition to the prior individual achievement test scores, the explanatory variables include indicator variables for black, Hispanic, female, eligibility for free or reduced price lunch, cohort, and community type (three separate) plus % black, % Hispanic, and % eligible for subsidized lunch for the school. The estimates for learning disabled come from separate regressions excluding all students ever classified with other disabilities. The samples include only students who stay at the same campus for 4th and 5th grades, and exclude all students who were classified as disabled in 4th grade.

Appendix Table A2. Estimates of Special Education on 4th & 5th Grade Test Score Gains, Controlling for Individual Fixed Effects and School Mobility and Excluding Schools with Missing Student Test Score Data (Huber-White adjusted t statistics in brackets, number of transitions in parentheses)

	All Disabilities		Learning Disabled	
	Math	Reading	Math	Reading
Total special education	0.09	0.03	0.14	0.04
	[5.23]	[1.44]	[4.04]	[1.08]
	(4,330)	(4,072)	(1,149)	(1,012)
Entrants into special education	0.16	0.08	0.18	0.08
	[5.82]	[2.56]	[4.29]	[1.38]
	(2,073)	(1,865)	(700)	(609)
Exits from special education	0.03	0.01	0.06	0.00
	[1.42]	[0.59]	[1.16]	[0.04]
	(2,257)	(2,207)	(449)	(403)
Individual fixed effects	yes	yes	yes	yes
School nonmovers only	yes	yes	yes	yes

Note: Each coefficient comes from an individual fixed effect regression that includes an indicator variable for cohort plus % black, % Hispanic, and % subsidized lunch for the school. The estimates for learning disabled come from separate regressions excluding all students ever classified with other disabilities. The estimates for entrants and exiters come from a single regression that allows for separate effects by transition timing. All employ samples restricted only to students who stay at the same campus for 4th and 5th grades and are in schools in which 100 percent of exiters or entrants into special education who took tests as regular education students also take tests as special education students.

Appendix Table A3. Estimates of Effects of Special Education on 4th & 5th Grade Test Score Gains, Controlling for Individual Fixed Effects and Changes School Quality with Regular Student Gains (Huber-White adjusted t statistics in brackets, number of transitions in parentheses)

	All Disabilities		Learning Disabled	
	Math	Reading	Math	Reading
Special Education Program	0.16	0.06	0.18	0.03
	[5.86]	[2.05]	[3.13]	[0.36]
Individual fixed effects	yes	yes	yes	yes
School nonmovers only	yes	yes	yes	yes

Note: These coefficients come from separate school fixed effect regressions using data aggregated by school and special education transition type. There are two observations for each school: the average difference between 5th grade and 4th grade achievement gains for students who enter special education in 5th grade; and the average difference in achievement gains for students not classified as disabled in either grade. The school average difference in gains is regressed on indicator variables for whether the students entered special education in 5th grade and cohort in a model that removes school fixed effects. The samples include only students who stay at the same campus for 4^{th} and 5^{th} grades, and exclude all students who were classified as disabled in 4^{th} grade

Appendix Table A4. Instrumental Variable Estimates of Special Education on 4th and 5th Grade Test Score Gains along with Coefficients for First Stage Regressions of Fiscal Instruments on Special Education Classification (Huber-White adjusted t statistics in brackets)

	Math	Reading
All disabilities	0.0 [0.00]	7.9 [1.11]
1st Stage Estimate (coefficient on excluded instrument)	0.002 [2.00]	0.001 [1.50]
Individual fixed effects	yes	yes
School nonmovers only	yes	yes

Note: Instrumental variables estimates of the effects of special education programs on performance with individual fixed effects. Instrument is the predicted revenue gain ($1,000s) from classifying an additional student as disabled).