INFERRING PROGRAM EFFECTS FOR SPECIAL POPULATIONS: DOES SPECIAL EDUCATION RAISE ACHIEVEMENT FOR STUDENTS WITH DISABILITIES?

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

Abstract—Most discussion of special education has centered on the costs of providing mandated programs for children with disabilities and not on their effectiveness. As in many other policy areas, inferring program effectiveness is difficult because students not in special education do not provide a good comparison group. By following students who move in and out of targeted programs, however, we are able to identify program effectiveness from changes over time in individual performance. We find that the average special education program significantly boosts mathematics achievement of special-education students, particularly those classified as learning-disabled or emotionally disturbed, while not detracting from regular-education students. These results are estimated quite precisely from models of students and school-by-grade-by-year fixed effects in achievement gains, and they are robust to a series of specification tests.

I. Introduction

One of the most discussed but least analyzed issues in education today is special education. Although a disproportionate amount of school funding goes to the education of handicapped children—perhaps as much as one-fifth of total current spending for slightly more than 10% of students—extraordinarily little evidence has accumulated about the effectiveness of special-education programs in raising achievement. Moreover, while evidence for New York and Texas suggests that special education programs may have crowded out regular-education spending, little evidence of the impact on achievement for non-special-education students exists.[1]

The paucity of evidence stems in part from the difficulty of isolating the causal effects of special education. A comparison of special-education and non-special-education students does not provide a valid measure of program effectiveness, because special-education students by definition differ in some significant respects, implying that achievement differences confound program effects with other factors (cf. Wagner and Blackorby, 1996). Similarly, the correlation between achievement for non-special-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 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 non-special-education students.

This type of problem is not specific to special education, but occurs in a variety of circumstances where programs are developed to serve special populations. For example, in considering programs to benefit the long-term unemployed, chronically ill patients, or teenage mothers, people outside of these groups generally provide poor comparison populations. Moreover, specific programs are often embedded within a set of other time-varying treatments. The methods developed in this paper transfer directly to a variety of such circumstances.

We are able to identify special-education effects by exploiting longitudinal information on individual students included in the UTD Texas Schools Project, which follows several entire cohorts of Texas elementary school students across grades. The large number of special-education students in this data set permits detailed investigations of the effects of special-education placement on student achievement, controlling for fixed student and school effects. Comparisons of academic performance before and after placement into special education provide much better evidence of program effects than existing cross-sectional data. Similarly, the identification of program effects on regular-classroom achievement with changes over time in the percentage of a school's students classified as special-education is superior to identification based on cross-sectional differences in the percentage classified as special-education.

Even controlling for all time-invariant unobserved differences among students and schools with the fixed-effects approach, changes in special-education status that are accompanied by other changes in students or schools still present a potential problem. If, for example, a deterioration in skills or school quality accompanies classification as disabled, fixed-effects models will tend to underestimate the impact of special education. If, conversely, a transitory downturn in prior-year achievement raises the probability of classification as disabled, these models will tend to

Received for publication July 8, 1998. Revision accepted for publication July 10, 2001.

* Stanford University, National Bureau of Economic Research, and University of Texas at Dallas; University of Texas at Dallas; and Amherst College and University of Texas at Dallas, respectively.

The analysis in this paper has been supported by grants from the Smith Richardson Foundation, the Packard Humanities Institute, 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 developed for special education in Texas. We also thank seminar participants at the University of California, Santa Cruz; Texas A&M University; and 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. Lankford and Wykoff (1996) document changes in expenditures on regular and special education in New York public schools during the 1980s. In the single existing study of the achievement effects on non-special-education students, Cullen (1997) finds evidence that fiscal crowding out harms these students, using aggregate school and district-level data.

The Review of Economics and Statistics, November 2002, 84(4): 584–599

overestimate the impact of special education by attributing the recovery from a temporary negative shock to the program. We address the issue of endogeneity bias in a number of ways, including the use of school-by-grade-by-year fixed effects and comparisons of achievement in nonadjacent years. We also investigate the possibility that manipulation of the test-taking population contaminates the results.

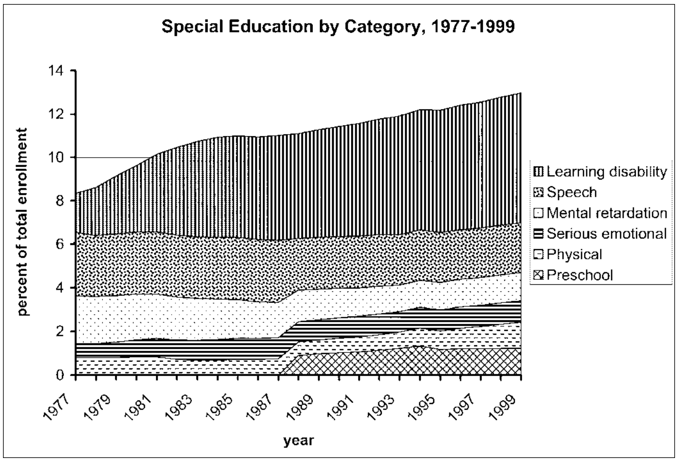
Following a baseline analysis of average special-education effects for students in all settings, we investigate the possibility that specific types of special-education programs produce systematically different outcomes. Of particular interest is the effect of mainstreaming on achievement. Texas law emphasizes the importance of educating students in the least restrictive environment, and the state has reinforced this policy through the use of fiscal incentives. The increase in mainstreaming may also affect non-special-education students by altering the student composition and resources in regular classrooms.

The primary results are straightforward. Special-education programs on average boost the achievement of students provided this special treatment, and it appears that schools target services toward students who derive larger benefits. 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. More surprisingly, achievement gains for students who do not receive special education are positively related to the percentage of students classified as special education, and there is little or no evidence that mainstreaming systematically harms non-special-education students. Whether it is the additional funding 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 do not constitute a comprehensive cost-benefit analysis of special education. They apply just to the special-education population taking the standardized Texas tests, and the findings may not generalize to more severely disabled students who are not tested. Moreover, neither costs nor other important outcomes are considered, though it should be noted that an important element of the general set of programs is the provision of extra services that would enable handicapped students to compete with other students.[2] Nevertheless, the evidence provides a convincing case that the special-education programs on average provide the intended benefits without reducing achievement for the non-special-education population.

[2] The National Academy panel studying special education concentrates much of its attention on the idea that special education students should be included in the standards for the entire school system (McDonnell, McLaughlin, & Morison, 1997). 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 databases.

FIGURE 1.—SPECIAL EDUCATION BY CATEGORY, 1977–1999



II. Background

The Individuals with Disabilities Education Act (IDEA), enacted in 1975,[3] translated concerns about the education of children with both physical and mental disabilities into federal law. This act prescribed a series of diagnostics, counseling activities, and services for disabled students. Although 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 & 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 needs. The existence of partial categorical funding from the state and federal governments and of intensive instruction for students creates incentives both for school systems to expand the population of special-education students and 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

[3] This act, P.L. 94-142, was originally the Education for All Handicapped Children Act and was retitled IDEA in 1990. It is commonly described 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).

[4] The financing of special education differs significantly across states and localities. In overall terms, approximately 8% of special-education funding is federal, some 56% comes directly from states, and the remainder is local. These shares are approximately equal to the shares of total elementary and secondary spending. Moreover, as with total funding of public schools, wide variation in the state funding formulas 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).

1999 in the population identified as disabled.[5] Despite the fact that overall public-school enrollment remained roughly constant over this period, the number of students classified as disabled increased from 3.7 million in 1977 to 6.1 million in 1999, causing the percentage of students classified as disabled to increase from 8.3% to 13%. Virtually all of the growth came from increases in students classified as learning-disabled, which grew from 22% to 46% of all disabled students over this period and from less than 2% to 6% of the total school population.[6] This category encompasses a continuum of learning conditions where it is difficult to describe and to apply precise cutoffs in evaluation and assessment. This discretion also leads to considerable variation in classification rates across states, districts, and time (Reschly, 1996; Lewit and Baker, 1996). The more clearly defined physical disabilities represent less than 10% of special-education students.

The expansion of special-education services and consequent expenditure increases have raised concerns about adverse impacts on resources and school quality for non-special-education students. Hanushek and Rivkin (1997) finds that special education accounted for roughly 20% of the increase in per student spending during the 1980s, slightly less than double the share of special-education students.[7] Thus, special education had a disproportionate effect, although certainly less than the overwhelming impact that some have suggested.

The fiscal impact of special education can rise significantly in times of fiscal stringency. Because of the legally 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. 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.[8]

III. The Texas Schools Microdata Panel

The cornerstone of this research is the analysis of a unique matched panel data set of school operations constructed by the UTD Texas Schools Project, a project conceived of and directed by John Kain. The data track the universe of three successive cohorts of Texas public elementary school students as they progress through school, beginning in 1993. Students who switch public schools within the state of Texas can be followed along with students who remain in the same school or district. For each cohort there are over 200,000 students in over 3,000 public schools. The substantial numbers of students from each school and the large number who change special-education status are especially important for the methodology pursued here. We use data for grades four through seven for the two older cohorts, and grades three through six for the youngest cohort, yielding three grades of achievement gains for each cohort. The youngest cohort attended fifth grade in 1996, while the oldest cohort attended fifth grade in 1994. The main regression sample includes 767,763 students and a total of 1,876,915 annual observations.

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 allow implicitly for time-invariant individual and school effects on achievement. Both special-education program effects and the effects of special-education program size on the achievement of regular classroom students are identified by changes over time in either special-education status or the percentage of students classified as special education. This methodology effectively controls for all time-invariant student and school-by-grade effects on achievement gains as well as school-by-grade effects that vary from year to year.

Beginning in 1993, the Texas Assessment of Academic Skills (TAAS) was administered each spring to eligible students enrolled in grades three through eight.[9] The criteria-referenced tests evaluate student mastery of grade-specific

[5] Data on special education come 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. See, however, Rothstein and Miles (1995).

[6] Note that students aged 3–5 in preschool programs appear to have increased 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, whereas those students were spread across categories before 1987–1988, in that year and after they were reported separately.

[7] These calculations 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 education spending. 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 in expenditure related to special education was clearly larger during the 1970s. 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.

[8] 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 by slow growth in the district's overall spending per student. This channeling of funds toward special education could add to voters' apparent discontent with spending growth in the 1990s.

[9] 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 somewhat more than 15% of students do not take the tests, either because of an exemption or because of repeated absences on testing days (see Table 3). These matters are explicitly considered in the analysis that follows.

TABLE 1.—DISTRIBUTION OF SPECIAL-EDUCATION STUDENTS IN TEXAS BY DISABILITY TYPE AND GRADE LEVEL

	Distribution by Type (%)			
Disability Type	Grade 4	Grade 5	Grade 6	Grade 7
Learning-disabled	60.6	66.0	69.6	70.6
Speech impairment	20.4	12.9	7.2	4.0
Emotionally disturbed	6.4	7.6	8.6	9.6
Mentally retardation	5.0	5.9	6.2	7.3
Other physical impairment	4.4	4.5	5.3	5.6
Orthopedic impairment	1.2	1.2	1.0	1.0
Auditory impairment	1.0	1.0	1.0	1.0
Visual impairment	0.4	0.5	0.5	0.5
Autism	0.6	0.5	0.4	0.4
Deaf and blind	0.0	0.0	0.0	0.0
Traumatic brain injury	0.0	0.0	0.1	0.1
	100.0	100.0	100.0	100.0
Special Education Rate	15.5	15.8	15.4	14.9
Observations	293,344	858,343	870,105	586,013

Grades are combined across different sample years; sample sizes differ because of the differing number of cohorts available at each grade, as shown in table 2.

TABLE 2.—PERCENTAGE OF STUDENTS CLASSIFIED AS DISABLED IN TEXAS, BY GRADE LEVEL AND YEAR

	Classification Rate (%)			
Year	Grade 4	Grade 5	Grade 6	Grade 7
1994		13.5		
1995	14.5	14.8	13.6	
1996		15.2	14.6	13.4
1997			14.5	13.9

subject matter. Unique IDs link the student records with the test data. We use test results for mathematics; results for reading are qualitatively similar though somewhat smaller in magnitude.[10] The tests each contain approximately fifty questions. Because the number of questions and average percentage of right answers 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 use of the raw percentage correct.

The student IDs also permit linking the student records with separate special-education 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 on the regular campus.

IV. The Texas Special Education Population

Table 1 describes the distribution of Texas public school students by disability type for students in grades four through seven. About 15% of students are classified as disabled in each grade, though the composition of those served by special education changes markedly as students age. The percentage of special-education students who receive therapy for speech impairments falls from 20% in fourth grade to 4% in seventh grade. Conversely, the percentage classified as learning-disabled—a disability category for which schools exert the most discretion in classification decisions—rises from 61% in fourth grade to 71% in seventh grade. These two categories, plus students classified as emotionally disturbed, account for over 80% of all students classified as disabled in grades 4 through 7. Most of the remaining students classified as disabled suffer from well-defined physical or mental disabilities whose shares of special-education enrollment remain fairly constant as students age.

Since the higher grades in Table 1 are observed in later years, a portion of the trend in classification rates represents general changes over time. Table 2 shows some increase in participation across years, but disaggregated data (not presented) confirm that the trends for specific disabilities described in Table 1 are quite similar for all three cohorts.

The subsequent analysis relies heavily on transitions into and out of special education, and, as seen in the top three rows of Table 3, for many students special education is not a career but a set of varying programs. Over 10% of students classified as disabled in fourth grade do not receive special education in the following school year, and 15% of students who receive special education in fifth grade have not received special education services in the previous year. The transition rates between grades five and six and between grades six and seven are similar, though the figures suggest that for these grades entrants are declining over time as a percentage of the total.

As expected from the grade patterns in Table 1, the transitions vary dramatically by disability type. A much higher percentage of students classified as speech-impaired exit than enter special education following the fourth, fifth, and sixth grades, whereas entrants and exiters constitute much more similar and much smaller shares for those classified as learning-disabled or emotionally disturbed. Nevertheless, large numbers of students enter or exit special education every year in each of the three largest disability categories.

Importantly, a substantial proportion of special-education students, including some who transition into and out of special education, do not complete the standardized tests. As seen in Table 4, slightly over 80% of students without identified disabilities have valid gain scores in a grade, in comparison with about 30% of those with disabilities.[11] Substantial variation exists by disability type: Gain scores

[10] Lyon and Fletcher (2001) and Lyon et al. (2001) argue that there are potential structural reasons for the lower responsiveness of reading scores than math scores. Their evidence suggests that reading problems must be addressed earlier than fourth grade in order to have much success. They also point to inadequate early diagnosis of reading problems and the subsequent confusion of learning disabilities and preventable reading difficulties.

[11] Valid gain scores require the completion of tests in consecutive grades. We emphasize the availability of gain scores because that is central to the subsequent empirical analysis.

TABLE 3.—TRANSITION RATES INTO AND OUT OF SPECIAL-EDUCATION PROGRAMS FOR SELECTED DISABILITY TYPES

	Transition Rates of Special-Education Students (%)					
	Grade 4 to 5		Grade 5 to 6		Grade 6 to 7	
Years	Entrants	Exiters	Entrants	Exiters	Entrants	Exiters
All Disabilities						
1994–1995	n.a.	n.a.	13.6	10.6	n.a.	n.a.
1995–1996	14.9	10.8	10.6	10.0	10.7	9.1
1996–1997	n.a.	n.a.	9.2	10.6	8.2	9.9
Learning-Disabled						
1994–1995	n.a.	n.a.	6.9	5.2	n.a.	n.a.
1995–1996	8.1	4.4	4.4	4.5	4.3	5.4
1996–1997	n.a.	n.a.	4.1	5.2	3.7	5.3
Emotionally Disturbed						
1994–1995	n.a.	n.a.	11.7	5.1	n.a.	n.a.
1995–1996	14.2	4.3	8.8	4.9	9.2	4.8
1996–1997	n.a.	n.a.	7.0	13.0	8.6	11.8
Speech-Impaired						
1994–1995	n.a.	n.a.	10.2	45.5	n.a.	n.a.
1995–1996	11.5	40.3	9.3	49.4	10.2	49.5
1996–1997	n.a.	n.a.	7.6	45.4	7.8	45.2

can be computed for less than 30% of learning-disabled and emotionally disturbed children, but roughly three-quarters of speech-impaired students. Special-education students are excused from the test if their Individualized Education Program (IEP) indicates that these tests are not an appropriate measurement instrument for them. 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 issues. First, if schools employ systematic patterns of selective test administration, the results for the tested population could be biased. Below we examine the sensitivity of the results to school test-taking criteria, particularly for the students who transition into special education. Second, some question arises 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.

TABLE 4.—PERCENTAGE OF TEXAS STUDENTS WITH VALID TEST DATA BY DISABILITY AND GRADE LEVEL

	Valid Tests (%)	
Disability Type	Grade 4	Grade 7
Speech impairment	74.8	77.2
Visual impairment	40.1	51.8
Auditory impairment	29.5	33.6
Other physical impairment	28.1	44.2
Orthopedic impairment	26.7	35.1
Unknown disability	23.2	25.4
Learning-disabled	24.2	39.3
Emotionally disturbed	22.6	36.5
Autism	5.5	12.6
Mentally retardation	0.2	1.2
Deaf and blind	0.0	0.0
Traumatic brain injury	0.0	35.0
All disabilities	30.3	35.6
Not classified as special education	81.5	81.9

Bilingual students are excluded.

V. Empirical Model

Analyses of school attributes and programs typically begin with a model in which student achievement levels are determined by the cumulative past influences of families and schools along with the student's abilities, motivation, and the like. This conceptual framework is often used for cross-sectional empirical analyses that relate achievement in a grade to family background and the characteristics of schooling in that grade. Such approaches have been heavily criticized, however, for a variety of legitimate reasons: ignoring past family and school inputs, inadequately describing current schooling conditions, ignoring programmatic placement rules, and others. To deal with these issues, models considering the growth in student achievement during a school year—often called value-added models—have been generally preferred. In difference form, past family and school attributes drop out, leading to direct analysis of how the flow of current school and family resources affect achievement growth.[12] Nonetheless, as conventionally analyzed using a single year of annual growth across students,

[12] These models have been estimated in difference form (i.e., $A_t - A_{t-1}$) and with prior achievement on the right-hand side. The choice largely reflects the scale of the achievement measures. The level explanatory variables are the cumulative flows of family and school factors, which in

a variety of issues crucial to special education—such as unobserved differences in student learning ability—remain important. Dealing with these deeper issues is the central focus of this analysis.

A value-added model that examines the growth in achievement during a school year provides the starting point for the empirical analysis. The equation

$$\begin{aligned}\Delta A_{igst} &\equiv A_{igst} - A_{i,g-1,s-1,t-1} \\ &= SE_{igst}\lambda + X_{igst}\beta + D_{gst}\theta + \gamma_i \\ &\quad + \delta_s + \omega_{gst} + \epsilon_{igst}\end{aligned} \tag{1}$$

models achievement gain (that grade's test score minus the score in the prior grade) for student i in grade g and school s at time t as a function of special-education status in that grade (SE), vectors of family characteristics (X) and school demographic characteristics (D), and four error components: an individual factor γ_i, a school fixed effect δ_s, a school-by-grade-by-year fixed effect ω_{gst}, and a random error ϵ_{igst}. The family characteristics include race, ethnicity, and gender, along with indicator variables for students who switch schools and students who are eligible to receive a free or reduced-price lunch.[13] Finally, the vector of school demographic characteristics includes the proportion black, proportion Hispanic, and proportion eligible for a reduced-price lunch.

Using the gain in achievement for a grade rather than the level of achievement as the dependent variable eliminates any fixed individual effects on the level of achievement (for example, differences in preparation for the first grade or past differences in teacher quality). This specification also handles variations in ability to the extent that ability affects the level of performance. Importantly, the explanatory variables must represent annual family and school flows into the education production process in order to correspond with the value-added outcome measure. In the case of the school characteristics the variables reflect special-education participation and peer group characteristics for that grade, the flow of new services and inputs. Family characteristics such as race/ethnicity are included because they may be systematically related to the annual rate of learning.

The ability to estimate the separate parameters of equation (1) and the interpretation of these depend on the data available and the estimation approach. For example, it can be estimated by OLS using cross-sectional data on student achievement gains, in which case the coefficient λ captures the average difference in test score gains between special-education and non-special-education students, controlling for observed family and school characteristics. But a cross-sectional approach cannot separately identify the individual and school components of the composite error term in equation (1). Thus, interpreting λ as the causal impact of special education requires that no error component be correlated with the probability of classification as disabled. Because selection into special education is almost certainly related to unobserved school and student characteristics, this assumption would likely be violated despite the fact that the value-added framework allows for fixed differences in the level of achievement.

We control for much of the confounding variation introduced by systematic but unmeasured differences in students and schools by making use of the matched panel structure, which provides multiple observations of achievement growth for each individual, multiple cohorts at each school, and the ability to follow students across different Texas schools. This data structure enables us to control directly for student, school, and school-by-grade-by-year fixed effects.

With panel data, the special-education effects are identified by the change in achievement gain for students who transition into or out of special education. In an extension, special-education transitions are divided into those entering special education and those who exit in order to investigate the existence of possible differences by transition pattern. Importantly, these latter estimates provide a different perspective on the effects of special education programs by comparing student performance with pretreatment learning growth in the case of entrants and to posttreatment learning growth in the case of exiters. For entrants, $SE_{\text{entry}} = 1$ if the student receives special education in grade g but not in grade $g - 1$, while for exiters, $SE_{\text{exit}} = 1$ if the student receives special education in grade g but not in grade $g + 1$. We estimate the effects for entrants and exiters separately by excluding other types of transitions from each of the samples.

Any differences in schools not perfectly correlated with the student fixed effect or the included covariates but correlated with the probability of classification as disabled would still contaminate the estimates. While most past analyses of schools have pursued an estimation strategy relying on measured attributes of schools (say, characteristics of teachers or spending per pupil), the inability to characterize differences among schools with any precision leaves this open to question. Even in models that remove school fixed effects, any temporal or across grade variation in school quality that is related to the probability of special-education classification would continue to bias the estimates. The addition of school-by-grade-by-year fixed effects, however, controls for such differences over time and among grades, making it highly unlikely that variations in school or teacher quality bias the estimated effects of special education.

difference form become simply the current values of family and school inputs (e.g., current class size or current family size).

[13] Roughly 20% to 25% of transitions into or out of special education are coincident with a school transfer, yielding only a slightly higher mobility rate than that for regular-education students (Cullen and Rivkin, 2003). Moreover, virtually all students in the sample change campuses at least once as they progress from middle to junior high school. Because we control for school-by-grade-by-year effects and for school switching, the disruption and change in school quality following a move do not contaminate the special-education estimates.

The complete fixed-effects models identify the effects of special education as the difference in achievement gains for students who transition into or out of special education and those who retain the same special education classification throughout the period, controlling for grade and for time-varying differences in teacher and school quality.[14] The advantage of eliminating the confounding influences of individual and school heterogeneity does come at a cost, however, because the effects of special education must be estimated entirely on the basis of students who transition into or out of special education during the periods of observation.

The validity of the fixed-effects approach relies on the assumption that the probability of being classified as special education is orthogonal to the remaining error. Though the fixed-effects approach eliminates the main potential sources of bias from family and schools, we conduct a series of specification tests to investigate the possibility that other complications contaminate the estimates.

VI. The Effects of Special-Education Programs on Special-Education Students

We examine annual test score gains in the fourth through seventh grades using three cohorts of students. The cohorts do not cover exactly the same grades: there is only one cohort with fourth-grade gains, whereas there are two cohorts with seventh-grade gains and three cohorts with fifth- and sixth-grade gains. Concern about measurement error, which is amplified in the fixed-effects form of estimation, led us to exclude the bottom 1% of test scores (roughly, students who scored lower than random guessing). The estimated coefficients remain largely unchanged by these deletions.

A. *Basic Results*

Special-education program effects are estimated for all special-education students, combined and separately for the categories of learning-disabled, emotionally disturbed, and speech-impaired. These categories, the three largest, encompass the disabilities where schools exert the largest degree of discretion in selection. From a policy view, a decision to expand or contract special education largely refers to a decision to expand or contract these categories. Additionally, they provide a useful contrast, because we expect special education to have its largest achievement impact on learning-disabled and emotionally disturbed students and a much smaller impact on students classified as speech-impaired. As Table 5 shows, the average achievement of students classified as speech-impaired, in all grades, is at least 0.7 standard deviations higher than the average for those classified as learning-disabled and at least 0.5 standard deviations higher than for those classified as emotionally disturbed.

Table 5.—Average Mathematics Score by Disability Type and Grade

	Normalized Mathematics Score			
	Grade 4	Grade 5	Grade 6	Grade 7
Learning-disabled	−0.84	−0.94	−1.04	−1.07
Speech-impaired	−0.14	−0.19	−0.28	−0.33
Emotionally disturbed	−0.69	−0.81	−0.91	−0.95
All disabilities	−0.61	−0.76	−0.94	−0.99
Not classified as special education	0.06	0.08	0.09	0.09

Test scores for the state are normalized to mean 0 and standard deviation 1 in each year and grade.

Table 6 reports baseline value-added models in addition to specifications that control for fixed effects. The top three rows present estimates for all disability categories combined; the remainder of the table presents estimates for the separate categories. All specifications provide two sets of estimates: one that combines entrants and exiters, and one that reports separate effects by transition type.[15] In addition to the indicator for special-education program participation, each regression includes dummy variables indicating eligibility for a subsidized lunch and a change of school. Specifications other than those removing school-by-grade-by-year fixed effects also include cohort-by-grade dummies and the proportions of students in the school who are black, Hispanic, and eligible for a free or reduced-price lunch, and the non-fixed-effects specifications add race/ethnic and gender dummy variables. The absolute values of Huber-White t-statistics adjusted for the clustering of students into schools are reported for all coefficients.[16]

Table 6 shows that the average effect of special education for all disabilities is positive once student heterogeneity is allowed for with fixed effects. In the absence of fixed effects (column 1), the estimate is 0, suggesting that differences in learning growth between regular- and special-education students exactly offset any special-education effects. The inclusion of student fixed effects, which effectively changes the treatment comparison, raises the coefficient to roughly 0.03. This and the subsequent fixed-effects estimates are statistically significant at conventional levels. The magnitude of the estimate is also relatively invariant to the inclusion of school-by-grade-by-year fixed effects that capture variation in school quality across grades and years (as seen in the final column).[17]

[14] The fixed-effects model controls for fixed differences in the rate of growth. An alternative and potentially more flexible approach is to include all past values of the dependent variable on the right-hand side. Because the test scores are noisy measures of actual achievement in addition to being endogenous variables, their inclusion on the right-hand side would introduce serious problems in the absence of instruments. We do not believe that any valid instruments can be found, and we consider below the possible effects of changes in student circumstances that are related to entry into or exit from special education.

[15] Separate regressions generate the entry and exit coefficients.

[16] The t-statistics are not adjusted in specifications that remove school-by-grade-by-year fixed effects.

[17] Alternative estimation that included just school fixed effects instead of school-by-grade-year fixed effects produced virtually identical point estimates of the special-education effects.

TABLE 6.—ESTIMATED EFFECTS OF SPECIAL-EDUCATION PROGRAMS ON MATHEMATICS TEST SCORE GAINS, BY TYPE OF DISABILITY AND TRANSITION

	Estimated Program Impact						
Transition Type	No Fixed Effects		Student Fixed Effects		Student and School-by-Grade-by-Year Fixed Effects		Observations
			All Disabilities				
Entry and Exit	0.00		0.03		0.04		1,876,915
	(1.16)		(4.17)		(6.76)		
Entry		0.00		0.09		0.08	1,832,799
		(0.31)		(7.47)		(9.99)	
Exit		0.00		−0.02		0.00	1,838,044
		(0.25)		(1.67)		(0.60)	
			Learning-Disabled				
Entry and Exit	0.00		0.04		0.06		1,828,834
	(1.53)		(2.70)		(6.73)		
Entry		0.00		0.07		0.11	1,797,589
		(0.95)		(3.44)		(7.99)	
Exit		0.00		−0.02		0.02	1,801,912
		(0.87)		(1.33)		(1.68)	
			Emotionally Disturbed				
Entry and Exit	−0.01		0.08		0.11		1,770,446
	(1.39)		(2.75)		(4.90)		
Entry		−0.01		0.10		0.15	1,746,037
		(1.39)		(2.32)		(4.70)	
Exit		−0.01		0.04		0.09	1,750,279
		(0.82)		(0.93)		(2.72)	
			Speech-Impaired				
Entry and Exit	0.01		0.00		−0.01		1,783,787
	(3.05)		(0.06)		(0.86)		
Entry		0.01		0.01		−0.02	1,750,516
		(1.66)		(0.52)		(1.43)	
Exit		0.01		−0.01		−0.02	1,761,260
		(2.16)		(1.14)		(1.89)	

Absolute values of Huber-White adjusted t-statistics in parentheses. Separate regressions are estimated for each disability type, and special-education students with other disabilities are excluded from these specification. In addition to the indicator for special-education program participation, each regression includes dummy variables, indicating cohort and grade, eligibility for a subsidized lunch, and a change of school, as well as the proportions of students in the school who are black, Hispanic, and eligible for a free or reduced-price lunch.

The average effect, however, combines different transitional populations and masks important treatment differences. We expect the special-education effect in a given grade to be larger for students who enter special education in that grade than for students who exit special education following the grade. Those who exit may have gained the skills needed to perform in regular classrooms or may not have benefited from the special services. In either case, exiting students likely depress the average program effect, which in their case is based on a comparison of achievement gains while receiving special services and achievement gains following the exit. Table 6 (columns 4 and 6) shows that the estimated program effects are much larger when derived from the students entering special education than from those who exit (although the estimates derived from simple comparisons with regular-education students, reported in column 2, still show no effect for either group). In fact, only one of the special-education effects for those who exit is statistically significant, and this estimate—for emotionally disturbed children—is only 60% as large as the corresponding coefficient for entrants. This pattern is consistent with both the targeting of services toward students who benefit most and positive effects that improve learning in the periods following program participation. Either of these causal linkages could lead to little or no difference in achievement gains during and after program participation—leading us to investigate below the question of whether schools actually target services toward those who benefit most.

Importantly, the entry-exit differential may also capture variations in program effects among disabilities, as the pattern of estimates is consistent with special education having a smaller achievement effect on the speech-impaired (the majority of those who exit) than on the learning-disabled (the largest share of entrants). Estimates for the learning-disabled, emotionally disturbed, and speech-impaired, including separate coefficients for entry and exit, are reported in rows 4 through 12 of Table 6. (All students classified with other disabilities in any of the three grades are excluded from these samples.) For those classified as learning-disabled or emotionally disturbed, the pattern of special-education effects is almost identical to that for all special-education students, including the much larger effects for entrants. In contrast, the effects of special education are very small and statistically insignificant for students classified as speech-impaired. While speech impairments may have adverse effects on reading comprehension (which could spill over into math achievement for some students), most receive targeted services for roughly one-half hour per week, leaving little reason to expect large program achievement effects for most 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 models capture a causal relationship.

Because special education likely affects academic performance even after leaving the program, the entry effects for the specific disabilities constitute the best estimates of effect sizes. Of course the extent that schools target services toward those who benefit most determines the degree to which these estimates overstate the impact on a typical student with the specified disability. In any case, the estimates suggest one year of special-education programming improves performance by 0.1 standard deviations, or a movement of 3–4 percentile points, depending upon where it is evaluated. Thus, one year of special services closes over one-tenth of the average achievement gap between those identified with learning disabilities or emotional problems and regular-education students. These effect sizes are roughly equal to the estimated gains from reducing the typical fourth- or fifth-grade class size by ten students (Rivkin, Hanushek, and Kain, 2001).

B. *Specification Checks*

Procedures for placing students into and out of special education and the limited testing of these students pose the most significant interpretative questions about the previous results. If the probability of classification is not orthogonal to the remaining error, the fixed-effects estimates will be biased. The standard bias calculations for the estimator of λ indicate that the stronger the correlation with the error, the larger will be the bias:

$$E(\hat{\lambda}) - \lambda = \mathrm{cov}(SE,\epsilon)/\mathrm{var}(SE). \quad (2)$$

Given that λ is identified by the difference in achievement gains inside and outside of special education, the estimates will be biased if there are unmeasured changes in personal well-being or the family environment (captured in ϵ) that are systematically related either to the decision to classify or declassify a student as disabled or to the completion of the TAAS examination itself. The simple fact that special-education status changes suggests that something else must also have changed in order to trigger the reclassification; consequently the potential for bias is quite clear.

A key question is why the student gets reclassified. Some actions are explicitly dealt with in our empirical specification. First, some teachers may be more likely than others to initiate classification as special-needs. Because most special-education students also spend time in regular classrooms, if the propensity to classify children is positively related to teacher quality [$\mathrm{cov}(SE, \epsilon) > 0$], the estimates are biased upward; if it is negatively related [$\mathrm{cov}(SE, \epsilon) < 0$], the estimates are biased downward; and if it is unrelated to teacher quality across the sample, no bias is introduced. Note, however, that inclusion of school-by-grade-by-year fixed effects controls for systematic differences in teacher quality by grade and year, effectively removing such variations in teacher quality from ϵ. Similarly, some schools may be more likely than others to initiate classification as special-needs, but the inclusion of school-by-grade-by-year fixed effects also eliminates any bias from systematic differences among schools.

More serious issues can arise if a student performance drop that is precipitated by nondisability factors such as personal or family problems (divorce, job loss, relationship concerns, etc.) also leads to classification as needing special education in the subsequent year. Permanent external changes that affect achievement in all periods do not introduce any problem, because they are controlled for by the individual fixed effects. A temporary downturn that triggers classification but is subsequently reversed, on the other hand, will bias the estimates upward, because it induces $\mathrm{cov}(SE, \epsilon) > 0$. In essence, the estimates of λ confound the actual special-education effect with the recovery from the negative shock, similarly to the phenomenon described as Ashenfelter's dip in the job-training literature (Heckman and Smith, 1999). Indicators for free-lunch eligibility and whether a student switches schools control for gross changes in family economic circumstances but are unlikely to take all relevant changes into account. Note finally that a student performance drop due to such temporary personal or family problems leading to special education classification in the year of the decline introduces a downward bias, because $\mathrm{cov}(SE, \epsilon) < 0$.

To examine possible biases introduced by temporary downturns, we compute interrupted panel estimates of special-education effects. Consider a student who faces a temporary downturn in grade $g-1$ and enters special education in grade g. Because the grade $g-1$ test score is both the post-test score in calculating the grade $g-1$ achievement gain and the pretest score in calculating the grade g gain, the temporary downturn would deflate achievement growth while not in special education and inflate it while in special education, both of which bias upward the estimated special-education effects based on a comparison of grades $g-2$, $g-1$, and g. If, however, information on gains in the year prior to special education entry (grade $g-1$) were not considered and the estimates were computed from comparisons of gains in just grades g and $g-2$, any upward bias should be much smaller. In this case the transitory dip during $g-1$ would be removed from the pre-program comparison, even though it would still inflate the gain in grade g.

Table 7 is based on the sample of students not in special education in their first two years in the data but who enter special education in grade g. (Students whose classification does not change during the entire period are also included.) Estimates in the first column were produced by samples restricted to grades g and $g-2$, whereas those in the second column were produced by samples restricted to grades g and $g-1$ for the same students. Any bias from coincidental temporary factors should be more severe in the second-column estimates based on data for adjacent years than in the first column. However, the interrupted panel estimates in the left column are quite similar in size to the corresponding full-sample estimates in Table 6 (column 6) and not significantly smaller than the estimates in the right column for all disabilities combined and for the learning-disabled and emotionally disturbed, providing evidence against bias from temporary downturns in the year prior to classification. On the other hand, the possibility that a contemporaneous temporary downturn that tends to accompany classification remains. While such effects cannot be identified with our data, their result would be a downward bias, suggesting that our estimates may be lower bounds for the true program effects.

The other aspect of nonrandom selection that may lead to an upward bias is the possibility that schools manipulate which students take the tests. 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. However, such manipulation is generally related to attempts to affect the level of school performance. Because special-education effects estimated here are identified by the difference in test score gains inside and outside 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 with 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 given that the state did not monitor achievement gains for special-education programs during the period under study.

TABLE 7.—INTERRUPTED-PANEL ESTIMATES OF SPECIAL-EDUCATION PROGRAMS ON TEST SCORE GAINS FOR STUDENTS WHO ENTER SPECIAL EDUCATION IN GRADE g BUT ARE NOT IN SPECIAL EDUCATION IN GRADES $g-1$ AND $g-2$

	Estimated Program Impact	
Population	Interrupted Panel, g and $g-2$	Conventional Panel, g and $g-1$
All disabilities	0.05	0.02
	(3.75)	(1.53)
Learning-disabled	0.10	0.02
	(4.71)	(0.82)
Emotionally disturbed	0.10	0.13
	(2.17)	(2.58)

Absolute values of Huber-White adjusted t-statistics in parentheses. Both panel data sets are constructed for the same students. Samples are restricted to students who either enter special education for the first time in the final year of the sample (g) or whose special-education status does not change. Estimation includes student and school-by-grade-by-yearfixed effects. In addition to the indicator for special-education program participation, each regression includes dummy variables indicating cohort and grade, eligibility for a subsidized lunch, and a change of school, as well as the proportions of students in the school who are black, Hispanic, and eligible for a free or reduced-price lunch.

TABLE 8.—ESTIMATED EFFECTS OF SPECIAL-EDUCATION PROGRAMS ON MATHEMATICS TEST SCORE GAINS INCLUDING JUST STUDENTS IN SCHOOLS THAT TESTED ALL STUDENTS BEFORE AND AFTER TRANSITION INTO SPECIAL EDUCATION (ALL DISABILITIES)

	Estimated Program Impact			
Transition Type	Schools with 100% Continuity in Testing		All Schools (from Table 6)	
Entry and Exit	0.04		0.04	
	(3.81)		(5.57)	
Entry		0.08		0.08
		(5.08)		(8.30)
Exit		0.01		0.00
		(0.91)		(0.56)

Absolute value of Huber-White adjusted t-statistics in parentheses. Estimation includes student and school-by-grade-by-yearfixed effects. Restricted sample includes students in schools in which 100% of students who enter or exit special education and took tests while not in special education also took tests as special-education students. In addition to the indicator for special-education program participation, each regression includes dummy variables indicating cohort and grade, eligibility for a subsidized lunch, and a change of school, as well as the proportions of students in the school who are black, Hispanic, and eligible for a free or reduced-price lunch.

Nevertheless, to assess the importance of test selection we repeat the fixed-effects regressions but include only students in schools for which 100% of the students who took tests while not receiving special-education services also took tests while in special-education programs. Slightly more than half of the students whose special-education status changes are excluded from the regressions. Nonetheless, as seen in Table 8, these results yield largely the same conclusions as those based on the full sample: Special-education programs improve student performance significantly. Even though the more stringent sample selection criterion reduces the sample sizes, the point estimates for students with all disabilities are virtually identical to those previously shown for the full sample (repeated from Table 6 in the final two columns).[18]

One other possible source of bias is that schools do have some discretion in classifying students and may somewhat arbitrarily select some students for special education out of all struggling students or out of the subset of students who exhibit problems that fall under the rubric of learning disabilities. The final specification check, based on the imprecise nature of classification, compares the change in achievement gains of special-education entrants with those of students who will enter special education in the future.[19] Such students may provide a better baseline achievement growth profile than regular-education students, particularly if special-education effects tend to decline with increasing

[18] Point estimates for the learning-disabled students from the more stringent sample are qualitatively similar, though somewhat smaller and less significant, than those from the full sample. The emotionally disturbed samples were too small to support estimation.

[19] We thank an anonymous referee for this suggestion.

grade level, say, because the effect of any disability becomes more pronounced as the subject matter becomes more difficult. Alternative estimates (not shown) of special-education effects for entrants in $g-1$ were derived from two comparison groups: (1) students who enter special education in the following year (g), and (2) students whose classification never changes between $g-2$ and g. To the extent that current entrants are more similar to future entrants than to other students, the first comparison group would provide better estimates of special education effects than the second. In fact, the alternative comparison groups yield virtually identical estimated program effects for students with all disabilities and for the learning-disabled (not reported).

While we directly address the potential impacts of temporary shocks to achievement, of manipulation in test taking, and of changes in school quality, it is always difficult to rule out the possibility that still other, unknown factors might contaminate the estimates. An alternative strategy would employ instrumental variables estimation to isolate exogenous variations in participation. It is generally difficult to identify valid instruments—variables related to the probability of special-education classification but otherwise unrelated to achievement gains—in situations of joint decision-making such as is involved in school programming. As Cullen (1997) has shown, however, changes in Texas school financing formulas during the 1990s offer some hope. Texas altered the additional dollars received for classifying a student as disabled, and, because the magnitude of the change in state revenue depended upon district wealth, these changes created variations over time and between districts in the fiscal incentives to classify students as disabled which may well be orthogonal to systematic changes in other determinants of achievement. Following Cullen (1997), we consider, as an instrument for special-education placement, the predicted change in state aid by district from classifying an additional student as disabled.[20] The first-stage estimates (not reported) reveal the expected positive relationship between the probability of receiving special education and the predicted revenue gain. In contrast to Cullen's work at the district level, however, the coefficient is only marginally significant at the 5% level and has little explanatory power. As a result, the instrumental variables estimates are very imprecise, and they do not provide any additional information.

[20] 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 experimented 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. Moreover, even the predicted revenue change might be related to determinants of achievement if other changes in state educational policy during this period varied systematically by district size and wealth.

TABLE 9.—COMPARISON OF ESTIMATED EFFECTS OF SPECIAL-EDUCATION PROGRAMS ON MATHEMATICS TEST SCORE GAINS FOR STUDENTS EXITING VERSUS STUDENTS CONTINUING IN SPECIAL EDUCATION

	Estimated Program Impact	
Year	Single Year of Special-Education Programs[a]	Two Years of Special-Education Programs[b]
	All Disabilities	
Entry year	0.01	0.19
	(0.30)	(12.17)
Second year[c]	0.03	0.14
	(1.40)	(8.52)
	Learning-Disabled	
Entry year	0.03	0.15
	(1.03)	(5.37)
Second year[c]	0.11	0.09
	(1.43)	(3.46)

[a] Sample of entrants into special education in grade $g - 1$ who exit in grade g.
[b] Sample of entrants into special education in grade $g - 1$ who remain in special education in grade g.
[c] Relative to pretreatment, $g - 2$.
Absolute value of Huber-White adjusted t-statistics in parentheses. Samples for each of the specifications are restricted to students whose transitions match the description at the top of the column or whose special-education status does not change. Estimation includes student and school-by-grade-by-year fixed effects. In addition to the indicator for special-education program participation, each regression includes dummy variables indicating cohort and grade, eligibility for a subsidized lunch, and a change of school, as well as the proportions of students in the school who are black, Hispanic, and eligible for a free or reduced-price lunch.

C. *Targeting and Dynamics*

An important question is whether the placement interactions of schools and families lead to targeting services toward students who receive the greatest benefits. While larger effects for entrants than for students who exit special education are consistent with the notion that schools target services where they are most effective, they are certainly far from definitive evidence.

Some insight can be obtained by comparing the achievement patterns for those in special education for a single year with those who remain in special education for two years.[21] The first column of Table 9 provides information on why the test score gains after exiting special education are similar to gains while in the program. The only transitions included in these regressions are those students not in special education in grade $g - 2$ (first year) who enter special education in grade $g - 1$ and exit again in grade g. Coefficients are reported for two variables: "entry year" means the estimated gain differential between achievement in the entry year ($g - 1$) and the prior year in regular education ($g - 2$), and "second year" means the estimated gain differential between achievement in the third year (g) and the pretreatment year ($g - 2$). Notice that in these regressions the coefficient on "second year" compares achievement gains for two periods in which the student is not classified as special-education. If special education helped these students so much that they no longer needed assistance, the gain in the year following exit from special education (g) should be significantly larger than the gain in the year prior to

[21] Ideally we would want to compare achievement patterns over more years, but we are limited to three observations for each student.

classification as special education ($g-2$). On the other hand, if these students were dropped from special education because it had little or no positive effect on achievement, one would expect little or no difference between their performance prior to and following the special intervention.

The estimates provide somewhat noisy evidence on the question of whether achievement rises for students who spend only one year in special education. Neither of the two coefficients is statistically significant for either the sample of students with all disabilities or those classified as learning-disabled, though all are positive, and the learning-disabled coefficient on "second year" is large. Unfortunately, the small number of the specified transitions among the learning-disabled reduces the precision of the estimates. Nonetheless, the small and insignificant coefficients in the specifications including all students classified as disabled suggests that those who exit special education after one year derive smaller benefits from the intervention, at least in terms of higher academic achievement.

The second column in Table 9 provides a preliminary look at the dynamics of special-education effects for students who remain in the program for at least two years. The only transitions included in these regressions are for those students not in special education in the first year ($g - 2$) who enter special education in the second year ($g - 1$) and remain in special education in the third year (g). The results indicate a positive effect each year for students who remain both years, but with diminishing returns. (As with the first column, both special-education variables (for $g - 1$ and g) are parameterized to measure the difference between achievement gains in the given year and the pretreatment period, $g - 2$.) The program impact declines by roughly 25% in the second year for the average participant and roughly 40% for the average learning-disabled student. Nonetheless, these students are clearly benefiting in the second year of the program.

Overall, the results in Table 9 are consistent with the hypothesis that schools target special-education services to students who derive greater benefits: the hypothesis that special education has no affect on mathematics achievement cannot be rejected for students who exit after one year in the program, and special education has a significant positive effect in both years for students who remain in special education for the second year.

D. *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, & Terman, 1996), but 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 to every extent possible in the regular classroom setting.[22] 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, as did the proportion of special-education students mainstreamed into regular classrooms.[23] At the same time, the use of mainstreaming appears to be a source of conflict with parents of students in regular education, who are worried that special-education students may detract from the education of their children.

TABLE 10.—ESTIMATED EFFECTS OF SPECIAL-EDUCATION PROGRAMS ON MATHEMATICS TEST SCORE GAINS FOR ALL SPECIAL-EDUCATION STUDENTS AND STUDENTS CLASSIFIED AS LEARNING-DISABLED, BY SETTING

Program Setting	Estimated Program Impact
All Disabilities	
All settings	0.04
Additional	(6.51)
Additional mainstream impact	0.00
	(0.55)
Learning-Disabled	
All settings	0.06
	(6.12)
Additional mainstream impact	0.01
	(1.33)

Absolute value of Huber-White adjusted *t*-statistics in parentheses. Estimation includes student and school-by-grade-by-yearfixed effects. In addition to the indicator for special-education program participation, each regression includes dummy variables indicating cohort and grade, eligibility for a subsidized lunch, and a change of school, as well as the proportions of students in the school who are black, Hispanic, and eligible for a free or reduced-price lunch.

While the objectives of mainstreaming go far beyond achievement gains, its impact on achievement is nevertheless important. Table 10 reports fixed-effects estimates of special education by mainstream status for all special-education students and separately for those classified as learning-disabled. The fixed-effects specifications, where the coefficients are identified by students who switch program type or special-education status, control for student heterogeneity. These estimates reveal no significant difference by treatment setting in the impact on achievement.

One difficulty in interpreting these results is that students are not randomly selected into programs, so this is not the type of unambiguous experiment that would be produced if students were randomly assigned to different settings.[24] Student fixed effects remove time-invariant characteristics, but any treatment assignment based on observing students in different settings could potentially contaminate the estimated impact of mainstreaming. A possible solution is the

[22] Mainstreamed special-education students receive all special services within a regular classroom, whereas nonmainstreamed students spend part or all of the day outside the regular classroom, say, in a resource room.

[23] The percentage mainstreamed rose from 5% to 10% in a single year.

[24] The simple estimates in column 1, which do not control for individual and school fixed effects, find positive effects of special education only for mainstreamed students—suggesting purposeful placement of students into different settings.

use of the previously described fiscal instruments, but the weak first-stage explanatory power of fiscal incentives for mainstreaming again makes instrumental variables estimates uninformative. Thus, while there is little evidence that mainstreaming reduces achievement gains for students currently in that setting, a policy of expanding mainstreaming to include other special-education students might lead to different results.

A second problem of interpretation comes from possible variations in the definition of what is and is not mainstreaming. For example, although mainstreamed students will spend the entire day in regular classrooms, students pulled out for treatment in resource rooms may still spend a majority of the day in regular classes. Unfortunately, no measure of actual exposure to non-special-education class activities exists for students not mainstreamed. Overall, the results at this point are consistent either with the setting not making a difference for student performance, with school officials being very effective at designing programs of study for each individual student, or with errors in variables such that the given categories do not provide meaningful information concerning the substantive differences among settings.

VII. The Effects of Special-Education Programs on Regular-Education Students

The final component of this analysis considers the effects of special-education programs on regular-education students. The most systematic investigation of this issue is found in 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 variables techniques. Her results show that increases in special-education costs not covered by state or federal sources reduce regular-education funding and achievement, a result 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 non-special-education students? Such changes may affect regular-education students in myriad 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 et al. (2001) 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 take advantage of the multiple student cohorts and estimate these models with a series of explicit controls for students and peers as well as student and school-by-grade fixed effects.[25] Essentially, the proportion classified as special-education replaces the special-education indicator variable in equation (1), and the sample is restricted to students not classified as disabled during the three-year sample period. We also include information on teacher experience and average class size of regular classrooms in some preliminary specifications that we discuss below, to control for other changes in school characteristics that might coincide with changes in special-education enrollment. However, since special-education enrollment might affect regular class sizes and teacher experience, we just report the reduced-form specifications that exclude these school characteristics.

The ability to control for school-by-grade fixed effects rather than just school fixed effects is quite important. While removing much of the confounding influences of other factors, student and school fixed effects alone may fail to eliminate all biases if school average achievement and proportion classified as special-education change in a systematic way as students progress through school. Consider the possibility that achievement for students in some schools tends to decline as the students age, particularly as they become adolescents, and that these are the students who also experience the largest increases in the proportion of their classmates who are classified as special education (or, more specifically, as learning-disabled or emotionally disturbed).[26] If only fixed individual and school effects were removed, the estimates would be identified by between-grade differences in the proportion in special education. In the case sketched above, the estimates would show a large effect of the proportion in special education on achievement, when in fact the relationship would be spurious and the decline in achievement would be brought about by other factors. In contrast, the removal of student and school-by-grade fixed effects means that coefficients are identified by cohort differences in changes in the proportion in special education as students progress through school. Such between-cohort differences likely emanate from a combination of random differences among cohorts in classification rates and changes in school district policies toward special education. It seems highly unlikely that schools or school districts would base special-education policy on the expected changes over time in achievement gains for specific cohorts, and thus the coefficients should provide consistent estimates of the effect of the proportion in special education on achievement for students not classified as special-education.

Table 11 reports estimated effects of the proportion in special education for simple regressions with just observed individual and school variables, for regressions with student fixed effects, and for regressions with student and school-

[25] School-by-grade-by-year fixed effects cannot be used, because all students attending grade g in school s in year t experience the same composition of students.

[26] We thank an anonymous referee for providing this example.

TABLE 11.—ESTIMATED EFFECTS OF THE PERCENTAGE OF STUDENTS CLASSIFIED AS DISABLED ON MATHEMATICS TEST SCORE GAINS FOR STUDENTS NOT RECEIVING SPECIAL EDUCATION, BY DISABILITY AND SETTING

	Estimated Regular Education Impact		
Type of Proportion	No Fixed Effects	Student Fixed Effects	Student and School-by-Grade Fixed Effects
Total special education:			
Proportion of special education	0.00	0.20	0.16
	(0.08)	(3.62)	(2.74)
Proportions by setting:			
Mainstreamed	0.00	−0.11	0.13
	(0.08)	(0.93)	(1.08)
Not mainstreamed	0.00	0.22	0.16
	(0.10)	(3.97)	(2.75)
Proportions by disability type:			
Learning-disabled or emotionally disturbed	−0.12	−0.12	−0.05
	(3.11)	(1.76)	(0.72)
Speech-impaired	0.18	0.48	0.19
	(1.60)	(2.77)	(1.15)
Other disabilities	0.27	0.60	0.43
	(4.68)	(6.99)	(5.10)

Absolute value of Huber-White adjusted t-statistics in parentheses. The results of three separate regressions are presented for each estimation approach (column). The first aggregates all special-education students into a single group, the second distinguishes among special-education students on the basis of setting, and the third distinguishes among them on the basis of disability type. In addition to the proportion in special education, each regression includes dummy variables indicating cohort and grade, eligibility for a subsidized lunch, and a change of school; the proportions of students in the school who are black, Hispanic, and eligible for a free or reduced-price lunch; and class size.

by-grade fixed-effects specifications.[27] Regardless of the estimation approach, there is no evidence that special education harms achievement in regular classrooms. The preferred fixed-effects results show 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 0.016 standard deviations. (A change of one special-education student in a class of 20 students would be a 5% change. The standard deviation in our sample of the percentage classified as disabled is 6%). While program selection factors related to which students avoid classification as disabled for the three-year period could again contaminate these estimates, their lack of impact in the prior analysis of special-education students and the fact that special-education students constitute only a small percentage of all students strongly suggest that endogenous selection into special education is unlikely to contaminate these results.

Within our data, we can investigate some of the transmission mechanisms that might link achievement and proportion classified as special-education. First, if special education affected achievement through fiscal influences on the quantity of resources devoted to regular education, the inclusion of the specific resource measures of teacher experience and class size should alter the coefficient on the percentage classified as disabled.[28] The results (not reported) show that the inclusion of these variables has virtually no effect on the special-education coefficients, providing preliminary evidence that resources are not driving the link between achievement and special-education classification rates.[29] Second, as shown in Table 11, treatment setting for special-education students (mainstreamed or not) makes no significant difference (although the effect of proportion mainstreamed is quite noisily estimated). Finally, if special education provides a means of removing unruly students or those experiencing great difficulty with the material, increases in the proportion classified as learning-disabled or emotionally disturbed would be expected to raise achievement for regular classroom students. To the contrary, the bottom panel of Table 11 shows that the proportion with "other disabilities" exhibits the strongest positive relationship for the achievement of regular-education students. Of course, changes in the proportion classified as learning-disabled or emotionally disturbed may also reflect changes in the proportion of students with behavioral or learning difficulties. All in all, the estimates in Tables 11 suggest that, as currently operated, special education does not harm and may even help regular-education students on average, though the precise underlying causes for such a positive relationship are not readily identified.

VIII. Conclusions

For good reason, previous discussions of special education have concentrated on issues related to costs. Outlays to provide schooling for students with identified handicaps average more than twice those for regular education. Yet the focus on costs has often obscured the fact that there is educational purpose in 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 large panel data set that follows gains of individual students

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

[28] It is important to note, however, that we use grade- and year-specific classification rates to identify the coefficients, 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. Special-education programs likely reduce effective class size for the regular education students—either because of pullouts or because of other resources devoted to the special-education students that allow for substitution of the regular teacher's time. (The class size measure reported by the regular-education teachers in a grade should include the special-education students assigned to the class, although it is probably subject to error.)

[29] 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.

across time and programs provides a unique window on program effects. Specifically, the repeated performance measures allow us to identify program effects by contrasting the achievement gains of students who experience both special education and regular education. Furthermore, because movements into and out of special education—particularly when disaggregated by specific disability—remain relatively rare phenomena, very large samples are required to obtain reliable estimates of program effects.

The estimates, which fully allow for any persistent individual handicapping conditions, ability differences, and time-varying grade-within-school differences in school quality, indicate that special-education programs on average have a significantly beneficial effect on performance. Our best estimate is that one year in a special-education program boosts average math scores by roughly 0.1 standard deviations over what would be expected in the absence of the intervention. As expected, program effects are much larger for students classified as learning-disabled or emotionally disturbed, conditions that directly impede classroom performance, than they are for speech impairments, a condition less likely to have a major effect on mathematics achievement. In addition, the evidence is consistent with schools targeting services toward those who benefit more.

An elaborate series of specification analyses allows for temporary achievement declines that might trigger placement in special education and bias upward the estimated effects on 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 of these appears to contaminate the estimates.

Similar estimation of achievement growth by students in regular education provides no evidence that higher rates of special-education classification detract from their performance—suggesting a much more benign view of special education than is typically found. Quite at odds with much of the general discussion, there is no evidence that mainstreaming adversely affects regular-education students. Note, however, that our analysis necessarily ignores any negative impacts of special education common to all schools in Texas, such as reduced overall state aid for regular education.

This analysis concentrates on average program effects, only minimally disaggregated by setting. Our other work (Rivkin et al., 2001), however, shows dramatic differences in achievement across teachers of regular-education students. Since many special-education students spend significant time with the regular classroom teacher, this by itself would be expected to have powerful effects on the achievement of special-education students. It is further reasonable to believe that similar variations in the quality of special-education teachers are important. Lyon and Fletcher (2001) further point to systematic differences in both teacher quality and in program content as being important in dealing with reading disabilities and reading problems treated as learning disabilities under special education. More analysis is needed to investigate the heterogeneity of performance across teachers, programs, and schools.

None of this analysis has considered costs, even though special education undoubtedly involves additional spending. In addition, the analysis has also concentrated exclusively on issues of academic achievement, even though 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, and the estimates are identified solely according to the performance of students who transition between regular and special education. These issues suggest that further analysis is required to understand both the generalizability of the results to the entire special-education population and the larger impact of these programs outside the achievement realm.

Finally, the methodologies of this paper are directly transferable to a number of other policy areas where inferring program effectiveness is made difficult by the nature of the populations being served. When programs are designed to treat very special groups—disabled students here—the population not falling into them is a poor comparison group. Evaluations of programs for other special groups—such as students with limited English proficiency, Medicaid recipients, individuals receiving psychiatric services, or people in drug rehabilitation—face similar problems.[30] Moreover, specific programs under consideration often overlap with other programs and services, making the identification of specific program effects difficult. The use of stacked panel data permits the separation of program effects from other treatment effects and other factors influencing performance.

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* 26:4 (1993), 344–370.

Cullen, Julie Berry, "Essays on Special Education Finance and Intergovernmental Relations." Ph.D. Dissertation, Massachusetts Institute of Technology (1997).

Cullen, Julie B., and Steven G. Rivkin, "The Role of Special Education in School Choice," in Caroline M. Hoxby (Ed.), *The Economics of School Choice* (Chicago: University of Chicago Press, 2003).

Hanushek, Eric A., and Steven G. Rivkin, "Understanding the Twentieth-Century Growth in U.S. School Spending," *Journal of Human Resources* 32:1 (1997), 35–68.

Hartman, William T., "Policy Effects of Special Education Funding Formulas," *Journal of Education Finance* 6 (Fall 1980), 135–359.

Heckman, James J., and Jeffrey A. Smith, "The Pre-programme Earnings Dip and the Determinants of Participation in a Social Programme: Implications for Simple Programme Evaluation Strategies," *The Economic Journal* 109 (July 1999), 313–348.

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).

[30] Evaluations in some areas, such as manpower training programs, have pursued longitudinal designs to infer program effects, but they are often subject to various complications similar to those considered here.

———, “The Allocation of Resources to Special Education and Regular Instruction” (pp. 221–257), in Helen F. Ladd (Ed.), *Holding Schools Accountable: Performance-Based Reform in Education* (Washington, DC: Brookings, 1996).

Lewit, Eugene M., and Linda Schuurmann Baker, “Child Indicators: Children in Special Education,” *The Future of Children* 6:1 (1996), 139–151.

Lyon, G. Reid, and Jack M. Fletcher, “Early Warning System: How to Prevent Reading Disabilities,” *Education Matters* 1:2 (2001), 23–29.

Lyon, G. Reid, Jack M. Fletcher, Sally E. Shaywitz, Bennett A. Shaywitz, Joseph K. Torgesen, Frank B. Wood, Ann Schulte, and Richard Olson, “Rethinking Learning Disabilities,” in Chester E. Finn Jr., Andrew J. Rotherham, and Charles R. Hokanson Jr. (Eds.), *Rethinking Special Education for a New Century* (Washington, DC: Thomas B. Fordham Foundation and Progressive Policy Institute, 2001).

Martin, Edwin W., Reed Martin, and Donna L. Terman, “The Legislative and Litigation History of Special Education,” *The Future of Children* 6:1 (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 Education Statistics, 1997).

Parrish, Thomas B., and Jay G. Chambers, “Financing Special Education,” *The Future of Children* 6:1 (1996), 121–138.

Reschly, Daniel J., “Identification and Assessment of Students with Disabilities,” *The Future of Children* 6:1 (1996), 40–53.

Rivkin, Steven G., Eric A. Hanushek, and John F. Kain. “Teachers, Schools, and Academic Achievement,” National Bureau of Economic Research working paper no. 6691 (revised) (2001).

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* 57:2 (1987), 125–152.

Wagner, Mary M., and Jose Blackorby, “Transition from High School to Work or College: How Special Education Students Fare,” *The Future of Children* 6:1 (1996), 103–120.