Causes and Consequences of Special Education Placement: Evidence from Chicago Public Schools*

Jessica L. Cohen†
Brookings Institution

Current Version: December 2007

Abstract

Despite the fact that one out of every seven U.S. public school students receives Special Education (SE) services, little is known about the direct impact of SE placement on students' social and academic outcomes. This paper exploits the strategic incentive to increase SE enrollment induced by a 1996 accountability policy in Chicago Public Schools to identify the impact of SE placement on high school completion, absenteeism and GPA. Pre-accountability performance characteristics of the school determined to what extent sanctions could be avoided by increasing SE placement, since SE students' scores were excluded from accountability measures. I construct an instrument that captures the strength of strategic incentives, and show that low-achieving students in high-incentive schools experienced the largest increase in SE placement. Using instrumental variables analysis and a panel of student data from Chicago Public Schools, I find that SE placement in elementary school reduces the probability of dropping out of high school and absenteeism for the marginal low-achieving student, while results on GPA are inconclusive. I provide evidence that these results are not driven by other changes taking place at high-incentive schools. The results suggest that low-achieving students benefit from SE placement for mild mental disabilities.

*I am very grateful to Amy Finkelstein, Michael Greenstone and especially David Autor and Esther Duflo for their guidance and extensive comments. I have gotten excellent feedback from David Abrams, Liz Oltmans-Ananat, Pascaline Dupas, Greg Fischer, Raymond Guiteras, Guy Michaels, Peter Schnabl, Sarah Siegel, and participants at the Brookings Institution Economic Studies Seminar, the Center for Global Development Seminar, MIT Public Economics Seminar and MIT field lunches. I thank Chicago Public Schools and the Consortium on Chicago School Research for the data used in this study, especially Vanessa Coca and Amy Nowell. This research was funded by a National Science Foundation Graduate Research Fellowship. All errors are my own.

†E-mail: jcohen@brookings.edu

1 Introduction

Nearly seven million children with disabilities currently receive Special Education (SE) services, roughly 13.5 percent of total enrollment in public and private schools. This number has been growing steadily since 1975 when the Education for All Handicapped Children Act (EHCA) was passed and when only one in five children with disabilities was being educated in the public school system (Singer and Butler 1987).[1] Figures 1a and 1b show the trend in SE enrollment in the U.S. over the past three decades. The level of SE enrollment (Figure 1a), as well as its prevalence among all students (Figure 1b), grew rapidly after the passage of EHCA and then steadily throughout the 1980s. SE has grown by 40 percent since the early 1990s, largely due to the increasing diagnosis of "soft disabilities" such as mild mental/behavioral disorders (e.g. Attention Deficit Hyperactivity Disorder) and the rising inclusion of pre-school and kindergarten students with developmental delays. Despite the fact that SE enrollment is so high and roughly 22 percent of total current educational spending is dedicated to these services, there is surprisingly little solid evidence of its benefits (OSEP 2002). The net benefit of SE services is unclear. SE may improve social and academic outcomes by providing disabled students with extra resources, attention and accommodations. On the other hand, SE may increase the risk of negative outcomes because of negative peer effects, lower expectations and stigma effects (Martinez 2004, Bear 1993, Meltzer 1998).

Students with disabilities experience poor social and academic outcomes at very high rates.[2] SE students are about 30 percent more likely to drop out of high school than students in the general population. Those who drop out are 50 percent less likely to be employed, while those who graduate are 80 percent less likely to attend post-secondary school. However, OLS estimates of the effect of SE cannot be interpreted as causal, since SE students come into the school system with higher risk factors for low achievement and poor social outcomes.[3] Beyond

[1] The Individuals with Disabilities Act (IDEA) is the current version of the EHCA.

[2] The following estimates are taken from the Digest of Education Statistics (2005) and from Wagner, et al. (2005).

[3] SE students are 50 percent more likely to have unemployed parents, 24 percent more likely to come from

observable background differences, students are screened for and placed in SE precisely because they are thought to be at risk of failing and disengaging from school. If all of the criteria for SE placement were well-defined and implemented uniformly, it would be straightforward to estimate the impact of SE by comparing outcomes for students close to the "disability cutoff". In fact, disability criteria are not straightforward or standardized and identification of the impact of SE requires an instrument for SE status. Some severely disabled students are clearly in need of SE services and some high-achieving students clearly are not. This paper uses an instrument that predicts SE placement for the marginal, low-achieving student, in order to measure the direct impact of SE placement on high school outcomes.

The introduction of accountability in Chicago Public Schools (CPS) provides an excellent opportunity to study the impact of SE placement. Jacob (2005) finds evidence that teachers and administrators responded strategically to the policy along a number of dimensions. One of these responses was a dramatic increase in SE enrollment, since school performance measures excluded scores for SE students. Figures 2a and 2b plot the trend in SE enrollment in CPS over the past decade. The fraction of CPS students receiving SE increased substantially (roughly 3 percentage points) following the introduction of accountability in 1996. While the number of students in SE grew by less than 1,000 in the three years prior to accountability (1994-1996), it grew by nearly 10,000 students in the three years after the policy was introduced. This paper exploits the abrupt change in SE enrollment to construct an instrument for SE placement in elementary school. I use an individual-level panel dataset from CPS administrative records (1994-2004) containing detailed background information and performance measures. This dataset allows me to follow students throughout their school years, observing SE placement in elementary school and subsequent high school outcomes.[4]

The instrument for SE status utilizes three sources of variation: First is the large temporal

households with less than a high school diploma and 28 percent more likely to come from single-parent households. (Wagner, et al., 2003)

[4]Some students do get placed in SE for the first time in high school, but this is much less common than placement in elementary school. SE placement in CPS is most common between the ages of ten and twelve.

variation coming from the introduction of accountability. Students in elementary school after the policy change were more likely to be placed in SE than students in previous years. The instrument also incorporates cross-school variation in the probability of SE placement. This second source of identification comes from differences in the strength of incentives to increase SE enrollment across schools.[5] The variation in incentives stems from performance characteristics of the school from the year before the policy was announced, which determined to what extent it could utilize SE enrollment to avoid disciplinary measures. I will illustrate that the largest increases in SE placement occurred in schools with pre-accountability performance measures putting them close to the disciplinary cutoff. The third source of variation is in student achievement (test scores) prior to accountability. This triple-difference strategy explores whether the increase in SE placement for students in high-incentive schools is largest for the lowest-performing students. The use of within-school variation in prior achievement purges all other changes occurring at high-incentive schools from the estimation, as long as these changes were not targeted differentially across achievement levels. I will present evidence that schools with strong incentives do not appear to have targeted resources toward low-achieving students.

In addition to a plausibly exogenous source of identification for SE placement, this study has a number of valuable features. Previous research on the impact of SE has been hindered not only by the endogeneity of SE placement, but also by the large amount of regional variation in the selection process, disability criteria and SE program characteristics. Any cross-sectional study of outcomes for SE students must try to control for many confounding sources of variation. Further, differences in the selection process make it difficult to interpret estimates of the impact of SE, since the characteristics of the marginal student are unclear. The data used in this study is detailed enough to characterize the (average) marginal SE student. Also, previous research on SE has been plagued by small sample sizes, making it difficult to measure its impact with any precision. The analysis here focuses on one large urban district, so it is not confounded by

[5] All specifications include year dummies to absorb temporal variation in SE enrollment that is common across schools, and school dummies to absorb fixed characteristics of schools correlated with SE enrollment.

regional variation in funding, disability criteria, population characteristics or program design. Since Chicago is the third largest school district in the country, it has a sizeable number of SE students (about 36,000), all of whom are captured in the CPS administrative data used in this paper.

I am aware of only one previous study with a large sample size that attempts to measure the impact of SE. Hanushek, et al. (2002) use the UTD Texas Schools dataset to analyze the impact of SE placement on math test score gains. Using an individual-level fixed effects design, they use variation in SE entrance and exit patterns over a three year period to estimate the impact of SE. They find that placement in SE increases growth in math scores by 0.1 standard deviations in the first year, but that this effect falls by 40 percent the following year. The main drawback of this study is that only 30 percent of the SE sample has test scores and these are likely the highest achieving SE students. Also, one would ideally like to measure the impact of SE on the basis of longer term outcomes, particularly since they find such a rapid deterioration in achievement gains.

I find that SE placement in elementary school reduces absenteeism and the probability of dropping out of high school, but results for GPA are inconclusive and noisy. These results are consistent with strong negative selection bias in OLS estimates. IV results are consistent across a variety of specifications and do not appear to be explained by other potential responses to the policy occurring at schools with strong incentives. The main concern is that increased incentives induced a reallocation of resources toward low-performing students. There are a number of reasons to believe that this type of bias is not present. First, while the accountability policy induced schools to try to improve ITBS scores, Jacob's (2005) results suggest that the policy did not lead to a general improvement in academic achievement. It is thus questionable how much we would expect the increased effort and resources aimed at improving ITBS scores to be reflected in high school outcomes. Second, if schools were to target resources differentially by student achievement, we would expect them to invest most in middle-achieving students since these are the students who could potentially improve a school's standing, but I find no evidence that outcomes improved most for middle-achieving students. Finally, if schools with strong incentives targeted resources toward low-achieving students, these efforts should be reflected in higher scores on the test used for accountability. However, I find that scores decreased for low-achieving students in high-incentive schools, suggesting that the exclusion restriction is valid.

The next section discusses selection into SE and the criteria for disability classification; it also addresses the mechanisms through which SE may be affecting outcomes. Section 3 discusses the accountability policy in CPS, Section 4 describes the data used in this study and Section 5 outlines the empirical strategy used to identify the impact of SE placement on high school outcomes. Sections 6 and 7 present first stage and reduced form results, respectively, and Section 8 compares OLS and IV estimates of the impact of SE. Section 9 discusses the robustness and external validity of the results and Section 10 concludes.

2 Special Education Assessment and Services

2.1 Disability Assessment and Special Education Placement

U.S. public schools are required to provide SE services to children with disabilities varying in severity from mild learning disabilities and speech problems to severe mental retardation and autism. The purpose of SE is to assist students who cannot function effectively in the regular classroom with an alternative curriculum tailored to their particular needs. The majority of students in SE (roughly 2/3) are being treated for some form of learning disability (LD). This discussion will focus on SE selection and services for students with LD, not only because it is the most common form of disability and the margin along which SE is growing nationally, but also because nearly all of the increase in SE placement that occurred in Chicago was for LD students.

Children that are diagnosed with LD are not exclusively low-achieving.[6] In fact, one of the

[6] The Special Education policies and procedures discussed here are those set forth by the Illinois State Board of Education (ISBE), but are based on national guidelines and are representative of procedures followed in other states. The most current version of the ISBE Special Education Policies and Procedures can be found at http://www.isbe.net/spec-ed/pdfs/policies.pdf.

criteria for the diagnosis of LD is a significant discrepancy between ability (usually gauged with an IQ test) and achievement in a particular subject.[7] LD is not only distinct from overall low-achievement, but students achieving very far below average are by definition excluded from the LD classification and instead must be tested for a Mental Handicap.[8] In order to qualify for SE the student must also be exhibiting difficulty functioning in school. Students who receive the LD diagnosis are generally slow learners—most of them scoring in the bottom quartile on reading and math tests—but it is not exclusively the very bottom scoring students that are diagnosed with this disability.

Typically, a student comes under evaluation for SE because his parent or teacher notices that he is struggling and requests a disability assessment.[9] On the basis of the child's academic and behavioral functioning, an evaluation team (composed of teachers, administrators and psychologists) makes recommendations for program modifications and extra assistance within the regular curriculum. If the student continues to struggle, a full disability assessment is conducted using a battery of achievement and aptitude tests.

SE for non-physical disabilities is unlike other public services that are provided to children shown to meet certain objective criteria. There is quite a bit of discretion in the decision to place a child in SE for a "soft disability" such as an emotional disturbance or LD and the characteristics of students who are treated for these disabilities vary widely across time, states and even across districts (Reschly, 1996; Lewit and Baker, 1996; Mercer et al., 1990). The appropriate guidelines for classification of LD, and how to get SE personnel to implement these guidelines consistently, have been a major topic of debate in the education literature over the past twenty years (Algozzine and Ysseldyke 1986). Though much of the empirical research is outdated, education researchers have found that, while many students are designated LD

[7] This discrepancy is, for most states, a two standard deviation difference between the student's score on an IQ test and an achievement test.

[8] A student must typically be scoring at least three standard deviations below the mean to be considered Mentally Handicapped (MH). As with the LD criteria, achievement scores cannot be the only criteria used in considering whether a child is MH. They must also exhibit difficulty executing age-appropriate tasks.

[9] Some disability screening is routine and given to all students, such as hearing and vision screening.

without meeting the criteria, significant numbers of regular education students who should be treated for LD go undiagnosed. In their paper on LD classification in Indiana public schools, McLeskey and Waldron (1991) administer IQ and achievement tests to a sample of SE students and find that, even after the Board of Education clarified LD classification criteria and initiated an aggressive information and training campaign, less than 2/3 of SE students being treated for LD met the state guidelines. Previous to this initiative, less than 1/3 of LD students met the criteria.[10] In a similar study, Shepard and Smith (1983) find that only 43 percent of Colorado LD students were appropriately placed according to state criteria.

The scope for discretion in the diagnosis of some disabilities has left room for strategic considerations in SE placement. Cullen (2003) uses variation in the amount of state aid a school district in Texas can receive by placing a child in SE to measure the impact of fiscal incentives on disability diagnoses. She finds that a large portion of the growth in SE placement in Texas over the 1990s can be explained by fiscal incentives. A number of studies on strategic responses to high-stakes-testing policies find that schools increase SE placement when these students' scores are excluded from accountability measures (Jacob, 2005; Figlio and Getzler, 2002; Cullen, 2006).

The vague and inconsistent criteria for SE placement also exacerbate the omitted variables bias in OLS estimates of the impact of SE. In the absence of strict and clear guidelines for a disability diagnosis, OLS estimates may be biased by selection based on characteristics of school personnel, parents, etc. For example, parents might switch schools to acquire or to gain a disability diagnosis (Cullen and Rivkin, 2003), examiners may selectively administer subtests to confirm their hypotheses (Mellard, 1985), and racial discrimination may be present in a teacher's recommendation that a child be screened for a disability (Losen and Orfield, 2002).

[10] Estimates were based on two stratified random samples (taken before and after the guidelines were clarified) of roughly 800 students. The "severe discrepancy" criterion in Indiana at the time was an 18 point (1.2 standard deviation) difference between actual and expected achievement scores, where the expectation is determined by a regression of achievement on IQ, sex, race, etc. for all of the students in the school or district.

2.2 Special Education Programs

Once a child is diagnosed with a disability her evaluation team designs an Individualized Education Program (IEP). The IEP specifies in which areas the child is falling behind, how extra assistance will be provided, and how the child's progress will be monitored. In crafting an IEP, the team is required by the Least Restrictive Environment (LRE) clause of the IDEA to keep the child in the regular classroom, learning the regular curriculum, to the greatest extent possible. Most children with LD spend three to four hours a week in a resource room getting extra assistance in the subject(s) in which they are struggling. Many students with mild disabilities are "mainstreamed" and never leave the regular classroom. These students are typically given extra assistance within the classroom by a teacher's aid.

The purpose of SE is not only to improve achievement, but to improve social outcomes and help students remain engaged in school. The two main components of the program are extra assistance and alternative standards for assessment and promotion. There are a number of potential positive and negative consequences to SE placement for the marginal, slow-learning student and *a priori* the net benefit of SE is unclear. Students may benefit from the specialized instruction and the extra resources and time to learn the material. They may also benefit from a tailored curriculum, with easier promotion standards and requirements for credit accumulation. Alternatively, placement may be harmful, since SE students spend more time with a low-achieving peer group. SE students may also be challenged less and be harmed by lower expectations of teachers, parents and peers (Martinez 2004, Bear 1993, Meltzer 1998). There may also be stigma effects, resulting in economic and psychological consequences to being called disabled. The empirical strategy in this paper identifies the net benefit of SE services, but cannot identify the separate contribution of each of these factors.

3 CPS Accountability Policy

In September 1996, CPS announced a new policy that held schools accountable for student performance. Elementary school performance was measured by the fraction of students reading at grade-level norms (i.e. 50th percentile) on the Iowa Test of Basic Skills (ITBS).[11] The ITBS is a nationally-normed test given every May to elementary school students, which in CPS includes grades 1 to 8. Schools in which fewer than 15 percent of students were reading at national norms ($Pnorms$) on the ITBS were placed on probation. Probationary status was based on scores from the previous May and in September of 1996 CPS announced that it was placing 71 of its 475 elementary schools on probation (Finnigan and O'Day 2003).

Schools on probation received a combination of increased assistance and monitoring, a loss of autonomy and the threat of more severe sanctions in the form of reconstitution.[12] Modest resources were given for teacher training and after-school programs. The central component of probation was the requirement that schools work with an "external partner" (individual consultants, university scholars and experienced administrators) who made recommendations for and oversaw professional development and school improvement. In order to be taken off of probation, schools had to increase the fraction of students reading at national norms to 20 percent. Schools with 1996 scores between 15-20 percent were also closely monitored and about 25 percent of them were placed on "remediation."[13]

Although many students showed marked gains in math and reading scores on the ITBS, it is unclear how much these gains reflect true increases in student achievement. Jacob (2005) illustrates that the new policy did not lead to comparable increases in achievement on a state-administered, low-stakes exam or on subjects in the ITBS that were not included in accountability measures. Improved scores were partly the result of increased stamina and test-specific skills. Teachers and administrators report experiencing enormous pressure to meet the new accountability standards (Jacob, et al., 2004). One way for faculty and administrators to improve school performance was by increasing SE placement. Since SE students' scores were not included in accountability measures, schools could mechanically improve performance by taking students

[11] High school performance was measured similarly, using the Test of Academic Proficiency.

[12] Reconstitution refers to removal of teaching and administrative staff and the potential closing of the school. CPS never moved to reconstitute any elementary schools.

[13] Of those elementary schools placed on probation in 1997, 75 percent were still on probation two years later. Of those placed on remediation in 1997, 70 percent went on probation the following year.

performing below the 50th percentile out of the testing pool and putting them in SE. Under the heightened scrutiny, all schools were concerned with performance and had an incentive to increase SE placement. Schools close to the probation/remediation cutoff, however, had stronger incentives to increase SE enrollment. That is, a school with *Pnorms* equal to 15 percent had stronger motivation to increase SE placement than schools with *Pnorms* at 5 or 25 percent, since these marginal schools could potentially affect probationary status.

Figure 3 plots the growth in Special Education enrollment between 1994-2000 for elementary schools close to the cutoff (*Pnorms* between 10-20 percent) and adjacent schools above and below (*Pnorms* between 0-10 percent and between 20-30 percent).[14] While there is barely any growth in SE between 1994-1996, SE in all of these schools grew rapidly after 1996. Consistent with the idea that schools near the cutoff had stronger incentives to alter the testing pool, these schools had faster growth in SE enrollment after 1996—growing by nearly 40 percent—than other schools. Figure 4 provides additional evidence that the growth in SE following the introduction of accountability occurred at "high-incentive" schools close to the probation cutoff. This figure plots the change in SE enrollment from pre- to post-accountability for schools ranked by 1996 *Pnorms*. The change in SE is most pronounced for schools near the probation cutoff and is lower for schools with very low or very high 1996 *Pnorms*.

The analysis will focus on the group of schools close to the disciplinary cutoff and adjacent schools, since better performing schools (and their Special Education programs) are different in a number of ways. Table 1 presents summary statistics for schools categorized by 1996 performance. Higher performing schools have a much smaller fraction of minority and low income students. Partly due to differences in SE resources and partly to population differences, higher performing schools also treat more students for speech problems and physical disabilities and fewer students for emotional and mental disabilities.

[14] Of the 27 elementary schools that had 0-10 percent of students reading at national norms, all but 2 were placed on probation in 1997. Of schools with 10-15 percent of students at norms (54 schools), 45 were placed on probation and the other 9 were placed on remediation in 1997. 15 of the schools in the 15-20 percent category (62 schools) were on remediation in 1997.

To succinctly capture the idea that schools closest to the probation cutoff have the strongest incentives to increase SE placement, I use the following formula for the cross-sectional component of the instrument:

$$I96 = 1 - \frac{|17.5 - Pnorms96|}{17.5} \tag{1}$$

This variable is equal to one for schools in which 17.5 percent of students were reading at national norms in 1996 and declines linearly from 17.5 to 0 percent and from 17.5 to 35 percent (illustrated in Figure 5). The 20 percent of CPS elementary schools with $Pnorms96 > 35$ are not included (this leaves 360 out of 450 schools in the sample). The instrument peaks at 17.5 percent, halfway between 15 and 20 percent, because schools at 15 had to increase performance by 5 percent to be removed from probation and schools at 20 were 5 percent away from being placed on probation.

4 CPS Data and Summary Statistics

The analysis uses individual-level panel data from the administrative files kept by CPS between 1994-2004. The dataset includes detailed information on student background, such as school, zip code, census tract, demographic information and free lunch, bilingual and Special Education status (as well as disability type). Student records also include test scores, GPA, attendance and credits accumulated. Student data is linked to aggregate school data such as racial composition, class sizes, and the fraction of students who are receiving a free or reduced-price lunch, limited-English-proficient and low income. CPS assigns each student an identification number that allows me to follow them over time, as long as they remain in the CPS system.[15]

The sample is composed of cohorts of students who are 13-years-old in September of each year. I only include students in regular and magnet elementary schools with $0 \leq Pnorms96 \leq$

[15] Unless they return to CPS, students who transfer out of the CPS system are not included in the analysis since I cannot observe their outcomes. There is no change in the rate of transfer after accountability. The accountability policy does appear to have marginally reduced outflows to private schools, but this change is among high-achieving students at high-performing schools and should not affect the analysis.

35.[16] For most students, this is the fall of their eighth grade year. I follow students aged 13 in years 1994-2000 through high school, observing GPA and attendance records at age 16 and dropout status at age 17.[17] The sample consists of roughly 20,000 13-year-olds each year, for a total of about 140,000 students. Each September, a student is recorded as active or inactive in the CPS system. When a student is inactive, they are assigned a leave code documenting the reason for their exit, including whether they have dropped out and for what reason. I follow the procedure outlined in Allensworth and Easton (2001) for calculating CPS cohort dropout rates, and record whether or not a student is still in school or has graduated by age 17 (i.e. four years after they enter the sample). GPA and attendance are measured in the fall semester of the year in which the student is 16. These last two outcomes are missing for some students, primarily for those who have dropped out by age 16, but dropout status is available for all students.

For the specifications using prior student achievement, I calculate decile scores on the ITBS reading section at ages 9 and 10.[18] Achievement is measured at this age to avoid the endogeneity of test scores with respect to the accountability policy.[19] Since achievement at age 10 is also potentially endogenous to SE placement, I run all specifications with and without the sub-sample of students who are in SE at age 10. I will show that their inclusion only increases precision with

[16]Elementary schools in CPS include students in grades 1-8 or K-8, and combination elementary/high schools include grades 7-12. I include 13-year-olds who meet the following conditions: 1) in a regular or magnet elementary school or combination elementary/high-school; 2) in a school that was subject to the accountability policy (this excludes charter schools, Special Education schools and other alternative schools); and 3) in a school with $Pnorms$ data in 1996 (this excludes 13-year-olds in 7 schools). A tiny fraction of the sample is thrown out because of inconsistent age data or missing data on whether or not the student is in SE at age 13. Inclusion in the sample is not at all based on where the student goes to high school or transfers after age 13.

[17]Since roughly 40 percent of CPS 17-year-olds have dropped out, I measure GPA and attendance at an earlier age to increase the number of students with these outcomes in my sample.

[18]CPS reports ITBS scores in grade equivalents, which are supposed to reflect a student's ability relative to his or her grade-level. However, there are a number of recognized shortcomings with using this metric to make comparisons over time and across grades. For example, these scores do not correct for changes in test form and difficulty across years. The scores used in this study are Rasch measures, which have been constructed by the Consortium on Chicago School Research and used in other studies of the CPS accountability policy such as Jacob and Lefgren (2002). These scores are constructed from an item-response model and measure the log-odds that a student of a given ability level answers a question of a given difficulty level correctly. Since the average student answers the question of average difficulty 50 percent of the time, this metric is centered on zero. I thank the Consortium on Chicago School Research for providing these Rasch measures.

[19]Students who are 13 in 2000 will have had one year of exposure to the policy, so their age 10 (but not age 9) test scores are potentially affected by the policy. However, using the average of scores at ages 9 and 10, rather than just scores at age 9, does not meaningfully affect the results.

no impact on the nature of the results.[20] About 3,000 students per year are dropped because they are missing age 10 test score data. For those students missing age 9 test scores, scores are predicted using age 10 scores and other background, school and neighborhood characteristics.[21] This leaves a sample of 127,000 students (105,000 with test score data) composed of roughly 17 percent SE students.

Since the majority of CPS students are quite disadvantaged (e.g. 90 percent of students are eligible for a free or reduced price lunch) the differences between SE and RE students are not as dramatic in Chicago as they are in many other districts (e.g. Blair and Scott 2002). Nevertheless, there are many reasons to suspect that students are negatively selected into SE and some of this is evident in Table 2, which presents mean characteristics for 13-year-old SE and RE students in CPS. Achievement discrepancies at age 10 are large: students who will be placed in SE at age 13 score nearly 3 deciles below students who will not. Future SE students are also roughly 1/5 of a grade level behind other students at age 10. Students in SE are more likely to be African-American and to be receiving a free or reduced price lunch. SE students are less likely to be Hispanic and bilingual. Some of the discrepancy between rates of SE for African American and Hispanic students likely comes from the ability of schools to provide extra help to Hispanic students who are struggling through Bilingual and Limited-English-Proficiency programs. Differences between school and neighborhood characteristics are minimal. There is virtually no difference between the poverty rates, class size and average *Pnorms* of the elementary school an RE student and an SE student attends. Similarly, there appears to be no notable difference between census tract poverty, employment and education rates for SE and RE students. On the basis of average characteristics, CPS students appear to

[20] The dataset does not include the year of SE placement so for younger cohorts I do not know with certainty whether or not the student was in SE at age 10. However, I know whether age 10 test scores were excluded from school reporting and this appears to be a good measure of SE status. Roughly 90 percent of 10-year-olds in SE for a mental or emotional disability have test scores excluded from reporting.

[21] Missing age 10 scores does not appear to differ by SE status at age 13. There are a number of reasons why it is not possible to predict age 10 scores for those missing data. First, if a student is missing age 10 scores, I cannot know whether her scores were excluded and thus whether or not she was in SE at age 10. Second, it is not possible to predict age 10 scores with scores at older ages since the latter will be endogenous to SE placement and to the accountability policy.

be selected into SE overwhelmingly on the basis of achievement.

5 Empirical Strategy

Let y be some high school outcome, SE be an indicator for being in Special Education at age 13, X be a vector of individual background variables (including an intercept) and Z be a vector of time-varying school characteristics. The SE program impact is captured by the β coefficient in the regression:

$$y_{ist} = SE_{ist}\beta + X_{ist}\eta_1 + Z_{st}\lambda_1 + \alpha_s + \gamma_t + \varepsilon_{ist} \tag{2}$$

where i, s and t indicate individual, elementary school and year, respectively, α and γ are elementary school and year fixed effects, respectively, and ε is a stochastic error term.

The outcome variables I consider are dropout status at age 17, and absenteeism and GPA at age 16. The purpose of SE is not only to improve achievement, but to improve educational attainment and help disabled students remain engaged in school. Dropout status at age 17 is a good measure of the overall net benefit of SE since it is highly correlated with both achievement and educational attainment. Since SE can in some cases loosen curricular requirements and promotional standards for students, one may be concerned that SE appears to reduce dropout rates only because it makes it easier to graduate. While I cannot test whether this is the case, I can estimate the impact of SE on high school absenteeism, which serves as a good measure of the impact of SE on school engagement. GPA is also an imperfect measure of the impact of SE, since SE students may be graded according to different standards than RE students. This is unlikely to be a serious concern for the marginal student in this study, however, since SE services for mild learning disabilities are minimal and school districts must adhere to the regular curriculum as much as possible.

I have already discussed a number of reasons why the main requirement for consistent estimation of β (i.e. $Cov(SE_{ist}, \varepsilon_{ist}|X_{ist}, Z_{st}) = 0$) is unlikely to be met here. Factors influencing SE selection such as parental characteristics, behavioral problems and poor infant health that may also influence high school outcomes are not observed and could bias estimates of β. [22]

Rather than comparing outcomes for students who were and were not in SE in elementary school, the estimation of β requires an instrument that predicts SE placement but, conditional on covariates, is uncorrelated with high school outcomes. Students in elementary schools after accountability were more likely to have been placed in SE than previous students, and the strategic nature of this response to accountability caused students in high-incentive schools to experience this increase most. Further, the increasing probability of SE placement was most pronounced for low-achieving students. I thus run two basic first-stage specifications. The first specification analyzes whether students in high-incentive schools experienced larger increases in SE placement after accountability, while the second specification analyzes to what extent this increase was experienced differentially by achievement level.

The basic first stage regression without prior-achievement is:

$$\begin{aligned} SE_{ist} &= \pi(I96_s \cdot POST_t) + X_{ist}\eta_2 + Z_{st}\lambda_2 + \kappa(Pnorms96_s \cdot POST_t) \\ &\quad +\theta(probation_s \cdot POST_t) + \alpha_s + \gamma_t + \zeta_{ist} \end{aligned} \quad (3)$$

where $I96$ is from (1) and is equal to one for elementary schools in which the fraction of students reading at national norms in 1996 is equal to 17.5 (i.e. it is constant within a school) and $POST$ is an indicator variable equal to one after 1996 and zero otherwise. The $(Pnorms96_s \cdot POST_t)$ term is included to pick up any changes in SE placement that vary linearly with school performance.The probation variable is a dummy equal to one for schools with $Pnorms96 \leq 15$ and the inclusion of this variable interacted with the $POST$ dummy is intended to control for any changes in SE placement induced by probation status. The baseline specification does not include individual and school-level covariates, and I will illustrate that their inclusion does not

[22] Some of this bias could be eliminated with an individual fixed effect but this is not practical in the estimation of the impact of SE. First, there is not much outward mobility from SE (except for speech and language disabilities, which are not the focus of this paper) and those who do leave SE are not representative. For example, roughly 90 percent of the first graders I observe in the CPS data are still in SE in eighth grade. Second, much of the selection bias that needs to be purged does not come from fixed characteristics of individual students. Rather, many students are identified as needing SE because of changes in their behavior and performance, which often result from changes in the child's environment.

significantly impact the estimate of π.

It is important that the new policy was announced after 1996 testing was completed. This allows me to construct an instrument that makes use of school performance prior to the policy, which is necessary since performance is not only endogenous to the policy change, but is also a function of SE placement itself. A positive and significant estimate of π indicates that students in high-incentive schools had larger increases in the probability of SE placement after accountability than students in schools with weaker incentives. I will illustrate through a series of specification checks that rates of SE enrollment did not differ for these schools prior to the accountability policy.[23]

This identification strategy uses the year and school in which a child is 13 to predict SE status. Using cohorts of 13-year-olds has the advantage of purging variation in SE status by age. Since SE status is basically an absorbing state, most children who were ever in SE are still in it at age 13. Data limitations prohibit me from using the number of years in SE as the independent variable, so the estimates presented here will be for the 13-year-old with an average number of years in SE. Measuring the dependent variable at age 13 also enables me to measure outcomes at age 17—I would not be able to observe long-term outcomes for younger students after the accountability policy.

In order for Equation (3) to yield an estimate of the direct impact of SE, it must be that the only change taking place at high-incentive elementary schools influencing future outcomes was the differential change in the probability of SE placement. Strong strategic incentives could have led schools to make other changes, however, so it is important to utilize within-school differences in SE placement. As illustrated in Table 2, the test scores of 10-year-old regular

[23] Another reason to use age cohorts is to avoid changes in grade cohort composition resulting from a change in CPS's social promotion policy in 1996. The policy required third, sixth and eighth graders that did not meet certain test score cutoffs to go to summer school. When they were retested at the end of the summer, students who again did not meet the CPS standard were retained. SE students were not subject to the new promotion standards, but many of their low-achieving peers were. Using age cohorts allows me to avoid changes in cohort composition, but the change in social promotion still introduces a concern regarding the exclusion restriction. As long as retention is not correlated with the strength incentives, however, changes in social promotion should not be biasing the results.

education students who will be placed in SE by age 13 are much lower than the scores of those who will not. Thus, prior student achievement should be a good predictor of SE placement within a school. The following specification measures the impact of being in a high-incentive school on SE placement by prior achievement level, measured by decile reading score at age 10. For the specifications using within-school variation to yield the direct impact of SE it must be the case that, in the absence of the increase in SE placement, students of different achievement levels in high vs. low-incentive schools would not have had differential changes in high school outcomes. The exclusion restriction could be violated if high-incentive schools targeted resources differentially across achievement levels. I will illustrate in the robustness section that this does not appear to be the case.

The basic specification incorporating prior student achievement is:

$$\begin{aligned} SE_{ist} &= \pi_1(dec_d^2 \cdot I96_s \cdot POST_t) + \pi_2(dec_d \cdot I96_s \cdot POST_t) + \pi_3(I96_s \cdot POST_t) \\ &\quad + \tau_1(dec_d^2 \cdot I96_s) + \tau_2(dec_d \cdot I96_s) + \sigma_d + \alpha_s + \gamma_t + \sigma_d \cdot \gamma_t + \zeta_{ist} \end{aligned} \tag{4}$$

Where *dec* is a variable ranging from 1-10 indicating a student's decile reading score, σ is a decile fixed effect, and all other variables are as defined above.[24] In this regression, a negative and significant estimate of the coefficient π_2 indicates that lower-achieving students in high-incentive schools had the greatest increase in SE placement. A positive estimate of the coefficient π_1 implies that this effect is greatest at the lowest deciles. I also estimate a more flexible version of (4), interacting decile score dummies with $(I96_s \cdot POST_t)$, in order to characterize the change in SE placement across the achievement distribution. In the next section I present coefficient estimates from these basic first stage specifications as well as a number of specification checks.

[24] Inclusion of $Pnorms96 * POST$ and $Probation * POST$ and their interactions with decile score does not affect the results.

6 Impact of Accountability on Special Education

It is evident from Figure 2 that SE grew dramatically in CPS after the introduction of accountability. Table 3 illustrates that this increase was largest for students in schools with strong strategic incentives. Column (1) presents the coefficient of interest from the basic specification (Equation 3), with standard errors clustered by school*year. I find a significant coefficient of .048 on the instrumental variable ($I96_s \cdot POST_t$). Given that the mean of the dependent variable (SE status) is .113 and that the mean of $I96$ is about .63, this coefficient estimate implies that a student at the average school experienced a 3 percentage point (29 percent) increase in the probability of SE placement. The impact of accountability was even larger for students in high-incentive schools. The probability of being in SE grew by 4.8 percentage points more for students in high-incentive schools than for other students. This represents a differential increase of 43 percent over just 7 years, a very large change for a variable that typically does not fluctuate much from year to year. Since the average school has about 75 13-year-olds, these estimates imply that the number of SE students increased from 9 to 10.5 in the average elementary school and from 9 to 12 in the high-incentive elementary schools.

The rest of the results in Table 3 serve as specification checks. Column (2) estimates the same basic specification, but includes students who were already in SE at age 10. The point estimate is slightly larger but not significantly different from the basic estimate in Column (1). Columns (3)-(5) again estimate the basic specification in Equation (3), but include individual background controls, time-varying school controls and Zip Code*Year fixed effects, respectively. Their inclusion does not affect the magnitude or significance of the coefficient estimate much and these controls will be even less important in the specifications incorporating test scores. Thus, it is possible that the accountability policy altered the population of students in some way other than the increase in SE (e.g. by bringing higher income students back to CPS), but these changes were either not correlated with changes in SE enrollment or were not experienced differentially by high-incentive schools.

The estimates in Columns (6)-(11) explore whether high-incentive schools had different pre-trends in SE enrollment and whether or not the form of the instrument is specified correctly. A useful variant of (3) is:

$$\begin{aligned} SE_{ist} &= \pi(I96_s \cdot POST_t) + \kappa(Pnorms96_s \cdot POST_t) + \theta(probation_s \cdot POST_t) \\ &\quad + \phi(I96_s \cdot d_{1996}) + \alpha_s + \gamma_t + \zeta_{ist} \end{aligned} \tag{5}$$

where d_{1996} is a dummy variable equal to one in 1996. The coefficient on $(I96_s \cdot d_{1996})$ allows me to check for pre-trends in SE in schools closer to the incentive region. Including $(I96_s \cdot d_{1996})$ in the basic specification (Column (6)) does not change the coefficient of interest and itself has a tiny and insignificant coefficient estimate. The coefficient estimates in Column (9) are from a more flexible form of Equation (3), which includes interactions between year dummies and $I96$ (the omitted category is $(d_{1994} \cdot I96_s)$). Prior to accountability the instrument does not predict SE placement at all, but the coefficient estimates become increasingly large and significant after 1996. These coefficients and their confidence interval are plotted in Figure 6.

One last specification check explores whether the impact of incentives on SE placement differs if we look at the “upper” (i.e. $Pnorms > 17.5$) or "lower" ($Pnorms \leq 17.5$) sides of the instrument. Finding similar estimates for each side of the instrument would be an encouraging indication that the correct structure has been put on the incentive effect of accountability and that the coefficient of interest is not being disproportionately weighted by certain schools. When both sides of the instrument are controlled for separately—allowing different slopes for schools with *Pnorms96* below and above 17.5—very similar coefficient estimates are recovered (an F-test that these coefficients are the same cannot be rejected). This suggests that the impact of strategic incentives on SE placement for high-incentive schools does not differ notably if we compare them to low-performing low-incentive schools or high-performing low-incentive schools.

Table 4 gives estimates of the impact of accountability on SE status for students of various backgrounds and disability types. The only ethnic group to experience a large and statistically significant increase in SE placement was African-American students. The probability of being in SE grew 34 percent more for African-American students in high-incentive schools than for those in low-incentive schools. Growth in SE enrollment was also very pronounced for girls. The probability of SE placement grew by a full 50 percent for girls in high-incentive schools, nearly 6 percentage points more than for girls in schools with weaker incentives.

The effect for Hispanic students is insignificantly negative. Some of this is likely due to the fact that high-incentive schools are roughly 70 percent African-American and have comparatively few Hispanic students. However, given that there is some evidence of a decline in SE placement for bilingual students (Column (5))—who are overwhelmingly Hispanic—there appears to be a strategic element to this as well. This is because most bilingual students' scores could also be excluded from accountability measures, so it was less beneficial for schools to place Hispanic students in SE. The finding that black students are targeted most when SE placements are strategic is consistent with Cullen's (2003) result that SE placements in response to fiscal incentives are largest for black students.

If schools were using SE placement strategically, we would expect them to place students in SE largely on the basis of achievement. Coefficient estimates in Columns (6) and (7) imply that, while SE placement grew by 35 percent for students with age 10 reading scores in the bottom quartile, there was no growth in SE placement for students scoring in the top quartile. SE placement for an emotional or mental disability grew by 2.9 percentage points (36 percent) for students in high-incentive schools, while there was no significant differential increase in the probability of placement in SE for a physical disability. The coefficient estimate in Column (8) implies that students in high-incentive schools who were already in SE at age 10, were less likely than students at other schools to still be in SE at age 13. This means that flows out of SE were higher—and SE students were somewhat older—at schools with strong-incentives.

Results presented in Table 5 explore how changes in SE placement differed by prior achievement and are estimated from the triple-difference specification in Equation (4). The point estimate from the most parsimonious specification, presented in Column (1), implies that SE was increasing fastest for students at high-incentive schools and that the probability of placement was decreasing at an increasing rate with decile score. I find no impact of incentives on

SE placement for students in the fifth decile, but find a large, significant effect for students scoring in the bottom decile. Coefficient estimates barely change at all across specifications and are again similar when the lower and upper sides of the instrument enter the specification separately. These estimates imply that the probability of being in SE grew by 11 percentage points more for bottom decile students who attended elementary schools with strong incentives than for students attending schools with weaker incentives. The average fraction of students in SE who scored in the bottom decile is 55 percent, so my estimates imply that these low-performing students experienced a 20 percent relative increase in SE placement.

The coefficients in Table 6 are from a more flexible form of Equation (4), in which reading decile dummies are interacted with the incentive measure. The coefficients from the basic specification are in Column (1) and are plotted along with a 95 percent confidence interval in Figure 7a. These same coefficients are reproduced in Figure 7b along with a plot of the increase in SE placement for each decile predicted by the triple-difference specification (Equation 4). The probability of being in SE increased substantially for students in high-incentive schools who scored in the bottom decile at ages 9 and 10. In all specifications, SE placement probability declines rapidly as decile score increases, and the impact of incentives on SE placement for higher achieving students is small and insignificant. The close overlap between the actual change in SE placement at each decile and the change predicted by the triple-difference specification suggests that a quadratic or higher-order polynomial is a reasonable functional form for capturing the relationship between incentives, prior achievement and SE enrollment. The next section explores whether this large increase in the probability of SE placement is reflected in outcomes for low-achieving students.

7 Impact of Accountability on High School Outcomes

Reduced form estimates of the impact of being in a high-incentive school on dropout status, absenteeism and GPA are presented in Table 7. They are estimated from variations of Equation (3). The reduced form estimates from these specifications are generally noisy, but suggest a slight improvement in dropout probability and absenteeism and no impact on GPA for students in high-incentive schools. When the sample is limited to students scoring in the bottom quartile at age 10, the part of the distribution in which SE increased the most, the results for dropout and absenteeism remain. The flexible specifications estimating year*I96 coefficients indicate a downward trend in dropout probability and absenteeism after accountability, with the estimates becoming more significant after 1997. Again, there is no discernable trend for GPA.

As with the first stage estimates, identifying the impact of strong incentives across the achievement distribution adds precision to the reduced form estimates. Table 8 presents coefficient estimates from the reduced form version of Equation (4). The results for dropout status in Column (1) indicate that dropout rates were decreasing fastest for low-achieving students in high incentive schools. I estimate no impact for students achieving at the fifth decile, but a larger negative coefficient for students at the bottom decile. The standard errors are large, but estimates are significant in about half of the specifications. Estimates for dropout status are similar across specifications and indicate that bottom decile students in high-incentive schools experienced a roughly 4 percentage point decrease in dropout probability relative to other bottom decile students. Since the average dropout probability for bottom decile students is about 46 percent, these estimates imply that being in a high-incentive school reduced the probability of dropout for these students by about 8-9 percent.

Panel B of Table 8 presents reduced form estimates for GPA by decile score. As with dropout rates, there is no effect on GPA at the fifth decile. Point estimates at the bottom decile are positive, which is consistent with the positive impact found for dropout rates, but are small and insignificant. I cannot reject the null hypothesis of no difference in GPA between students at high and low-incentive schools. Results for absenteeism are presented in Panel C. The coefficient estimate in Column (1) implies that absences (in the fall semester) increase by roughly 1 day with decile score. As with dropout rates, the coefficient on the square of decile score is negative, implying that as decile score increases, absenteeism increases at a decreasing rate. Thus in the region of the achievement distribution in which SE is increasing the fastest, absenteeism and dropout probability are decreasing fastest. Coefficient estimates for absenteeism at the mean decile are small and inconsistent across specifications, but estimates at the bottom decile are statistically significant at roughly -2. The average number of days absent for bottom decile students is 14, so these estimates imply that bottom decile students at high-incentive schools experienced a reduction in absenteeism by about 15 percent more than students at schools with weaker incentives.

Reduced form estimates from the flexible specification interacting $(I96_s \cdot POST_t)$ with decile score dummies are presented in Table 9 and plotted in Figures 8a-8c. The coefficients are noisy but confirm that dropout rates and absenteeism are decreasing for the lowest deciles.

8 OLS and IV Estimates of the Impact of SE on High School Outcomes

It was apparent from comparing background, school and neighborhood characteristics in Table 2 that CPS students in SE are not significantly more disadvantaged than RE students. They do, however, have much lower test scores and are more likely to have been retained than their RE peers. One might expect these large differences in achievement to show up in high school outcomes and this is reflected in the OLS estimates in Column (1) of Table 10. SE students are 10 percent more likely to have dropped out of high school at age 17, have a .5 lower GPA and are absent nearly 3 days more per semester than RE students. However, OLS differences for SE and RE students with test scores in the bottom two deciles (Column (2)) are much smaller and, in the case of dropout, are not different at all.

The fact that OLS differences are nearly eliminated when only controlling for test scores at age 10 suggests that OLS estimates of the impact of SE are biased by selection. I find IV estimates that support this hypothesis. Coefficients in Columns (3) – (5) are estimated from the quadratic specification defined in Equation (4) and Columns (6) – (7) are from a cubic version of this specification. These estimates are consistent with SE placement in elementary school reducing the probability of dropping out of high school by about 35 percent. While I cannot reject the OLS results in Column (2), most of the IV results are significantly different from the basic OLS estimates in Column (1). Some of the specifications are borderline significant, implying that we can reject the null that SE has no impact on dropout at the 10-20 percent level. Results for absenteeism are more precise, generally implying that SE placement in elementary school reduces absenteeism by about 10 days per semester. This point estimate is very large, and is probably picking up the reduced dropout probability, since some students who drop out of CPS are not officially recorded as having dropped out until after they are absent extensively. This explains why there are some students who have more than 40 days—about 2/3 of a semester—absent. IV results for absenteeism when the 5 percent of the sample with more than 40 days absent is thrown out suggest that SE reduces absenteeism by a more reasonable 5 days per semester. Results for GPA are noisy and highly insignificant, and I cannot reject OLS estimates that the impact of SE on high school GPA is negative.

IV estimates in Columns (5) and (7) include *I96*Post* as an instrument, in addition to its interaction with test scores, while all of the other columns include only the latter. The fact that IV estimates are basically unchanged when cross-school variation is used in addition to within-school variation is an encouraging indication that there were not other important changes going on within high-incentive schools that are biasing the results. Columns (4) and (6) use the upper and lower sides of $I96$ as separate instruments. IV estimates from these specifications are very similar to the baseline specifications, but are more precise, which suggests that the model is correctly specified.

9 Robustness

The main concern about the validity of the IV estimates is that, in addition to the large increase in SE enrollment, high-incentive schools increased resources for low-performing students. One might expect from the previous literature on high-stakes-testing in Chicago that much of the improvement in test scores was the result of gaming and an improvement in test-specific-skills. Also, since I have provided strong evidence that schools were acting strategically and were very aware of how to improve their standing, if schools were to target resources one would expect it to be toward middle-achieving, rather than low-achieving, students.

It is possible, however, that schools with strong incentives generally improved academic achievement among low-performing students. In order to explore this possibility, I present the reduced form results after controlling for ITBS scores at age 13. These scores should reflect the maximum impact of strategic incentives on achievement, since any changes high-incentive schools undertook would have been directed toward increasing achievement on the test used for accountability. If achievement increased among low-achieving students, and these changes in achievement influenced high school outcomes, then one would expect the IV estimates to decrease in magnitude after controlling for age 13 test scores. The problem with simply controlling for these scores is that they are also a function of SE status and this is a channel that we do not want to eliminate from the IV estimates. While it is not appropriate to control for age 13 scores in the main specifications, controlling for these scores in the reduced form provides an important robustness check.

Reduced form estimates from the basic triple-difference specification (Equation (4)) after controlling for achievement at age 13 are presented in Table 11.[25] The first and second columns of each panel give coefficient estimates after controlling for reading decile score and reading decile score fixed effects, respectively. These coefficients are very similar to the main reduced form estimates (Table 8), again predicting no effect of strong incentives on outcomes at the 5th decile but an improvement in outcomes for students in the bottom decile. In fact, controlling for endogenous achievement *increases* the magnitude of the reduced form effect at the bottom decile (e.g. from 4 to 5 percentage points for dropout rates and from -2 to -2.5 days for absences). These results suggest that the improvement in outcomes for low-achieving students in high-incentive schools cannot be explained by increasing achievement.

It is also important to consider how the increase in SE placement induced by accountability may have influenced resources for both SE and RE students. One possibility is that the quality

[25] The within-school difference is still based on test scores at age 10.

of SE services is declining, e.g. from over-crowding of SE classes and fewer resources to spend on SE materials. It could also be that resources are being diverted away from regular education programs in order to support higher SE enrollment. Both of these changes would imply that I am underestimating the positive impact of SE. The overwhelming majority of Special Education students in Chicago Public Schools are classified as learning disabled. This is the least costly disability to treat and the majority of the cost is devoted to instructional services (Chambers, et al. 2003). Funding for SE students is legally mandated and can be taken away from other programs in the school budget. Expenditures on instructional services make up the majority (66 percent) of SE program costs for LD students (Chambers et al., 2003). Illinois state law mandates a maximum student-teacher ratio of 20 in resource rooms and schools need to apply for exemptions to this law. I estimated high-incentive schools to have increased the number of 13-year-old SE students from about 9 to 12. If other age groups experienced similar increases, the change in the number of SE students likely necessitated the hiring of one additional SE teacher. In most states, districts are responsible for about 45 percent of SE program expenses (another 45 percent and 10 percent are paid by the state and federal government respectively). State funding in Illinois is based on actual SE program expenses, which is generous relative to most other states that try to limit over-classification for SE by reimbursing on the basis of expected costs. Thus, the increase in SE placement in high-incentive schools may have led to some reduction in resources for RE students, but the impact was likely minimal. Without school-level budget information, however, I am unable to estimate the impact of additional SE costs on resources for RE students or on the quality of SE services.

If student-teacher ratios are declining for RE students, either because fewer SE students are in the classroom or because a teacher's aid is brought into the classroom to assist mainstreamed SE children, I could be overstating the positive impact of SE. All SE students—and particularly those with mild learning disabilities—are by law required to be in the regular classroom learning the regular education curriculum to the greatest extent possible (Martin, et al. 1996).[26] Students

[26] This is outlined in the "least restrictive environment" clause of the 1997 Individuals with Disabilities Act.

with LD typically spend 3-4 hours per week in a resource room getting extra assistance for the subject they are struggling in and the rest of the time in the regular classroom. It is therefore not likely that regular education students were benefiting much from smaller student-teacher ratios due to SE students being taken out of the classroom. If the majority of new SE students at high-incentive schools were mainstreamed it is possible that RE students were benefiting from the increase in SE classification. However, one would expect the benefit of lower student-teacher ratios to be reflected in achievement scores and I find in Table 11 that changes in achievement cannot explain the improvement in outcomes for low-achieving students.

A natural question that arises in any IV analysis is how representative the marginal student is of the average student. In this case, we want to know whether the impact of SE for the marginal student placed in SE because of strategic incentives is similar to the impact for those placed in the absence of these incentives. Whatever the decision criteria are prior to accountability, the policy change clearly leads high-incentive schools to put more weight on ITBS scores in its decision. As discussed in Section 2, SE placement criteria are vague and leave room for discretion, but schools are still required to demonstrate the likely presence of a disability and so are not completely unrestrained in "strategically" placing students in SE. I have found no evidence that these students are any "more" or "less" disabled than previous SE students. It could be that the net benefit is different for students placed in SE for strategic reasons than for traditional SE students. Hanushek et al. (2002) provide suggestive evidence that SE may be targeted toward students who will benefit most from the program. If the focus on ITBS scores causes schools to place students in SE who benefit less from the program, then the results presented here may be underestimates of the beneficial impact of SE.

CPS has been very cautious about removing SE students from the regular classroom since the loss of the Corey H. class-action suit (1992) in which it was accused of illegally segregating SE students.

10 Conclusion

Observational studies of SE students find that they are much more likely to suffer poor social and academic outcomes than their regular education peers. However OLS estimates of the impact of Special Education cannot be interpreted as causal since they are likely to be biased by omitted variables in the selection process for SE. This paper provides the first quasi-experimental evidence of the impact of SE placement on high school outcomes. I illustrate that low-achieving students benefit from SE placement for mild mental disabilities. IV estimates indicate that SE placement reduces absenteeism and the probability of dropping out of high school. I find no conclusive evidence that SE services impact high school GPA.

While these results suggest that students benefit from SE services in terms of educational attainment and engagement in school, they do not speak to the academic impact of SE. Absenteeism and educational attainment are typically highly correlated with achievement, but SE students face alternative curricular requirements, assessment metrics and promotional standards. It is therefore unclear whether the reduction in dropout probability is stemming from enhanced achievement or from easier graduation requirements. While my results indicate a net benefit to SE placement, further investigation of the mechanisms by which SE influences outcomes is necessary for understanding how to design SE services optimally. I also do not have detailed enough expenditure data to analyze the cost-effectiveness of SE services relative to other interventions for struggling students. Evidence on the relative efficiency of these programs is clearly essential to understanding how effectively SE services are delivering resources to children at risk of negative outcomes.

References

[1] Algozzine, B. and J. Ysseldyke. 1986. "The Future of the LD Field: Screening and Diagnosis." Journal of Learning Disabilities 19: 394-398.

[2] Chambers, J. G., Shkolnik, J. and M. Perez. 2003. "Total Expenditures for Students with

Disabilities, 1999-2000: Spending Variation by Disability." Special Education Expenditure Project Report.

[3] Cullen, J. B. and R. Reback. 2006. "Tinkering Toward Accolades: School Gaming Under a Performance Accountability System." NBER Working Paper # 12286.

[4] Cullen, Julie Berry. 2003. "The Impact of Fiscal Incentives on Student Disability Rates." Journal of Public Economics 87(7-8): 1557-89.

[5] Cullen, J. B. and S. G. Rivkin. 2003. "The Role of Special Education in School Choice" in C. Hoxby, ed., The Economics of School Choice, Chicago, IL: University of Chicago Press, pp. 67-106.

[6] Department of Education. 2005. Digest of Education Statistics.

[7] Figlio, D. and L. Getzler. 2002. "Accountability, Ability and Disability: Gaming the System." NBER Working Paper # 9307.

[8] Finnigan, K. and J. O'Day. 2003. "External Support to Schools on Probation: Getting a Leg Up?" Consortium on Chicago School Research. Chicago, IL.

[9] Hanushek, E., J. Kain and S. Rivkin. 2002. "Inferring Program Effects for Specialized Populations: Does Special Education Raise Achievement for Students with Disabilities?" Review of Economics and Statistics 84(4): 584-599.

[10] Illinois State Board of Education, Special Education Unit. 2006. "Local Education Agency Policies and Procedures to Govern the Administration of Services to Students with Disabilities." As of June 2006 this document could be found at http://www.isbe.net/spec-ed/pdfs/policies.pdf.

[11] Jacob, Brian. 2005. "Accountability, Incentives and Behavior: Evidence from School Reform in Chicago." Journal of Public Economics 89(5-6): 761-796.

[12] Jacob, B. and L. Lefgren. 2004. "Remedial Education and Student Achievement: A Regression-Discontinuity Analysis." Review of Economics and Statistics 86: 226-244.

[13] Jacob, B. and L. Lefgren. 2004. "The Impact of Teacher Training on Student Achievement: Quasi-Experimental Evidence from School Reform Efforts in Chicago." Journal of Human Resources 39(1): 50-79.

[14] Jacob, R. S. Stone and M. Roderick. 2004. "Ending Social Promotion: The Response of Teachers and Students." Consortium on Chicago School Research. Chicago, IL.

[15] Kling, J., J. Leibman and L. Katz. 2005. "Experimental Analysis of Neighborhood Effects." NBER Working Paper # 11577.

[16] McLeskey, James and Nancy L. Waldron. 1991. "Identifying Students with Learning Disabilities: The Effect of Implementing Statewide Guidelines." Journal of Learning Disabilities 24(8): 501-506.

[17] Mercer, C.D., King-Sears, P. & Mercer A. 1990. "Learning Disabilities Definitions and Criteria used by State Education Departments." Learning Disability Quarterly 13: 141-152.

[18] Lewitt, E. M. and L. S. Baker. 1996. "Child Indicators: Children in Special Education."The Future of Children 6(1): 139-151.

[19] Litt, J., Taylor, H., Klein, N. and M. Hack. 2005. "Learning Disabilities in Children with Very Low Birthweight: Prevalence, Neuropsychological Correlates and Educational Interventions." Journal of Learning Disabilities 38(2): 130-141.

[20] Losen, D. and G. Orfield, eds. 2002. Racial Inequity in Special Education. Cambridge, MA: Harvard Education Publishing Group.

[21] Martin, E., Martin, R. and D. Terman. 1996. "The Legislative and Litigation History of Special Education." The Future of Children 6(1): 25-39.

[22] Mellard, D. F. 1985. "Report of the Learning Disability Field Study." Office of Education of the Exceptional Children of the Superintendent of Public Instruction, Frankfurt, KY.

[23] Perlmutter, B. and Parus, M. 1983. "Identifying Children with Learning Disabilities: A Comparison of Diagnostic Procedures across School Districts." Learning Disability Quarterly 6: 321-328.

[24] Reschly, Daniel J. 1996. "Identification and Assessment of Students with Disabilities." The Future of Children 6(1): 40-53.

[25] Shepard, L. A. and M. L. Smith. 1983. "An Evaluation of the Identification of Learning Disabled Children in Colorado." Learning Disability Quarterly 6: 115-127.

[26] Singer, J. D. and J. A. Butler. 1987. "The Education for All Handicapped Children Act: Schools as Agents of Social Reform." Harvard Educational Review 57(2): 125-152.

[27] Thurlow, M., Ysseldyke J. and A. Casey. 1984. "Teachers Perceptions of Criteria for Identifying Learning Disabled Students." Psychology in the Schools 21: 349-55.

[28] Wagner, M., Newman, L. Cameto, R., Garza, N. and Levine P. 2005. After High School: A First Look at the Postschool Experiences of Youth with Disabilities. A Report from the National Longitudinal Transition Study-2. Menlo Park, CA: SRI International.

[29] Wagner, M., Marder, C., Levine, P., Cameto, R,. Cadwallader, T. and Blackorby, J. (with Cardoso, D. and Newman, L.). 2003. The Individual and Household Characteristics of Youth with Disabilties. A Report from the National Longitudinal Transition Study-2. Menlo Park, CA.: SRI International.

Figure 1a: Total U.S. Students in Special Education (Millions)

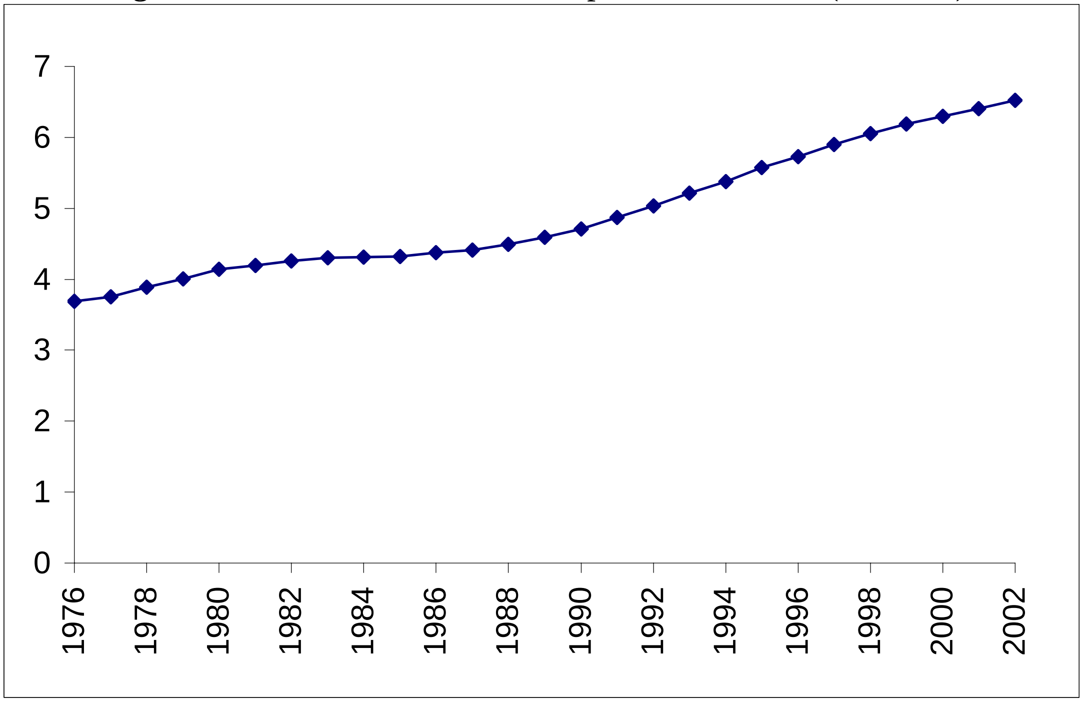

Figure 1b: Fraction of U.S. Enrolled Students in Special Education

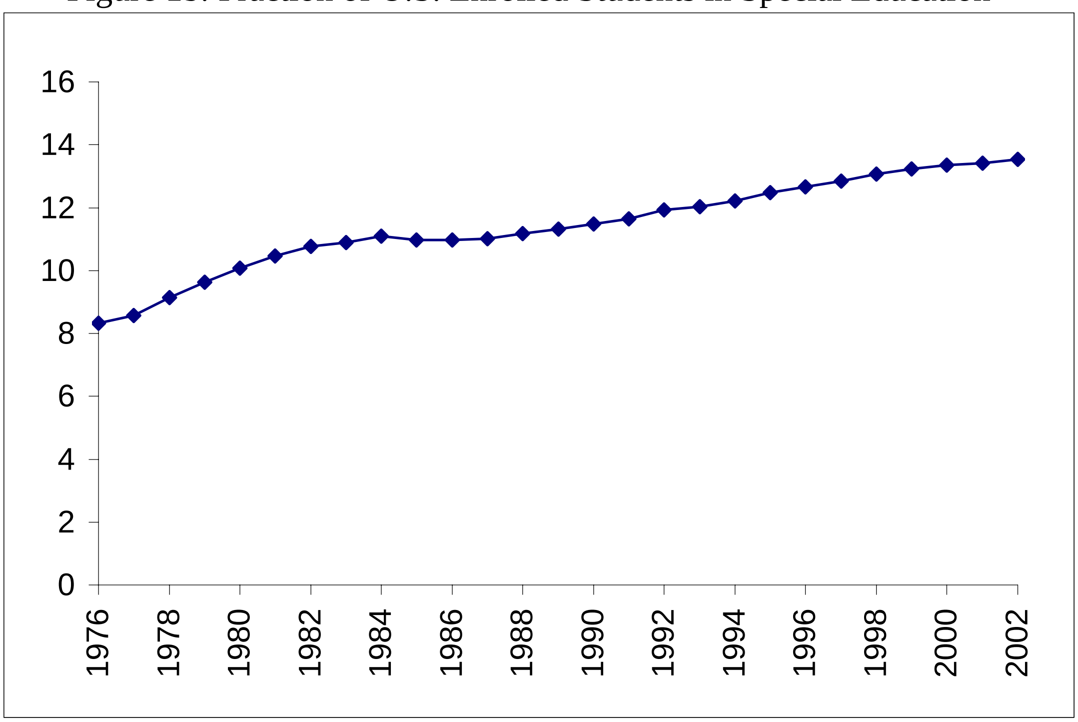

Figure 2a: Fraction of CPS Elementary School Students in Special Education

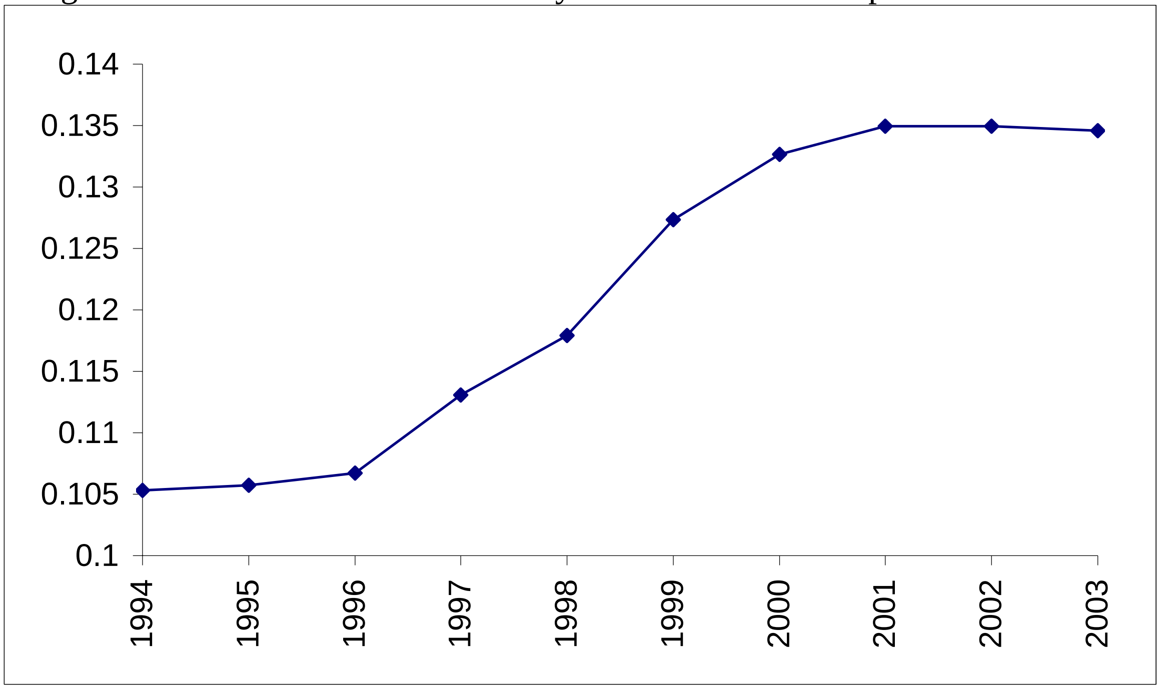

Figure 2b: Number of CPS Elementary School Students in Special Education

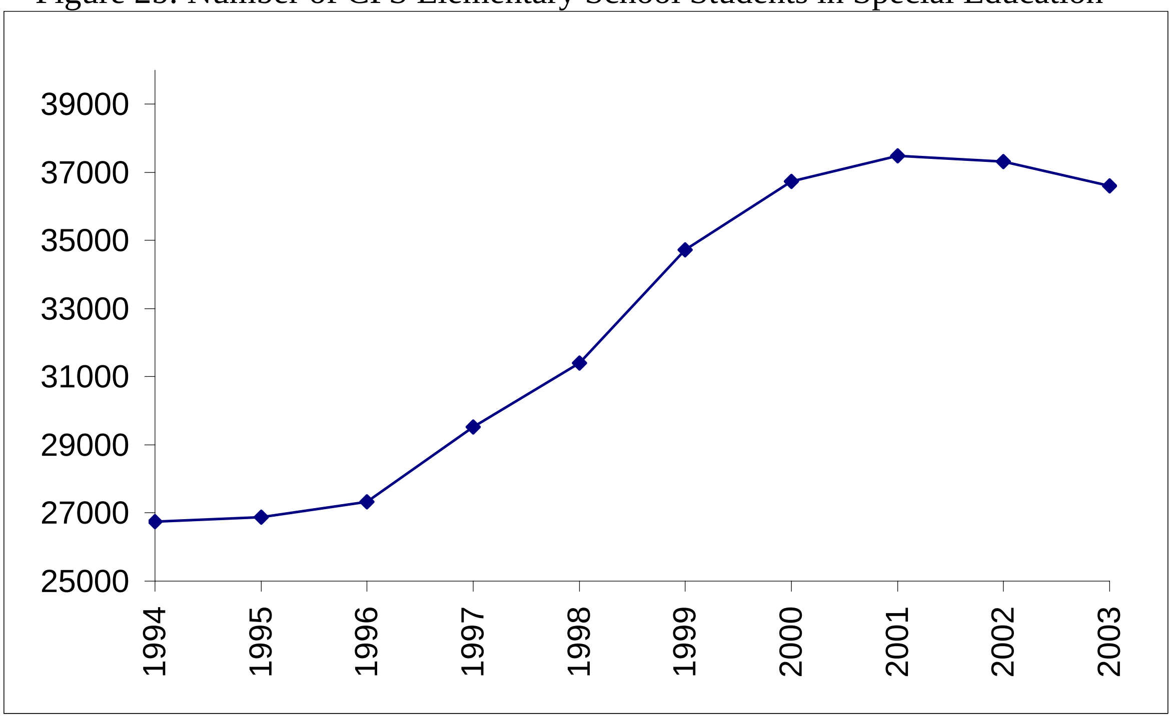

Figure 3: Growth in SE Enrollment for Schools Around, Above and Below Disciplinary Cutoff

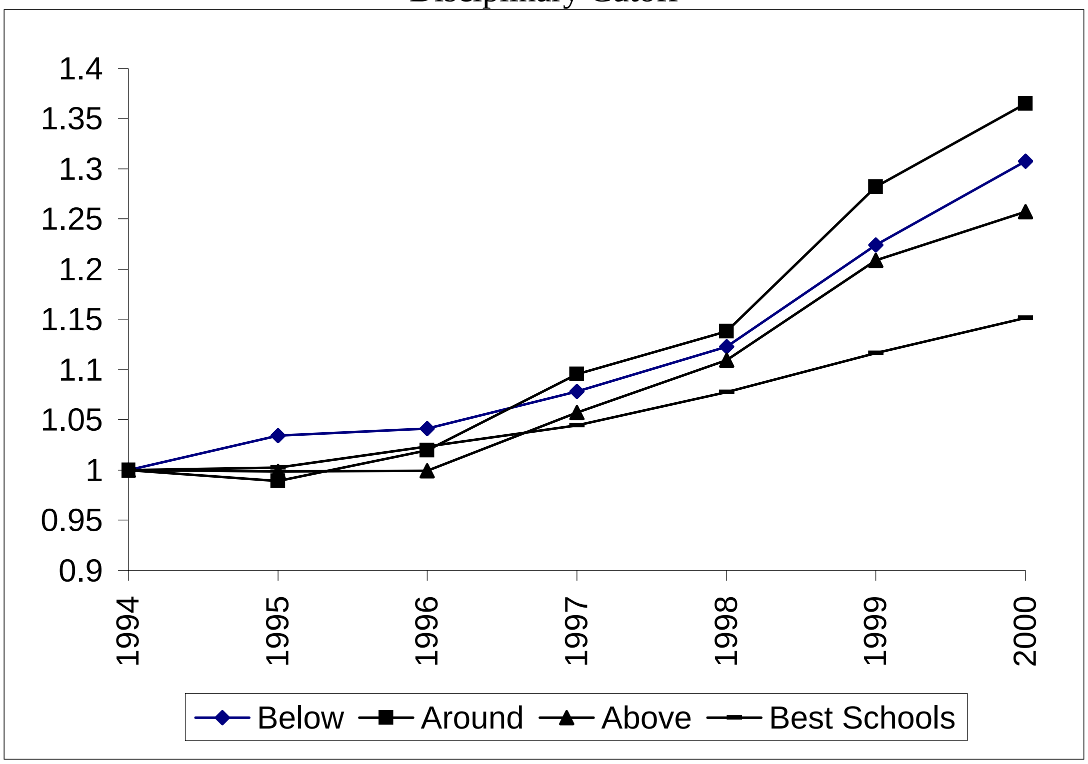

Figure 4: Change in Special Education Enrollment by 1996 Pnorms

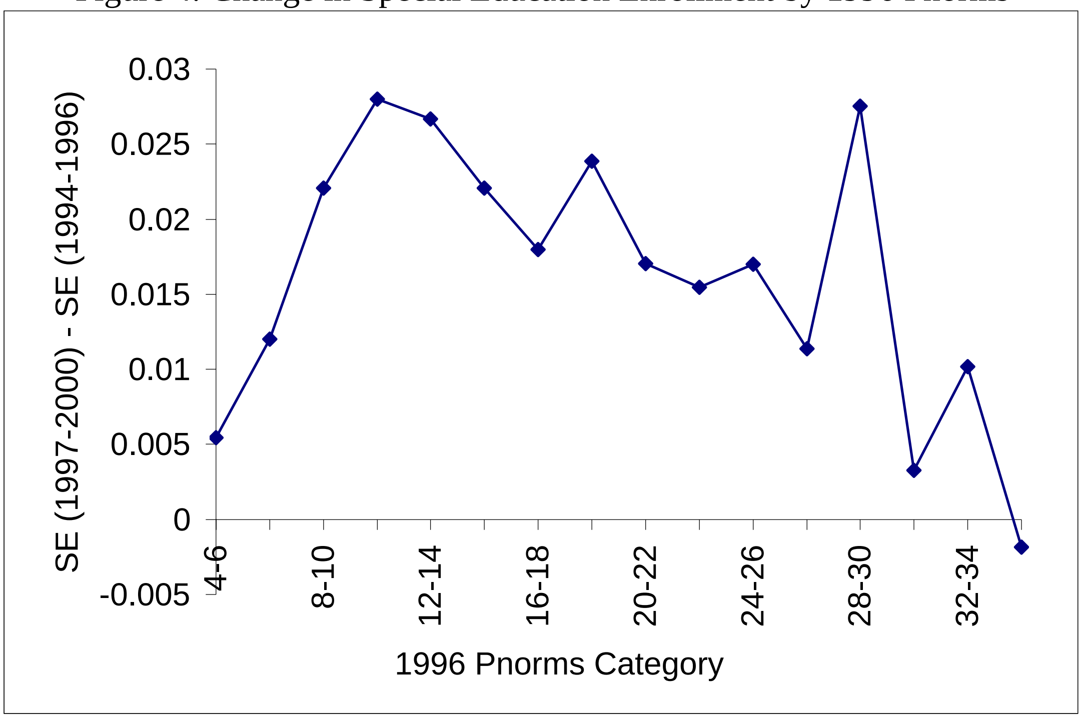

Figure 5: Illustration of Instrument I96

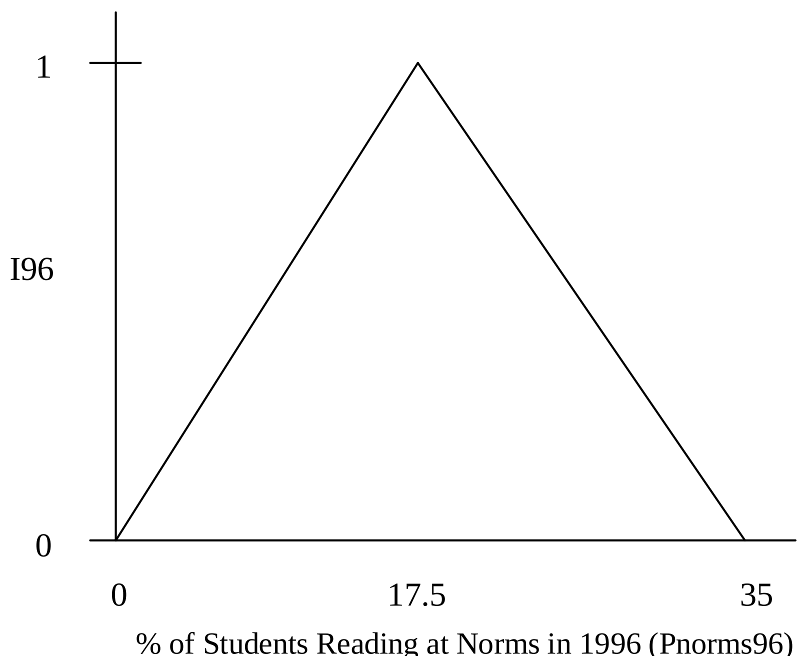

Figure 6: Impact of School Incentives on the Probability of Special Education Placement

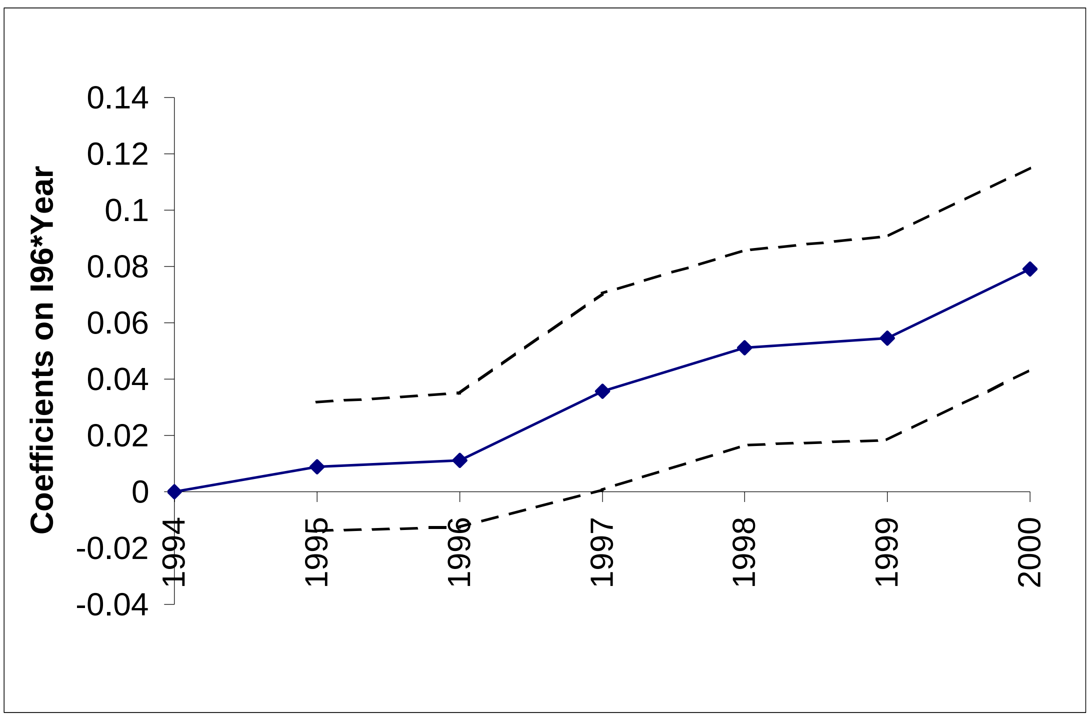

Figure 7a: Impact of School Incentives on the Probability of Special Education Placement by Prior Achievement

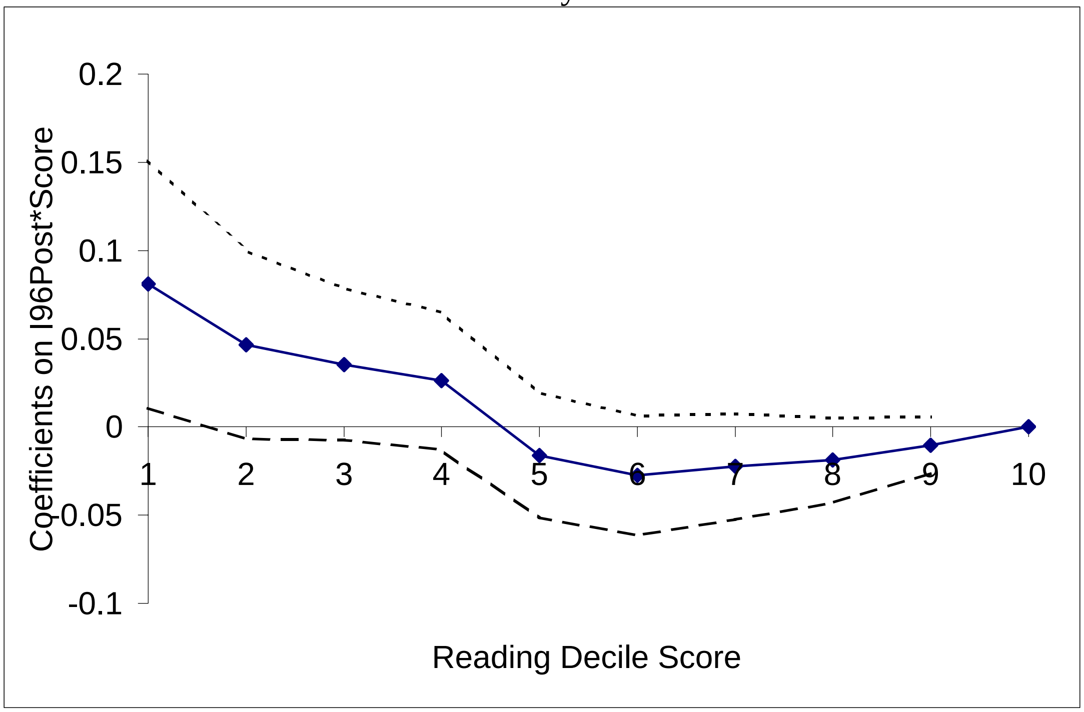

Figure 7b: Impact of School Incentives on the Probability of Special

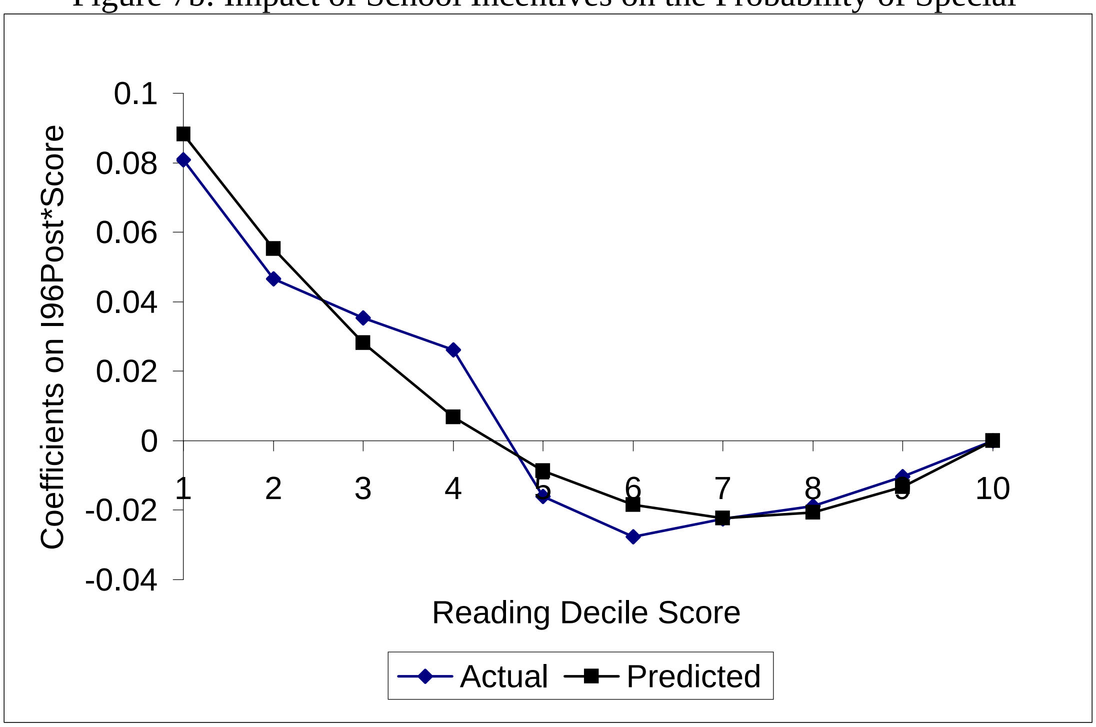

Figure 8a: Impact of School Incentives on Dropout Probability by Prior Achievement

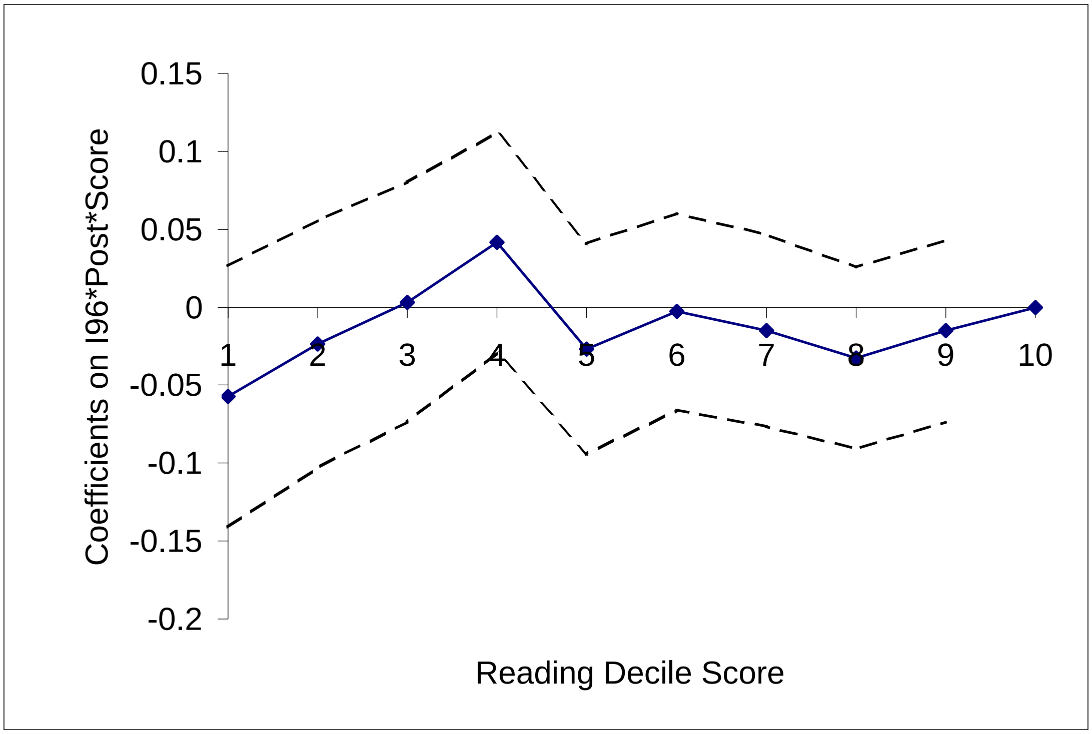

Figure 8b: Impact of School Incentives on School Absences by Prior

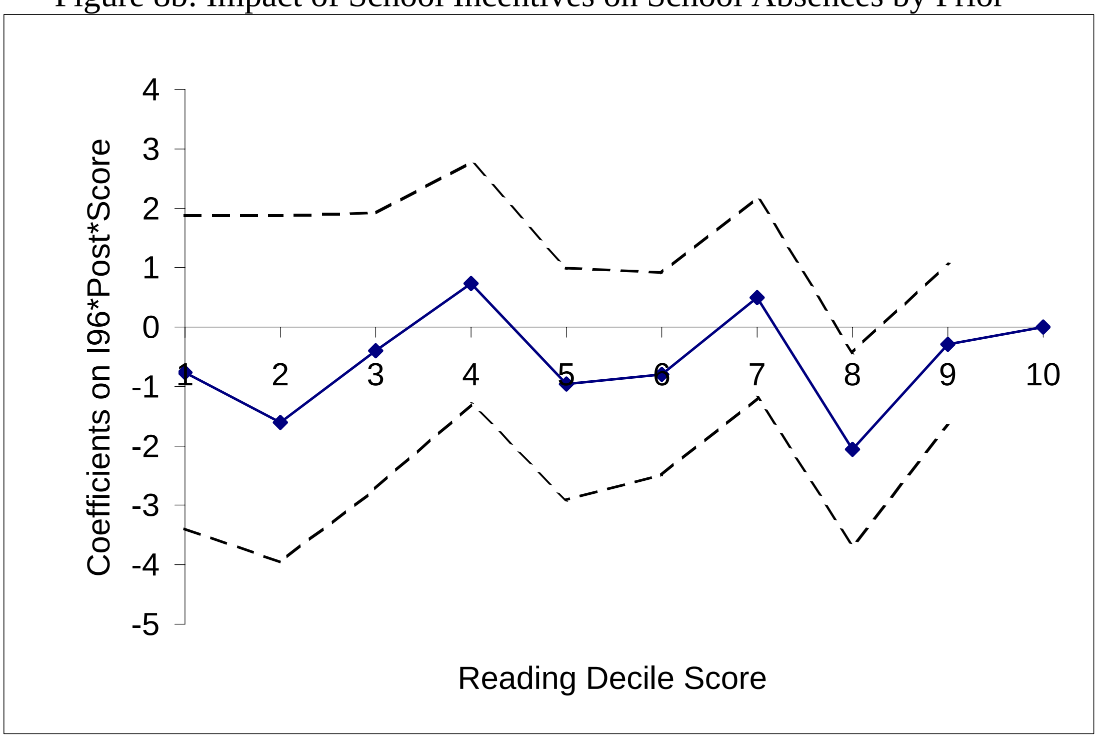

Figure 8c: Impact of School Incentives on GPA by Prior Achievement

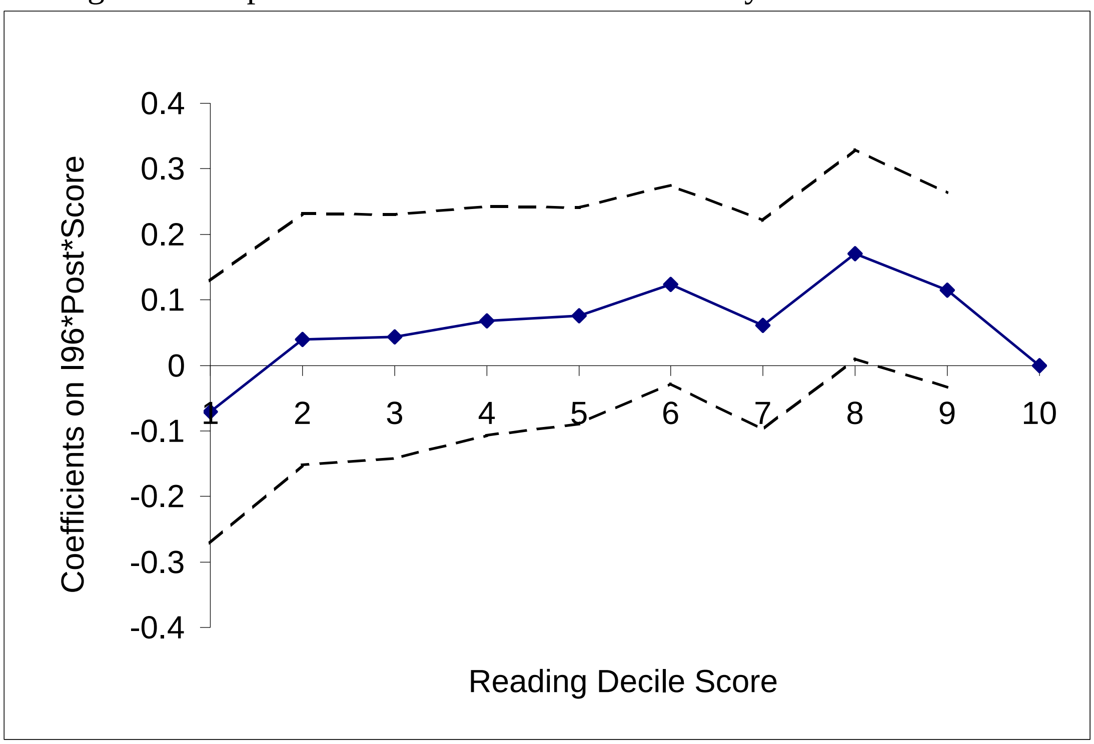

Table 1: Various Characteristics of Schools by Percent of Students Reading at National Norms in 1996 (*Pnorms96*)

	(1)	(2)	(3)	(4)	(5)
	1996 *Pnorm* Category:				
	0 - 10%	10-20%	20-30%	30-50%	50-100%
Total Enrollment	555	691	743	690	563
% Special Ed	.130	.120	.117	.132	.142
% with Learning Disability	.070	.069	.070	.083	.082
% with Behavioral Disorder	.012	.011	.010	.009	.008
% with Speech Disability	.019	.018	.018	.019	.027
Share of SE with LD	.546	.591	.606	.645	.578
Share of SE with EBD	.094	.089	.084	.069	.055
Share of SE with SPL	.164	.172	.168	.160	.221
% above Norms in Reading	0.08	0.16	0.25	0.38	0.67
Size of a 3rd Grade Class	21	23	24	26	28
% African-American	.921	.767	.622	.383	.281
% Hispanic	.078	.222	.309	.391	.226
% Limited-English-Proficient	.048	.107	.156	.180	.113
% Low-Income	.959	.940	.913	.802	.479
Obs	27	136	135	92	66

Notes: Sample is all regular and magnet CPS elementary schools
Each column is from an OLS regression of the school characteristic indicated by rov
on a variable indicating the 1996*Pnorm* category of the school.

Table 2: Means of Demographic Variables for Special Education (SE) and Regular Education (RE) 13-year-olds in CPS

	SE	RE	Difference (SE - RE)
	(1)	(2)	(3)
	A. Background Characteristics		
Male	.656	.466	.190
	(.004)	(.002)	(.004)
Black	.674	.611	.063
	(.005)	(.001)	(.004)
Hispanic	.263	.331	-.067
	(.004)	(.001)	(.003)
Free/Reduced Price Lunch	.910	.892	.018
	(.006)	(.001)	(.002)
Bilingual	.263	.347	-.084
	(.004)	(.001)	(.004)
Grade (Age 10)	3.798	3.992	-.193
	(.040)	(.006)	(.010)
Reading Score Decile (Age 10)	3.092	5.720	-2.628
	(.063)	(.009)	(.030)
Math Score Decile (Age 10)	2.801	5.777	-2.976
	(.063)	(.009)	(.030)
	B. School Characteristics		
% Reading at Norms (1996)	.213	.215	-.002
	(.002)	(.000)	(.001)
% Low Income	.914	.910	.004
	(.007)	(.000)	(.001)
% LEP	.152	.158	-.006
	(.002)	(.001)	(.002)
% Black	.619	.615	.004
	(.006)	(.001)	(.004)
% Hispanic	.317	.327	-.010
	(.004)	(.001)	(.003)
8th Grade Class Size	.255	.261	-.006
	(.002)	(.000)	(.000)
	C. Census Tract Characteristics		
% Above Poverty	.704	.708	-.004
	(.006)	(.001)	(.002)
% Males Employed	.639	.644	-.005
	(.005)	(.001)	(.002)
Ave. Years of Education	12.037	12.000	.037
	(.108)	(.003)	(.009)
Observations	21,383	105,718	

Notes: Standard errors are in parentheses. Sample is all CPS 13-year-olds in regular and magnet elementary schools.
Only students in schools with *Pnorms96* <36 are included (these are the 80% of elementary schools in which the fraction of students reading at national norms was below 36%).
Each column is from an OLS regression of the demographic characteristic indicated by row on a dummy indicating Special Education status at age 13.

Table 3: First Stage: Impact of Incentives from Accountability Policy (I96) on SE Enrollment
Dependent Variable: Special Education Status at Age 13

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
*I96*POST*	.048	.058	.024	.043	.051	.050	.043	.051			
	(.014)	(.018)	(.014)	(.016)	(.015)	(.015)	(.016)	(.015)			
I96 *1995									.009	.005	.023
									(.012)	(.012)	(.014)
I96 *1996						.007	.009	.002	.011	.011	.013
						(.010)	(.011)	(.011)	(.012)	(.013)	(.014)
I96 *1997									.036	.026	.044
									(.018)	(.018)	(.019)
I96 *1998									.051	.041	.609
									(.018)	(.018)	(.019)
I96 *1999									.055	.045	.045
									(.018)	(.019)	(.019)
I96 *2000									.079	.073	.073
									(.018)	(.020)	(.020)
Including Students in SE at Age 10		x									
Indiv. Controls			x								
School Controls				x							
*Zip Code*Year*					x						
Lower Side of Instrument							x			x	
Upper Side of Instrument								x			x

Notes: Standard errors in parentheses account for correlation between observations in the same elementary school and year.
Sample is all CPS 13-year-olds in regular and magnet elementary schools with age 10 test scores who were not in SE at age 10 (except in Column (2)).
Only students in schools with Pnorms96<36 are included (these are the 80% of elementary schools in which the fraction of students reading at national norms in 1996 was below 36%).
I96 is a variable reflecting the strength of incentives to increase SE placement at each elementary school and is defined in Equation (3) in the text.
I96 is equal to one for schools with the strongest incentives and zero for those with the weakest incentives. *Post* is a dummy equal to one after 1996.
Excluded variable in Columns (9) - (11) is *I96* *1994.
Each column is an OLS regression of a dummy variable indicating SE status at age 13 on the independent variables indicated by row.
All specifications include year and school fixed effects and control for *Pnorms96*POST* and *Probation *POST*, where *Pnorms96* is the fraction of students reading at norms in 1996 and *Probation* is a dummy variable equal to one for schools placed on probation in 1997.
Coefficients in Columns (7) and (8) are from the same regression where the lower (*Pnorms*<17.5) and upper (*Pnorms*>17.5) sides of *I96* are separate variables (coefficients in Columns (10) and (11) are estimated in the same way).
Background controls include gender, race, age in grade 10, bilingual status and free/reduced price lunch. School controls include ethnic composition of school, fraction low income, fraction LEP and 3rd and 8th grade class size.
To interpret magnitudes, note that the mean of the dependent variable is roughly .12 (.17 in Column (2)) and the mean of *I96* is about .60.

Table 4: Effect of Accountability Policy on Special Education Status for Various Populations

	Dependent Variable:									
	Special Education Status at Age 13								SE Status (Mental / Emotional Disability)	SE Status (Speech / Language Disability)
Sample Restricted to:	Black	Hispanic	Male	Female	Bilingual	Bottom Reading Quartile (Age 10)	Top Reading Quartile (Age 10)	In SE at Age 10	All	All
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
*I96*POST*	.058	-.010	.031	.058	-.051	.139	.002	-.074	.039	.003
	(.020)	(.035)	(.026)	(.018)	(.037)	(.040)	(.015)	(.047)	(.016)	(.004)
Mean of Dep. Variable:	.168	.139	.222	.115	.125	.407	.022	.906	.137	.010
Observations	78954	40579	63298	63802	41917	26261	26164	8885		

Notes: Standard errors in parentheses account for correlation between observations in the same elementary school and year.
Sample is CPS 13-year-olds in regular and magnet elementary schools. Only students in schools with *Pnorms96* <36 are included (these are the 80% of elementary schools in which the fraction of students reading at national norms in 1996 was below 36%).
I96 is a variable reflecting the strength of incentives to increase SE placement at each elementary school and is defined in Equation (3) in the text. *I96* is equal to one for schools with the strongest incentives and zero for those with the weakest incentives.
Post is a dummy variable equal to one after 1996. Columns (1) - (8) are from OLS regressions of a dummy variable indicating SE status at age 13 on *I96*POST* and is restricted to students with the characteristics indicated by column (each characteristic is a dummy variable).
Columns (9) - (10) are from OLS regression of a dummy indicating SE status for the disability indicated by column on *I96*POST*.
All specifications include year and school fixed effects and control for*Pnorms96*POST* and *Probation*POST*.
To interpret magnitudes note that the mean of *I96* is about .60.

Table 5: First Stage using Prior Achievement: Effect of Incentives from Accountability on Special Education by Reading Decile Score

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
*I96*POST*	.156	.147	.138	.150	.165	.155	.159
	(.044)	(.044)	(.048)	(.044)	(.050)	(.045)	(.051)
*I96*POST*Decile*	-.046	-.042	-.041	-.045	-.049	-.047	-.043
	(.015)	(.014)	(.016)	(.015)	(.017)	(.015)	(.017)
*I96*POST*DecileSquared*	.003	.003	.003	.003	.003	.003	.003
	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)
Effect at 5th Decile	.010	.011	.005	.006	.008	.007	.019
	(.011)	(.010)	(.012)	(.012)	(.013)	(.012)	(.013)
Effect at 1st Decile	.113	.109	.100	.108	.120	.112	.119
	(.032)	(.032)	(.035)	(.031)	(.036)	(.032)	(.037)
Not in SE at Age 10		x					
Indiv. & School Controls			x				
*Zip Code*Year*				x			
Data at Age 16					x		
Lower Side of I96						x	
Upper Side of I96							x

Notes: Standard errors in parentheses account for correlation between observations in the same elementary school and year. Sample is all CPS 13-year-olds in regular and magnet elementary schools with age 10 test scores (students in SE at age 10 are excluded in Column (2) and only students with outcomes at age 16 are included in Column (5)). Only students in schools with *Pnorms96* <36 are included (these are the 80% of elementary schools in which the fraction of students reading at national norms in 1996 was below 36%).
I96 is a variable reflecting the strength of incentives to increase SE placement at each elementary school and is defined in Equation (3) of the text. *I96* is equal to one for schools with the strongest incentives and zero for those with the weakest incentives. *POST* is a dummy variable equal to one after 1996.
Test scores are measured using the average of a child's reading scores at ages 9 and 10 on the Iowa Test of Basic Skills and are then categorized by decile. Each column is an OLS regression of a dummy variable indicating SE status at age 13 on *I96*POST* and its interaction with decile score and decile score squared. All specifications include all two-way interactions and year, school and decile fixed effects. Coefficients in Columns (6) and (7) are from the same regression where the lower (*Pnorms* <17.5) and upper (*Pnorms* >17.5) sides of *I96* are separate variables. Background controls include gender, race, age in grade 10, bilingual status and free/reduced price lunch. School controls include ethnic composition of school, fraction low income, fraction LEP and 3rd and 8th grade class size. The mean of the dependent variable and of *I96* are .16 and .60, respectively.

Table 6: First Stage using Prior Achievement: Coefficients on the Interactions between Decile Score Dummies and Incentive Measure (I96*POST)

Dependent Variable: Special Education Status at Age 13					
	(1)	(2)	(3)	(4)	(5)
*I96*POST*Decile1*	.096	.091	.095	.111	.081
	(.036)	(.039)	(.036)	(.040)	(.036)
*I96*POST*Decile2*	.038	.043	.037	.054	.047
	(.029)	(.032)	(.029)	(.032)	(.027)
*I96*POST*Decile3*	.028	.028	.027	.036	.035
	(.242)	(.026)	(.025)	(.026)	(.022)
*I96*POST*Decile4*	.026	.034	.026	.036	.026
	(.023)	(.025)	(.023)	(.024)	(.020)
*I96*POST*Decile5*	-.013	-.005	-.013	-.008	-.016
	(.021)	(.022)	(.021)	(.022)	(.018)
*I96*POST*Decile6*	-.022	-.021	-.022	-.015	-.028
	(.020)	(.021)	(.020)	(.021)	(.017)
*I96*POST*Decile7*	-.014	-.010	-.015	-.004	-.019
	(.017)	(.019)	(.018)	(.018)	(.012)
*I96*POST*Decile8*	-.013	-.015	-.013	-.015	-.010
	(.014)	(.015)	(.014)	(.015)	(.008)
*I96*POST*Decile9*	-.007	-.007	-.007	-.001	-.010
	(.009)	(.010)	(.009)	(.010)	(.008)
P-value	.052	.119	.062	.026	.015
Indiv. & School Controls		x			
Zip Code*Year			x		
Has Data at Age 16				x	
Not in SE at Age 10					x

Notes: Standard errors in parentheses account for correlation between observations in the same elementary school and year.
Sample and variable definitions are discussed in the notes to Table 5
Each column is an OLS regression of a dummy variable indicating SE statu at age 13 on the interaction between *I96*POST* and each decile score dummy
The excluded category is *I96*POST*Decile10*. All specifications include score main effects, all two-way interactions and year and school fixed effects
The reported P-value tests the hypothesis that the coefficients on the interactions between *I96*POST* and decile score dummies are jointly equal to zero

Table 7: Reduced Form: Effect of Incentives from Accountability Policy (I96) on Various High School Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	A. Dropout Status at Age 17			**B. GPA at Age 16**			**C. Absences at Age 16**		
*I96*POST*	-.010	-.011		.009	-.030		-.853	-.262	
	(.010)	(.012)		(.023)	(.028)		(.316)	(.375)	
I96 *1995			.007			-.052			.805
			(.041)			(.093)			(1.554)
I96 *1996			-.026			.125			-1.251
			(.043)			(.092)			(1.602)
I96 *1997			-.026			.035			-2.933
			(.047)			(.102)			(1.790)
I96 *1998			-.056			.024			-3.197
			(.047)			(.097)			(1.566)
I96 *1999			-.089			.114			-2.782
			(.045)			(.096)			(1.537)
I96 *2000			-.043			.052			-2.358
			(.041)			(.095)			(1.657)
Bottom Quartile		x	x		x	x		x	x

Notes: Standard errors in parentheses account for correlation between observations in the same elementary school and year.
Sample and variable definitions are discussed in the notes to Table 5.
Dropout status is a dummy variable equal to one if the student has dropped out by age 17. GPA and absences are measured in the fall semester beginning in the September in which the student is age 16.
Excluded variable in Columns (3), (6) and (9) is *I96*1994* . Each column is an OLS regression of the dependent variable indicated by column on the independent variables indicated by row. All specifications include year and school fixed effects. To interpret magnitudes, note that the mean of the dependent variable in Panel A is .36, in Panel B is 1.89, in Panel C is 11.39 and the mean of *I96* is about .60.

Table 8: Reduced Form using Prior Achievement: Effect of Incentives from Accountability on Various High School Outcomes by Decile Reading Score

	(1)	(2)	(3)	(4)
	A. Dep. Variable: Dropout Status at Age 17			
*I96*POST*	-.061	-.074	-.071	-.029
	(.042)	(.046)	(.041)	(.048)
*I96*POST*Decile*	.025	.026	.026	.016
	(.017)	(.019)	(.017)	(.019)
*I96*POST*DecileSquared*	-.002	-.002	-.002	-.002
	(.001)	(.002)	(.001)	(.002)
Effect at 5th Decile	.006	.002	.000	.007
	(.017)	(.019)	(.017)	(.017)
Effect at 1st Decile	-.038	-.051	-.047	-.015
	(.028)	(.031)	(.028)	(.033)
	B. GPA at Age 16			
*I96*POST*	.056	.004	.023	.059
	(.092)	(.099)	(.092)	(.111)
*I96*POST*Decile*	-.028	-.008	-.032	-.029
	(.037)	(.039)	(.037)	(.041)
*I96*POST*DecileSquared*	.003	.001	.004	.003
	(.003)	(.003)	(.003)	(.003)
Effect at 5th Decile	-.001	-.003	-.048	-.002
	(.035)	(.038)	(.037)	(.036)
Effect at 1st Decile	.031	-.002	-.006	.033
	(.063)	(.069)	(.063)	(.077)
	C. School Absences at Age 16			
*I96*POST*	-3.035	-2.357	-2.879	-3.157
	(1.424)	(1.543)	(1.407)	(1.614)
*I96*POST*Decile*	.933	.917	.989	1.000
	(.534)	(.575)	(.533)	(.586)
*I96*POST*DecileSquared*	-.078	-.082	-.084	-.083
	(.045)	(.048)	(.045)	(.048)
Effect at 5th Decile	-.326	.169	-.420	-.225
	(.530)	(.575)	(.538)	(.536)
Effect at 1st Decile	-2.181	-1.522	-1.974	-2.239
	(1.005)	(1.092)	(.988)	(1.140)
Indiv. & School Controls		x		
*Zip Code*Year*			x	
Not in SE at age 10				x

Notes: Standard errors in parentheses account for correlation between observations in the same elementary school and year. Sample and variable definitions are discussed in notes to Table 7. Test scores are measured using the average of a child's reading scores at ages 9 and 10 on the Iowa Test of Basic Skills and are categorized by decile. Dropout status is a dummy variable that is equal to one if the student has dropped out by age 17. GPA and absences are measured in the fall semester of the school year in which the student is 16.

Each column is an OLS regression of the dependent variable indicated in the first row of each panel on *I96*POST* and its interaction with decile score and decile score squared. All specifications include score main effects, all two-way interactions and year, school and decile fixed effects.

To interpret magnitudes, note that the mean of *I96* is .60, and the mean of the dependent variable is .36, 1.9 and 11.4 in Panels A, B and C, respectively.

Table 9: Reduced Form using Prior Achievement: Coefficients on the Interaction between Decile Score Dummies and Incentive Measure (I96*Post)

	Dependent Variable:					
	Dropout Status at Age 17		GPA at Age 16		Days Absent at Age 16	
	(1)	(2)	(4)	(5)	(7)	(8)
*I96*POST*Decile1*	-.009	-.038	-.059	-.015	-.558	-.041
	(.034)	(.038)	(.085)	(.091)	(1.196)	(1.274)
*I96*POST*Decile2*	-.008	-.030	.011	.047	-.577	.037
	(.032)	(.035)	(.078)	(.085)	(1.074)	(1.145)
*I96*POST*Decile3*	.030	.005	-.081	-.052	-.046	.508
	(.029)	(.032)	(.071)	(.076)	(.978)	(1.054)
*I96*POST*Decile4*	.029	.008	-.083	-.046	.802	.924
	(.026)	(.030)	(.065)	(.071)	(.902)	(.978)
*I96*POST*Decile5*	.043	.023	-.091	-.088	.273	.995
	(.024)	(.026)	(.057)	(.062)	(.793)	(.855)
*I96*POST*Decile6*	.017	-.005	-.078	-.055	.739	.945
	(.021)	(.024)	(.053)	(.056)	(.686)	(.738)
*I96*POST*Decile7*	-.008	-.019	-.012	-.016	.415	.756
	(.020)	(.022)	(.048)	(.052)	(.631)	(.687)
*I96*POST*Decile8*	.012	.009	-.004	-.004	-.058	.285
	(.017)	(.019)	(.044)	(.048)	(.545)	(.600)
*I96*POST*Decile9*	.006	.000	-.024	-.018	.788	.835
	(.017)	(.019)	(.042)	(.045)	(.237)	(.548)
Indiv. & School Controls		x	x	x	x	x
Zipcode*Year		x		x		x
P-Value	.030	.133	.028	.068	.176	.422

Notes: Standard errors in parentheses account for correlation between observations in the same elementary school and year.
Sample and variable definitions are discussed in the notes to Table 7.
Each column is an OLS regression of the dependent variable indicated by column on the interaction between *I96*POST* and each decile score dummy. The excluded category is *I96*POST*Decile10*. All specifications include score main effects, all two-way interactions and year and school fixed effects. The reported P-value tests the hypothesis that the coefficients on the interactions between *I96*POST* and decile score dummies are jointly equal to zero.

Table 10: OLS and IV Estimates of the Relationship Between Special Education Placement and High School Outcomes

(1)	(2)	(3)	(4)	(5)	(6)	(7)
OLS		IV				
A. Dropout at Age 17						
.097	-.003	-.313	-.362	-.370	-.430	-.436
(.004)	(.007)	(.323)	(.252)	(.312)	(.273)	(.344)
[.000]	[.635]	[.333]	[.150]	[.235]	[.116]	[.205]
B. GPA at Age 16						
-.499	-.154	-.098	.289	.020	.402	.002
(.011)	(.014)	(.596)	(.516)	(.604)	(.561)	(.627)
[.000]	[.000]	[.869]	[.575]	[.974]	[.474]	[.997]
C. Absences at Age 16						
2.794	.767	-9.442	-10.131	-12.324	-13.859	-16.419
(.144)	(.232)	(8.323)	(6.959)	(8.126)	(7.765)	(9.148)
[.000]	[.001]	[.257]	[.145]	[.129]	[.074]	[.073]

Notes: Columns (1) and (2) are OLS regressions of the dependent variable indicated by Panel on Special Education status at age 13. Column (1) is for the entire sample and Column (2) is for the bottom quartile only. Standard errors are in parentheses and P-values are in square brackets. Columns (3)-(7) are IV regressions (using the entire sample) of the dependent variable indicated by Panel on Special Education status, where SE status is intrumented with I96post interacted with test scores. Columns (3)-(5) use a quadratic function of test scores and Columns (6) and (7) use a cubic specification. All test scores are categorized by centile.Specifications in Columns (4) and (6) use both sides of *I96* separately. Specifications in Columns (5) and (7) include *I96*post* as an instrument in addition to its interaction with decile score, whereas all other specifications just use *I96*POST*score.*

Table 11: Robustness: Reduced Form Estimates of the Impact of Accountability on High School Outcomes Controlling for Endogenous Achievement

	(1)	(2)	(3)	(4)	(5)	(6)
	A. Dropout at Age 17		B. GPA at Age 16		C. Absences at Age 16	
*I96*POST*	-.071	-.072	.122	.121	-3.254	-3.269
	(.043)	(.043)	(.092)	(.092)	(1.451)	(1.454)
*I96*POST*Decile*	.025	.025	-.040	-.040	.910	.911
	(.017)	(.017)	(.036)	(.036)	(.539)	(.540)
*I96*POST*DecileSquared*	-.002	-.002	.004	.004	-.071	-.072
	(.001)	(.001)	(.003)	(.003)	(.045)	(.045)
Effect at 5th Decile	.000	.000	.013	.014	-.491	-.504
	(.017)	(.017)	(.034)	(.034)	(.528)	(.528)
Effect at 1st Decile	-.048	-.048	.086	.085	-2.415	-2.429
	(.029)	(.029)	(.063)	(.063)	(1.024)	(1.026)
Age 13 Reading Score	x		x		x	
Age 13 Score Fixed Effects		x		x		x

Notes: Standard errors in parentheses account for correlation between observations in the same elementary school and year. Sample and variable definitions are discussed in notes to Table 7. Specifications are identical to those estimated in Column (1) of Table 8 but control for various functions of decile reading score measured at age 13.