Available online at www.sciencedirect.com
ScienceDirect
Journal of Econometrics 142 (2008) 731–756
JOURNAL OF Econometrics
www.elsevier.com/locate/jeconom

Breaking the link between poverty and low student achievement: An evaluation of Title I[☆]

Wilbert van der Klaauw*

Microeconomic and Regional Studies Function, Federal Reserve Bank of New York, 33 Liberty Street, New York, NY 10045, USA

Available online 21 May 2007

Abstract

This study provides an evaluation of the impact of Title I funding of compensatory education programs on school finance and student performance in New York City public schools during the 1993, 1997 and 2001 school years. Estimates based on a regression-discontinuity approach indicate that the program was unsuccessful in improving student outcomes in high-poverty schools in New York City during this period, and may in fact have had adverse effects during the earlier years in our sample. Less evidence of a negative effect is found for the 2001 school year. These findings are related to the way in which the federal funds were spent.

JEL classification: I20; I22; I28

Keywords: Regression-discontinuity design; Federal aid; Compensatory education; School finance

1. Introduction

Title I of the Elementary and Secondary Education Act of 1965, represents the largest federal program for K-12 education.[1] It currently provides more than $12 billion in annual financial assistance to state and local education agencies for the expansion and improvement of instructional programs to meet the special needs of low-achieving students from schools with high concentrations of poverty. The goal of Title I financed supplementary educational services in reading and mathematics is to help overcome an inferior educational environment associated with growing up in a low-income family and attending schools with high concentrations of disadvantaged students.

Most studies of its effectiveness have indicated that Title I has not fulfilled its original expectation of closing the achievement gap between economically disadvantaged students and their more advantaged peers. Partly in

[☆]The views and opinions offered in this paper do not necessarily reflect those of the Federal Reserve Bank of New York or the Federal Reserve System as a whole.

*Tel.: +1 212 7205916; fax: +1 212 7201844.
E-mail address: Wilbert.Vanderklaauw@ny.frb.org

[1]Even though the program has alternatively been named Title I by Democratic administrations and Chapter I by previous Republican administrations, throughout the paper I simply refer to it as the Title I program.

0304-4076/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.jeconom.2007.05.007

response to these findings, the 1994 Improving America's Schools Act, which reauthorized the program, introduced several changes in the funding and operation of Title I compensatory education services. Further reforms were introduced in its most recent reauthorization as part of the No Child Left Behind Act of 2001, which was signed into law at the beginning of 2002.

This study provides a new evaluation of the impact of Title I funding on school finance and student performance at public elementary and secondary schools in New York City, the largest school district in the country. How, and to what extent, has Title I funding made a difference for New York City's schools and students? How have the recent changes in this federal program affected school spending decisions and school performance? How is the effectiveness of Title I school programs related to the way Title I funds are spent? To investigate these issues, this study uses unique school and grade level data on student performance, student backgrounds and school budgets of New York City public schools during the period 1993–2001.

The analysis differs from earlier evaluations of Title I in several important respects. First, it proposes and implements an evaluation method that better addresses the non-random allocation of funds and services. In comparing average achievement test scores of students and schools receiving Title I services or funds with those who do not, existing evaluations of Title I programs do not (or not sufficiently) account for the non-random selection of students and schools for participation in the program, so that corresponding impact estimates are likely to be biased. Even in the absence of any impact of Title I services we would not expect students in Title I schools (which are selected because of their high poverty rates) to perform as well as students in schools not receiving Title I funds. Similarly, without remedial education we would not expect participating students within a Title I school, who were selected for such services based on their past performance, to perform equally well as the other students in the school. This study adopts a regression-discontinuity approach which better accounts for this selection problem. More specifically, it exploits a discontinuity in the rule that determines Title I eligibility, where schools with poverty counts above the district average are eligible for Title I funds, while most other schools are not. While schools near the average will have comparable poverty counts, their Title I status will differ, which provides the basis of an RD evaluation of the impact of Title I on school performance.

Second, in contrast to most existing studies which aim to assess whether Title I has succeeded in narrowing the achievement gap, this study tries to answer the more relevant question of whether Title I enabled schools and students to perform better than they would have without Title I funding. Some evidence exists that in fact without the program, these students and schools would have fallen further behind. However, given the previously discussed concerns about the reliability of the methods used to obtain these results, the evidence to date on this true causal effect of Title I remains purely suggestive.

In evaluating the program's effectiveness, multiple outcome measures are considered, including grade retention, suspensions, school attendance, student mobility rates and a set of reading and mathematics test scores and gain scores. The impact estimates are then related to school budget data to assess how Title I funds were spent. Separate analyses for the 1992–1993, 1996–1997 and 2000–2001 school years provide a first indication of the impact of recent legislative changes in this program on New York City public schools.

The estimates indicate that Title I has been ineffective at raising student performance, and in fact appears to have had adverse effects during the 1993 and 1997 school years. Less evidence of adverse effects is found for 2001. The absence of an improvement in student performance is linked to the finding that Title I receipt actually does not translate into a statistically significant increase in per-pupil expenditures. This is partly due to the fact that Title I funds account for only a small proportion of per-pupil total expenditures, as well as to apparent crowding out effects. The estimation results are also consistent with the perceptions of many educators that common Title I funded remedial education programs have been ineffective.

The paper is organized as follows. The next section describes the Title I funding program. This is followed in Section 3 by a summary of earlier evaluations of the program's effectiveness and a description of the evaluation approach adopted in this study. Section 4 discusses the data set and presents the main empirical findings of the program's impact on student performance. Section 5 provides a sensitivity analysis and discusses the interpretation of the estimates. An analysis of school budget data and an assessment of the longer term impacts of Title I funding are presented in Section 6, while Section 7 concludes.

2. The Title I funding program

Since 1965, when Congress passed the Elementary and Secondary Education Act, Title I's main goal has been to "help close the educational achievement gap between economically disadvantaged students and their more advantaged peers by providing funding for supplementary educational services in reading and mathematics to low-achieving students in low-income elementary and secondary schools". The idea of compensatory education had evolved during the early sixties as part of the 'war on poverty', and was based on the argument that students from low-income families who live in areas with a high concentration of other poor families were twice disadvantaged. They not only have fewer non-school learning opportunities, but also attend schools that, for a variety of reasons, are not as able to aid them in achieving academic success as schools with an affluent student body.

With appropriations in fiscal year 2005 exceeding $12 billion, Title I is the largest federal program for K-12 education. It currently reaches more than twelve and a half million children each year, covering about 75% of all elementary schools and nearly 50% of all middle and secondary schools. Aside from some recent changes discussed later, the way in which Title I funds have been distributed to schools and students through the years has changed little and can be characterized by a four-stage allocation procedure.

In the first stage, the federal government allocates funds to counties on the basis of a formula that includes the number of school-aged children in the county from low income families (measured using Census data) and the average educational expenditure per pupil in each state. In the second stage state education agencies allocate the county funds to the individual school districts, using the same or a more recent measure of the number of children from poor families. Poverty may be measured using counts of children from families receiving welfare benefits and/or counts of children eligible for free lunch under the National School Lunch program.

Each school district then ranks its schools in terms of the poverty rate of each school's attendance area. Schools with a poverty rate above the average for the district, are eligible for Title I funds, while schools below the average are not. Funds are then allocated to eligible schools based on the number of low-achieving students (typically the number scoring below the district average on a standardized achievement test) in the school. While districts are required to rank attendance areas in this fashion, they have some flexibility in choosing which variables to use to define poverty. In addition, they have the choice of adopting instead a 25% rule where all attendance areas with a poverty rate of 25% or higher are designated as eligible, while those below 25% are ineligible. One important exception to determining school eligibility solely by its poverty ranking is that attendance areas that were selected under the general ranking requirements the year before may continue to be served for one additional year even if they are otherwise no longer eligible (the 'grandfathering' rule or 'hold-harmless' provision).

Finally, in the fourth stage, eligible schools select students for compensatory education services on the basis of educational need (determined by performance on standardized achievement tests or by teacher recommendation), regardless of family income. Schools receiving Title I funding must provide supplementary instructional services in the basic skills areas and grades where students are the most educationally disadvantaged, and then only to low achieving students in those grades.

The type of educational services provided is left at the discretion of districts and individual schools. The most frequently used delivery systems include pull-out programs (where participating students are taken out of the regular classroom for part of the day to receive instruction in reading and mathematics by separate teachers, materials and equipment), in-class programs (where students receive assistance within the regular classroom setting either from a Title I teacher or aide working with the regular classroom teacher), pre-school programs and extended day and summer school programs.

While there has been a recent shift towards alternative models, throughout most of its history the pullout program has been the most common method of delivery. Two different fiscal accountability requirements introduced soon after the program's introduction contributed to its popularity. The first required local districts to prove that they were contributing as many local resources to the Title I schools as to any other school in the district (the 'comparability' provision). A second regulation required schools and districts to prove that federal dollars were only used to supplement and not supplant other services the educational disadvantaged otherwise would have received paid for by non-federal sources (the 'supplement, not supplant' provision). Schools found that the simplest way for them to demonstrate compliance with these rules and regulations was to adopt the pull-out program in which reading specialists and teacher aides who taught the separate compensatory education classes were paid for entirely with federal funds, and their presence in the school was clearly an additional school activity.

While the basic funding structure of the Title 1 program has remained essentially unchanged during the past 40 years, there were several changes during this period in the fiscal and educational accountability imposed on the use of these funds. One of these changes, introduced as part of the 1988 reauthorization of the program, was to encourage schools with poverty rates over 75% to adopt so-called schoolwide projects. In a schoolwide project (unlike the traditional 'targeted assistance projects') eligible schools are no longer required to target services only to children deemed eligible for Title I, but can use these funds to upgrade the entire educational program of the school, for example by reducing class sizes. The same reform also introduced additional accountability standards, requiring that schools in which Title I participants do not make positive gains in average achievement (as measured by average gains on nationally normed standardized achievement tests in reading and mathematics) are required to develop a comprehensive plan for program improvement. If student performance after implementation of this plan continues to fall short of the standard, the state's education authority and the district will then jointly make additional revisions to the plan.

Disappointing results from several national assessments of the program led the Clinton Administration to propose a drastically restructured Title I program which was reauthorized by Congress as part of the 1994 Improving America's Schools Act. This legislation introduced new requirements and systems for assessment and evaluation of Title I. The new law required states to develop their own standards for the acquisition of real world competencies in reading, mathematics, science and other subjects, develop new curriculum frameworks that will permit all students to meet these standards, to develop their own methods of assessment to measure what students know and can do, and to monitor the progress of schools in improving the achievement of low-income students in meeting state standards. This framework replaced the use of standardized, norm-referenced testing as a standard for program success.

The new program also made several changes to the way funds are allocated. First, it removed a perverse incentive in the old system which penalized successful schools by making the allocated Title I funds for an eligible school dependent on the number of low-income children, rather than the number of low-achieving students in the school. Second, it allocated additional funds to the highest poverty areas and schools, so that beginning with the 1998–1999 school year, schools receive higher Title I grants per child the higher its poverty rate. It also expanded the number of schools eligible to operate schoolwide projects, by lowering the eligibility threshold from 75% to 60% in the 1995–1996 school year and to 50% in the 1996–1997 school year.

The most recent reauthorization of the program as part of the No Child Left Behind Act of 2001, emphasizes increased accountability for states, school districts, and schools. It requires states to implement statewide accountability systems covering all public schools and students, based on challenging state standards in reading and mathematics, annual testing of all students in grades 3–8, and the development of annual statewide progress objectives for all students. School districts and schools that fail to make adequate yearly progress toward statewide proficiency goals will, over time, be subject to improvement, corrective action, and restructuring measures. Schools that meet or exceed these objectives or close achievement gaps, on the other hand, will be eligible for State Academic Achievement Awards.

The revised program also includes greater choice for parents and students, particularly those attending low-performing schools. School districts must give students attending schools identified for improvement, corrective action or restructuring, the opportunity to transfer to a better public school, which may include a public charter school, within the school district.[2] Schools that want to avoid losing students, along with the portion of their annual budgets typically associated with those students, will have to improve or, if they fail to make adequate yearly progress for five years, run the risk of reconstitution under a restructuring plan.

[2]For students attending persistently failing schools, districts must permit low-income students to use Title I funds to obtain supplemental educational services from an approved public or private sector provider selected by the students and their parents.

In exchange for the strong accountability for results, the revised program gives states and school districts additional flexibility in the use of federal education dollars.[3]

3. Evaluating the effectiveness of Title I

Before discussing previous evaluations of the Title I program, it is useful to consider several factors which would lead us not to expect to find a large positive effect on student achievement, followed by a consideration of several compensating factors. First, despite being the largest federal program for elementary and secondary education and despite its large coverage, the program actually constitutes a relatively small share of total federal spending and of total school expenditures on K-12 education. For example, in fiscal year 2003 on-budget federal support for elementary and secondary education amounted to approximately $60 billion, of which $11.7 billion was accounted for by Title I, which was only slightly larger than the $10.8 billion spent through the second largest federal program for K-12 education, the Child Nutrition Program which includes the National School Lunch and School Breakfast Programs (US Department of Education, 2006; US Department of Agriculture, 2006). Moreover, total on-budget federal spending on K-12 education accounted for only 9.3% of total elementary and secondary institution expenditures, of which 5.7% were expenditures by the Department of Education (Sonnenberg, 2004).[4]

Second, despite existing fiscal accountability regulations aimed at preventing this, Title I's impact may be diluted by crowding out effects. In fact, evidence of states and cities substituting away some of their own funding to Title I schools has been found at the national level by Gordon (2004). Third, evidence of the overall effectiveness of public school spending is mixed, with some research reporting a modest positive relationship between student performance and per-pupil school resources (Krueger, 1998) and others finding no such relationship (Hanushek, 1998).

On the other hand, there are several aspects of the Title I program which mitigate these negative expectations for finding a large effect in this study. First, as Title I funding is not allocated equally across schools, for many schools it could represent a non-trivial proportion of the school budget. Second, fiscal and educational accountability restrictions associated with Title I receipt cause these funds to be specifically targeted for remedial education. These restrictions guarantee that most of the funds make it directly to the classroom, and they may trigger corrective action if schools do not show significant improvement among Title I students. Therefore, even if such funds do not significantly raise total per-pupil expenditures in a school, the program may still have a significant effect on school performance. Third, the evaluation in this paper is based on more recent data covering the 1993–2001 period, which as discussed earlier represents a period of significant reform. While the data predate the 2001 No Child Left Behind Act, they should incorporate the short and medium term effects of the 1994 reforms, and specifically the increased targeting of funds and the delinking of increased fund allocations and poor student performance.

Earlier evaluations of the Title I program have been mainly confined to several large federally mandated assessment studies, including the Sustaining Effects Study (Carter, 1984), the National Assessment of Chapter I (Kennedy et al., 1986), the Chapter I Implementation Study (Millsap et al., 1993) and the Prospects Study (Abt Associates, 1997). Many of the earlier evaluation studies were surveyed in Mullin and Summers (1983). The more recent national assessments, as well as several local area evaluations, were included in a meta-analysis by Borman and D'Agostino (1996). Like the Mullin and Summers survey, this study concluded that while Title I appears to have had a small positive effect on standardized test scores, it has not fulfilled its original expectation of closing the achievement gap between at-risk students and their more advantaged peers.

These evaluations were predominantly based on a norm-referenced model using standardized achievement tests in reading and mathematics, in which the standing of Title I participants relative to the national norm group on a pre-test is compared to these students' standing relative to the norm group on a post-test. The

[3]More specifically, it allows them to transfer up to 50% of the federal funding they receive under each of four major funded State grant programs to any one of these programs, or to Title I. The covered programs include Teacher Quality State Grants, Educational Technology, Innovative Programs, and Safe and Drug-Free Schools grants.

[4]The Child Nutrition Program is administered by the United States Department of Agriculture. In addition to the Title I program, the US Department of Education funds special education and school improvement programs as well as the Head Start program.

national norming sample in this case serves as the 'no-treatment' control group. Several studies have discussed the unreliability of the norm-referenced model as evaluation method (Linn, 1979, 1981; Tallmadge, 1982; Powers et al., 1983; Jaeger, 1979; Anderson, 1991; Davis, 1991). The model assumes that without program services, a student would stand at the same level on the post-test as on the pre-test, which would be indicated by a zero change score. This is commonly referred to as the 'equipercentile assumption'. In order for the assumption to be sound, norming samples would need to be nearly identical to Title I groups. Since, at a given pre-test score, students selected for the program are more likely to come from schools with high concentrations of poor children (and usually more limited resources) they are often not representative of the norm group. Selected students are also more likely to be slower learners who therefore are more likely to fall further behind each year. Moreover, the norm group itself includes significant numbers of students receiving Title I services. Consequently, the groups used to norm tests often do not serve as adequate 'no-treatment' control groups necessary to assess treatment related growth.

In addition to these selection bias and control group contamination issues, the approach suffers from a 'regression to the mean' effect, where due to randomness in test outcomes students with extreme scores on a pre-test tend to drift toward the average score on the post-test independent of any intervention (initial status will be correlated with gain scores). This will cause program gains to be overestimated, while the two previous problems are more likely to lead to an underestimation of the program effect.[5]

Similar problems exist with other evaluation methods that have been used in the literature involving comparisons of students who received Title I services with control groups of students who did not. When the no-treatment group consists of students from the same school or grades as the treatment group, students selected for Title I services will, by design, have lower achievement levels and motivation, and are more likely to be from poorer families. In addition, the non-Title I students may benefit from the pull-out of lower achieving students from the classroom (i.e. there may be positive spill-over effects).

A comparison of Title I students with comparable students in a non-Title I school is also problematic because the latter will generally be in a lower poverty school, with more resources than the former. Given these problems with the various evaluation approaches it is not surprising that impact estimates in the literature have been found to be very sensitive to the particular evaluation approach adopted (Borman and D'Agostino, 1996; Myers, 1986; Puma et al., 1993).

Another concern with many earlier evaluations of Title I relates to the main hypothesis they have tried to test, which is whether Title I has been successful in reducing the achievement gap between economically disadvantaged students and their more advantaged counterparts. One may argue that a more relevant assessment of its effectiveness would be to test instead whether Title I students learned more than they would have learned without Title I. In other words, it would be useful to know whether Title I helps prevent students from falling further behind, even if it were ineffective in actually reducing the achievement gap.

In this study I adopt a different approach to evaluating the program's effectiveness. Using school level data and a regression-discontinuity (RD) approach, the idea is to compare otherwise similar Title I and non-Title I schools to estimate the average impact on student achievement of attending a Title I school. In doing so, several outcome measures will be considered besides standard measures of achievement based on standardized tests, such as attendance and suspension rates. Moreover, estimates of the program's impact will be related to school expenditure patterns. Unfortunately, the school level data used in this study do not allow us to distinguish between students who actually did and did not receive Title I services. However, the focus on average student performance can be motivated by the fact that in most of the Title I schools in our sample the former group represents the majority of students, and our estimates will also capture potential spill-over effects on non-recipients of Title I services.

To describe the evaluation approach adopted in this paper let y_i be a measure of educational performance for school i, and let d_i be an indicator equal to 1 if the school receives Title I funds, and 0 if not. Further, let $y_i(1)$ be the outcome if school i were a Title I school, and $y_i(0)$ the outcome if it were not a Title I school. Then the actual outcome we observe equals $y_i = d_i y_i(1) + (1 - d_i) y_i(0)$. The model for the observed outcome can

[5]The regression to the mean effect also implies that using norm-referenced gain scores for identifying schools for program improvement will lead to overidentification of small projects and projects serving students with higher pre-test scores.

then be written as

$$y_i = \beta + \alpha_i d_i + u_i, \tag{1}$$

where $\alpha_i = y_i(1) - y_i(0)$ and $y_i(0) = \mathrm{E}[y_i(0)] + u_i = \beta + u_i$.

Comparing average outcomes of Title I schools and non-Title I schools would generally not provide us with an unbiased estimate of the average treatment effect $\mathrm{E}[\alpha_i]$ as

$$\begin{aligned}
&\mathrm{E}[y_i(1)|d_i = 1] - \mathrm{E}[y_i(0)|d_i = 0] \\
&\quad = \mathrm{E}[\alpha_i] + (\mathrm{E}[y_i(0)|d_i = 1] - \mathrm{E}[y_i(0)|d_i = 0]) \\
&\qquad + \Pr(d_i = 0)\ (\mathrm{E}[y_i(1) - y_i(0)|d_i = 1] - \mathrm{E}[y_i(1) - y_i(0)|d_i = 0]) \\
&\quad = \mathrm{E}[\alpha_i] + (\mathrm{E}[u_i|d_i = 1] - \mathrm{E}[u_i|d_i = 0]) \\
&\qquad + \Pr(d_i = 0)\ (\mathrm{E}[\alpha_i|d_i = 1] - \mathrm{E}[\alpha_i|d_i = 0]).
\end{aligned}$$

If average outcomes for Title I schools and non-Title I schools differed even in absence of Title I, or if average outcome gains resulting from Title I funding were different for both groups of schools, one or both of the last two terms in the equation will not be zero. As described in Section 2, within a given school district Title I status is determined on the basis of the percentage of students in the school attendance area from low-income families. Because poverty itself is strongly correlated with student outcomes (which of course is the whole reason for Title I to exist) the two bias terms are unlikely to be zero.

To solve this selection bias problem, I use the fact that eligibility for receipt of Title I funding within a given school district is determined by the poverty rate of each school's attendance area in such a way that schools with a poverty rate below a given cutoff are ineligible and receive no Title I funds at all, while all schools with a poverty rate above the cutoff are Title I recipients. Ignoring for the moment the one exception to this rule (the hold-harmless provision), Title I status is assigned by the deterministic rule:

$$d_i(s_i) = 1\{s_i \geqslant \mathscr{S}\},$$

where s_i is school i's poverty rate, $\mathscr{S}$ is the cutoff poverty rate and $1\{\}$ is the indicator function. This corresponds to the selection rule of a sharp RD design (Thistlethwaite and Campbell, 1960). Now by focusing on schools with poverty rates near the cutoff level, it is possible to evaluate the causal impact of Title I by comparing the performance of schools with poverty rates just above the cutoff to that of schools with poverty rates just below the cutoff.

More formally, as shown by Hahn et al. (2001) and van der Klaauw (2002), if in the case of a sharp design both $\mathrm{E}[u_i|s]$ and $\mathrm{E}[\alpha_i|s]$ are continuous in s at $\mathscr{S}$ then the difference between average outcomes in schools with poverty rates just above and below the threshold identifies

$$\lim_{s\downarrow\mathscr{S}} \mathrm{E}[y|s] - \lim_{s\uparrow\mathscr{S}} [y|s] = \mathrm{E}[\alpha_i|s = \mathscr{S}], \tag{2}$$

which is the average effect of Title I funding on schools with a poverty rate close to $\mathscr{S}$.[6]

If we include cases where otherwise ineligible schools may be receiving Title I funding because they were eligible to receive it in the previous year, the probability of being a Title I school as a function of the school attendance area's poverty rate $\mathrm{E}[d_i|s] = \Pr[d_i = 1|s]$ will no longer be a 0–1 step function but will still be a function that is discontinuous in s at $\mathscr{S}$. This would therefore represent the case of a so-called 'fuzzy RD' design. As shown by Hahn et al., under the same two continuity assumptions listed above and an additional local monotonicity assumption similar to that in Imbens and Angrist (1994),

$$\frac{\lim_{s\downarrow\mathscr{S}} \mathrm{E}[y|s] - \lim_{s\uparrow\mathscr{S}} \mathrm{E}[y|s]}{\lim_{s\downarrow\mathscr{S}} \mathrm{E}[d|s] - \lim_{s\uparrow\mathscr{S}} \mathrm{E}[d|s]} = \lim_{e\downarrow 0} \mathrm{E}[\alpha_i|d_i(\mathscr{S} + e) - d_i(\mathscr{S} - e) = 1], \tag{3}$$

which represents the local average treatment effect (LATE) for the subgroup of schools for which Title I status changes discontinuously at the cutoff point. These are the schools that would not receive Title I funding if their poverty rates fell just below $\mathscr{S}$ (i.e. their treatment assignment would be zero as in $d_i(\mathscr{S} - e) = 0$), but would be Title I schools if their poverty rates exceeded $\mathscr{S}$ (treatment assignment $d_i(\mathscr{S} + e) = 1$). In our case

[6]Note that in the sharp design we actually only require $\mathrm{E}[\alpha_i|s]$ to be right-continuous at $\mathscr{S}$.

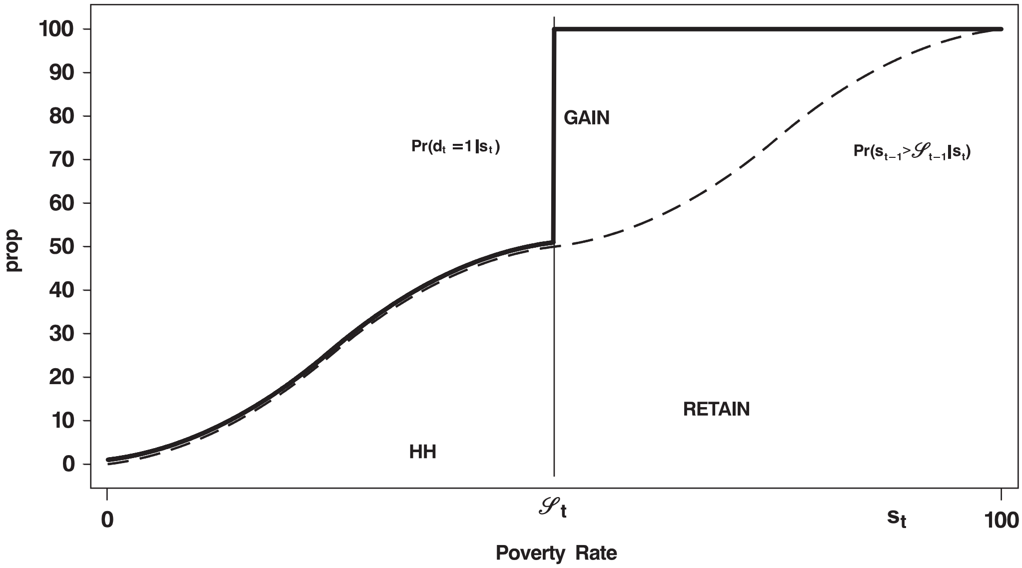

Fig. 1. Characterization of treatment effect: case with variation in school poverty rate and fixed cutoff.

this represents the subgroup of schools not covered by the hold-harmless provision, i.e. those whose poverty level fell below the cutoff in the previous year.[7]

As will be explained in more detail below, this preliminary characterization of the (local) average treatment effect identified in our case applies to a simplified scenario in which schools close to the poverty cutoff have comparable histories of Title I receipt. While under certain conditions one could expect these to be similar, under others they may differ in which case a comparison of schools on either side of the cutoff captures not just a difference in current Title I status, but also in previous Title I receipt. For this to occur, changes in poverty rates would have to be strongly correlated with past Title I receipt. More generally, in case of dynamic treatment assignment, the interpretation of the treatment effect defined in (3) will depend on the nature of the year-to-year variation in a school's poverty rate relative to the cutoff point that could lead to a change in a school's Title I status. For example, if each school's position relative to the poverty cutoff rate has remained fixed throughout the past few years (so that we are in effect evaluating the effect of permanent Title I assignment several years ago), then the estimated effect represents the cumulative effects of multiple years of Title I funding compared to the receipt of no funding. If, on the other hand, the year-to-year variation in poverty rates and Title I eligibility is such that schools just left and just right of the cutoff have comparable pre-assignment histories of Title I receipt then, as argued below, (3) will measure the short-run effect of gaining (or retaining) Title I funding.

Before considering the former scenario in greater detail, it is useful to examine the latter with the aid of two special cases. In the first, the year-to-year variation in treatment status is solely due to changes in a school's poverty rate. In the second case, changes in status are instead caused by a change in the poverty cutoff rate. Fig. 1 shows how, with a fixed poverty cutoff, random changes in a school's poverty rate determine the subgroup of interest. The dashed curve in the figure represents the fraction of schools for which the previous year's poverty rate exceeded the eligibility cutoff in that year, as a function of the current period's poverty rate, $\Pr(s_{t-1} \geqslant \mathscr{S}_{t-1}|s_t)$. The solid curve represents the percentage of schools receiving Title I funds in current year t as a function of s_t, where $\mathscr{S}_t$ is the poverty cutoff rate in year t. Both curves overlap at poverty values below this cutoff rate. Note that the area under the $\Pr(s_{t-1} \geqslant \mathscr{S}_{t-1}|s_t)$ curve left of $\mathscr{S}_t$, marked HH, represents the fraction of schools who would be receiving funds under a hold-harmless provision. Similarly, the fraction of schools retaining eligibility on the basis of their high current poverty rate is represented by the area labeled

[7]Of course, if we observed the school's poverty rate in the previous year, which is not the case in our data, we could simply restrict our evaluation to this set of schools in which case the evaluation problem would correspond to that of a sharp RD design.

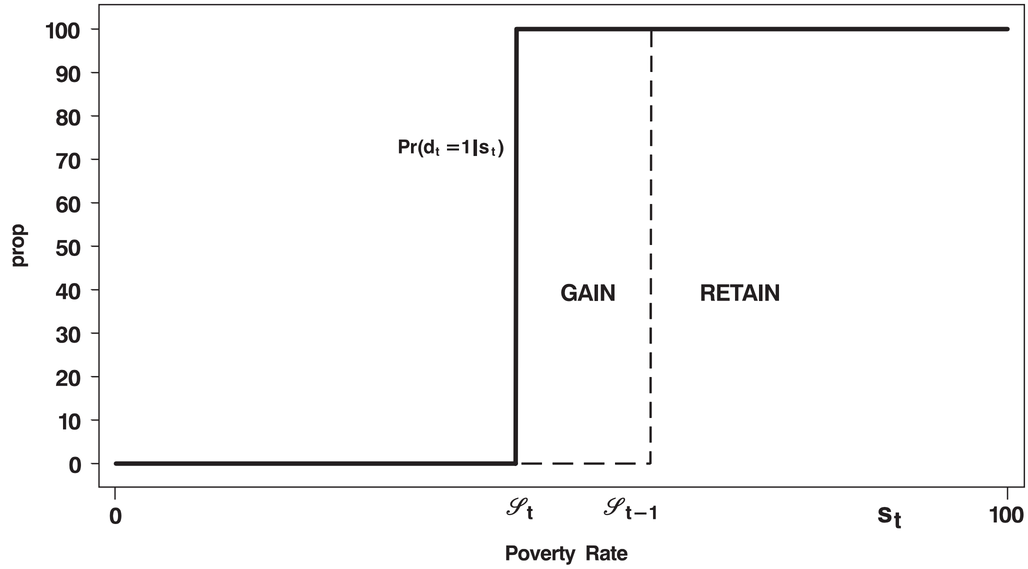

Fig. 2. Characterization of treatment effect: case with change in cutoff and fixed school poverty rate.

RETAIN. With $\Pr(s_{t-1} \geqslant \mathscr{S}_{t-1}|s_t)$ continuous at $\mathscr{S}_t$, the ratio in (3) defined at the point of discontinuity in $\Pr(d_t = 1|s_t)$, identifies the effect of Title I status for the newly eligible schools (represented in the figure by the area marked GAIN), with poverty rates close to $\mathscr{S}_t$. In the case without a hold-harmless provision, which applies to one of the years in our data, the probability of Title I receipt $\Pr(d_t = 1|s_t)$ is a simple 0–1 step function, in which case the RD approach would identify the (weighted) average effects of retaining and newly gaining Title I status for that year.

Fig. 2 considers the case where each school's poverty rate remains the same, but where the poverty cutoff rate changes between years $t-1$ and t. In case of a drop in the cutoff rate to $\mathscr{S}_t$ we have $\Pr(d_t = 1|s_t) = 1\{s_t \geqslant \mathscr{S}_t\}$, where all schools near the cutoff had poverty rates lower than the previous year's cutoff and therefore are not protected by the hold-harmless provision, in which case (3) again identifies the effect on schools that newly gain Title I eligibility.[8]

An alternative scenario described earlier, is one in which there is no year-to-year variation in treatment status either because school poverty rates and the poverty cutoff remain fixed, or because they all change by the same amount. In this case, the ratio in (3) measures the cumulative effects of receiving Title I funding for multiple years. This represents a special case of Fig. 1, where instead of being continuous in s_t, $\Pr(s_{t-1} > \mathscr{S}_{t-1}|s_t)$ is identical to the current award rate $\Pr(d_t = 1|s_t)$ which is a 0–1 step function with a jump at $\mathscr{S}_t$. The scenario corresponds to one in which treatment either equals continuous long-term receipt of Title I funds, or no receipt of such funds post-assignment. Fig. 1 could then represent the first year following assignment, with a hold-harmless provision in place, but in subsequent years this provision would no longer be relevant and $\Pr(s_{t-1} > \mathscr{S}_{t-1}|s_t)$ would become a 0–1 step function.

Therefore, while in case of random year-to-year fluctations in school poverty rates one would generally identify the average effect of gaining (and retaining) Title I status, in the case where Title I eligibility if fixed or only assessed on an infrequent basis, the estimated effects generally identifies the cumulative impact of continuous receipt of Title I funds as opposed to receiving none during the same period. An intermediate case is one where for a fraction of schools their position relative to the poverty cutoffs remains stable, while for the remainder it varies randomly across years. In such a case the changes in the school poverty rates and in the

[8]Note that this case is equivalent to one where each school's poverty rate increases by the same amount, while the cutoff rate either declines or increases by a smaller amount. In case of an increase in the cutoff rate (not shown in Fig. 2), because of the hold-harmless provision, we would not find a discontinuity at the new cutoff, but only at the previous period's cutoff rate. As will be shown later, we observe clear discontinuities in $\Pr(d_t = 1|s_t)$ at $\mathscr{S}_t$ in each year, so this case is not of relevance here.

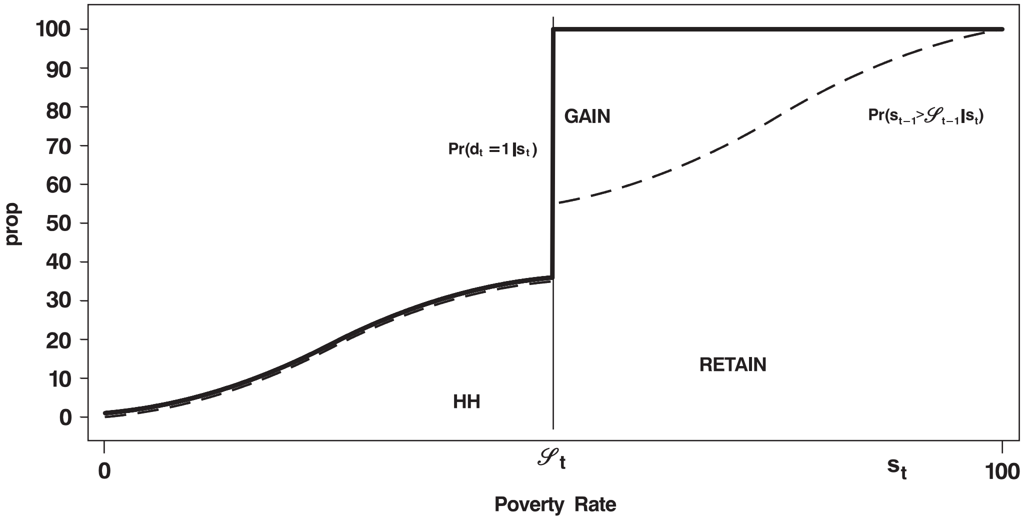

Fig. 3. Characterization of treatment effect: case with variation in school poverty rate and discontinuity in lagged recipiency rate.

poverty cutoff rate would be such that a smaller fraction of schools just left of the cutoff received Title I funds in the previous year, than among schools just right of the cutoff, leading to a discontinuity in the lagged recipiency rate at the cutoff. This scenario, where $\Pr(s_{t-1} \geqslant \mathscr{S}_{t-1}|s_t)$ is discontinuous at $\mathscr{S}_t$, is shown in Fig. 3. The local average treatment effect defined in (3) then represents a (weighted) average of the effect of long-term Title I receipt for the schools whose Title I status remains constant, and of newly gaining Title I status for the group of schools at risk of gaining or losing Title I status.[9]

The actual year-to-year variation in Title I recipiency in our data is likely to represent a combination of the scenarios described in the figures, where schools gain and lose eligibility because of both changes in the school's poverty rate, and in the district's eligibility threshold. In addition, we will find at least for one of the years considered here, that schools with poverty rates just above the cutoff rate have a significantly higher past recipiency rate than those with poverty rates just below it, indicating that poverty rate changes are strongly related to previous Title I status. While the exact source of the variation in eligibility status over time and the presence of a hold-harmless provision affect the interpretation of the effect estimate, it does not compromise the validity of the RD approach as long as schools on either side of the cutoff have comparable potential outcomes under identical sequences of Title I assignments.

How reasonable are the required continuity assumptions underlying the RD approach for our evaluation problem? If the eligibility rule for Title I receipt and the value of the poverty cutoff were well known by schools, and if schools can influence the measurement of the school attendance area's poverty rate, then the continuity assumptions underlying the RD approach may be violated. The latter would occur if schools with poverty rates just exceeding the cutoff differed on average from those just below the cutoff in characteristics other than previous Title I receipt, in such a way that their potential outcomes (average student achievement) with and without Title I funds could be expected to be different. Schools are well aware of Title I eligibility rules, and may be able to influence the calculation of the school's poverty rate to some extent. For the schools in our sample, the calculation of the poverty rate was based on the number of students in the school who received free lunch in the previous school year, and the number of children in the school attendance area from families receiving welfare benefits. The school can be expected to have some control over the former, as the number depends not only on the number of students who are eligible but also on the number of students who apply for free lunch.

A school's ability to affect its Title I status also depends on the predictability of each year's cutoff poverty rate. For the school district considered in this study, the cutoff poverty rate was calculated each year as the

[9]Note that a similar graph could be drawn to describe the case where the hold-harm provision only applies to a random subset of schools.

average poverty rate in the entire district and therefore varied across years. In case of an unknown and hard-to-forecast cutoff rate, even if each school had some control over the calculation of its poverty rate, there would be little reason to expect the continuity assumptions above to be violated.[10] If, on the other hand, it is possible to get a reasonably accurate forecast of the yearly cutoff rate, and one has sufficient control over the school's poverty rate, then schools left and right of the cutoff may differ. A divergence in previous Title I eligibility rates implies that subsequent changes in poverty rates were not independent of past Title I status, which suggests that there is scope for heterogeneous behavioral responses to reflect other school differences. I consider this possibility later when presenting the estimation results.

To estimate the local average treatment effect two estimators are adopted, the 'local Wald' estimator discussed in Hahn et al. (2001) and the two-stage estimator discussed in van der Klaauw (2002). The local Wald estimator is defined as

$$\hat{\alpha} = \frac{\bar{y}_{>\mathscr{S}} - \bar{y}_{<\mathscr{S}}}{\bar{d}_{>\mathscr{S}} - \bar{d}_{>\mathscr{S}}}, \tag{4}$$

where $\bar{x}_{<\mathscr{S}}$ and $\bar{x}_{>\mathscr{S}}$ represent, respectively, the sample average of variable x in a small interval left and right of the cutoff point. Thus the estimator is numerically equivalent to a Wald estimator applied to a subset of observations close to the cutoff (a 'discontinuity sample') which uses an indicator for having a value of s that exceeds the cutoff as instrument. Instead of optimally choosing the interval width, we implement this estimator by applying it to two different discontinuity samples.

The two-stage estimator instead uses the entire sample of observations and involves estimating a control-function augmented outcome equation

$$y_i = \beta + \delta d_i + k(s_i) + w_i, \tag{5}$$

where the control function $k(s)$ represents a specification of the conditional mean function $\mathrm{E}[u_i|s_i]$, and where d_i is replaced by a first-stage estimate of the propensity score $\mathrm{E}[d_i|s_i]$ specified as

$$\mathrm{E}[d_i|s_i] = \Pr(d_i = 1|s_i) = \gamma 1\{s_i \geqslant \mathscr{S}\} + g(s_i). \tag{6}$$

The discontinuity in the propensity score function at the cutoff value $\mathscr{S}$ is measured by γ. In implementing this estimator we specify $g(s)$ to be a continuous piecewise quadratic function of s.[11] The continuous control function $k(s)$ is estimated semi-parametrically, using a power series approximation $k(s) \approx \sum_{j=1}^{J} \eta_j \cdot s^j$, where the number of power functions, J, is estimated from the data by generalized cross-validation. The estimate of δ obtained in this manner represents an estimate of the average treatment effect defined in (3) (see van der Klaauw, 2002). Note that the two-stage estimator relies on additional smoothness assumptions on $\mathrm{E}[u_i|s]$ and $\mathrm{E}[\alpha_i|s]$. Not only does it impose continuity of the conditional expectation functions as a function of s over its whole domain, it also assumes it to be differentiable over its domain.[12]

4. Estimated impacts on student achievement

The analysis in this study is based on school-level data from the largest school district in the nation, New York City, on all public elementary and middle schools in 1993, 1997 and 2001. These data were collected by the New York City Board of Education's Office of Research, Evaluation, and Assessment, and provided by NYU's Institute for Education and Social Policy. The data set contains school level information on a set of student characteristics which varies across years. These include poverty rates, ethnic composition as well as performance measures, including grade attendance, suspension, and criminal incidence rates, and average reading and mathematics test scores and (in 1993) gain scores by grade. These data have been matched to school budget data as well as information on average teacher characteristics.

[10]For the same reason, it would be reasonable in this case to assume continuity of $\Pr(s_{t-1} \geqslant \mathscr{S}_{t-1}|s_t)$ and $\Pr(d_{t-1} = 1|s_t)$ at the poverty rate cutoff $\mathscr{S}_t$ in Fig. 1.

[11]Specifically $g(s)$ was parameterized as $g(s) = \lambda_0 + \lambda_1 s + \lambda_2 s^2 + \{\lambda_3(s - \mathscr{S}) + \lambda_4(s - \mathscr{S})^2\} \cdot 1\{s \geqslant \mathscr{S}\}$. Estimates based on a piecewise fifth order polynomial specification of $g(s)$ were essentially the same as those for the piecewise quadratic.

[12]This approach is similar to the Robinson-type estimator proposed by Porter (2003) which also assumes continuity and differentiability of $\mathrm{E}[u|s]$ and $\mathrm{E}[\alpha_i|s]$.

Table 1
School characteristics—sample means

Variable	1993			1997			2001		
	Non-Title1	Title1	t-Value	Non-Title1	Title1	t-Value	Non-Title1	Title1	t-Value
Students									
Enrollment[a]	788	867	−2.91	842	896	−1.80	810	836	−0.86
% free lunch[b]	36.5	83.7	−41.78	40.3	87.0	−36.18	39.4	84.8	−33.20
% special educ	5.1	6.1	−2.23	5.4	7.3	−4.37	4.8	7.0	−5.44
% white[c]	50.4	7.6	25.79	46.9	7.2	23.00	43.5	6.3	21.55
% black	18.5	43.8	−13.41	17.3	42.8	−13.21	17.3	41.4	−12.41
% other	31.1	48.7	−10.52	35.7	50.0	−8.06	39.1	52.4	−7.38
Staff									
Pupils per teacher	19.2	15.8	2.45	18.0	16.1	9.96	16.4	13.8	12.36
% licensed[d]	92.4	81.2	15.28	93.7	79.0	20.50	91.9	79.2	16.85
% Masters+	74.7	60.6	15.67	93.4	88.1	13.97	83.5	70.7	17.49
% tch 0–2 yrs in school				17.9	17.7	0.23			
% tch >2 yrs in school							32.8	41.8	−6.85
% tch 6+ yrs in NYC	78.7	68.6	10.26	75.5	67.5	8.98	60.4	50.2	10.39
Average teacher salary[e]	51,655	48,401	4.98	66,373	60,760	13.90	63,451	50,274	20.50
School									
% middle school	22.1	21.9	0.05	21.6	21.1	0.15	24.1	22.4	0.51
Per pupil expenditure[f]	3,623	4,125	−7.75	8,053	8,491	−3.63	8,748	8,944	−1.42
City									
Average poverty rate		62.23			66.66			68.28	
Obs	231	584		216	628		216	635	
%	28	72		26	74		25	75	

The reported t-values for each year represent t-statistics for a test of equality of means or proportions between Title I and non-Title I schools.

[a]Enrollment is measured by the October 31st register.

[b]Calculated as the percentage of students who applied for free or reduced-price lunch as of October 31.

[c]The student's ethnicity as reported by classroom teachers in their official classes as of October 31.

[d]The percentage of teachers who are licensed include tenured and probationary appointees and excludes substitute teachers including those uncertified in the subject area in which they are serving.

[e]Teacher salary amounts for 1997 and 2001 include fringe benefits and are measured in 2001 dollars.

[f]Expenditures in 1993 calculated based on mid-year budget allocations. All expenditure amounts measured in 2001 dollars.

Table 1 presents average school characteristics by Title I status for each of the three school years. The table shows a small gradual increase in the overall percentage of schools eligible for Title I funding, from 72% in the 1992–1993 school year to 75% in 2000–2001, mainly reflecting an overall increase in the skewness of the poverty rate distribution, where a larger fraction of schools have poverty rates exceeding the city average.[13] By design Title I schools have a larger percentage of students from low-income families. This corresponds to large differences in the racial composition of schools, with whites representing less than 8% of the student population in Title I schools, while they represent between 43% and 50% of students in non-Title I public schools. Title I schools generally have smaller pupil–teacher ratios, but a smaller percentage of teachers at Title I schools are licensed, have a Masters or higher degree or have more than 5 years of teaching experience in New York public schools. Not surprisingly therefore we find average teacher salaries to be lower in Title I schools. The table also shows per pupil expenditure amounts to be higher on average in Title I schools, although the difference has been declining over time.

Figs. 4–6 show for each of the three school years a piecewise quadratic fit as well as averages across 5 point (non-overlapping) intervals of the percentage of schools with Title I funding as a function of the school's poverty rate. In New York City the latter is calculated as the maximum of the percentage of students in the previous year who received free lunch, and the proportion of school aged children in the school attendance

[13]Table A1 in the appendix presents a cross tabulation of Title I recipiency status across the three years.

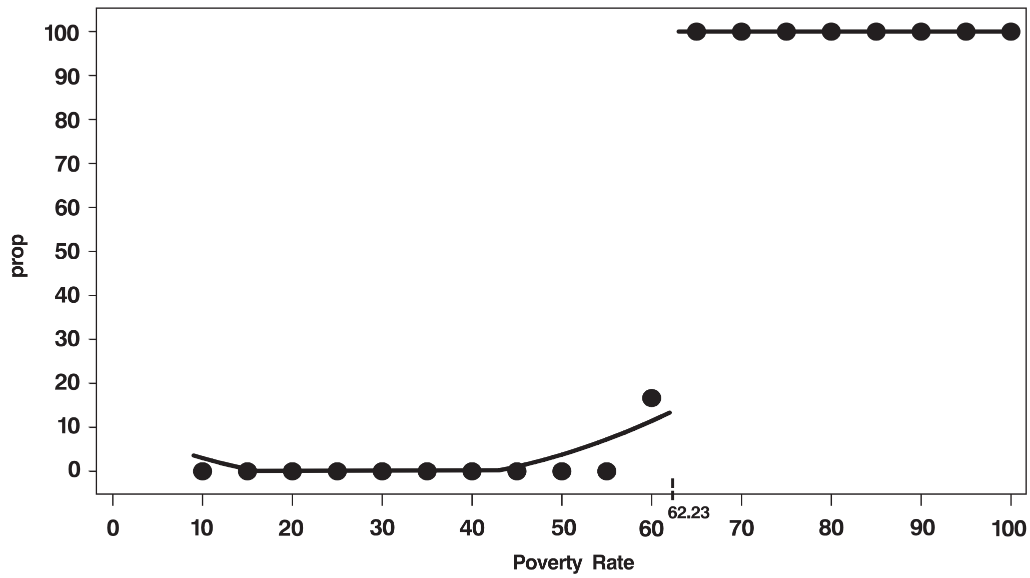

Fig. 4. Title I recipiency rate in 1993.

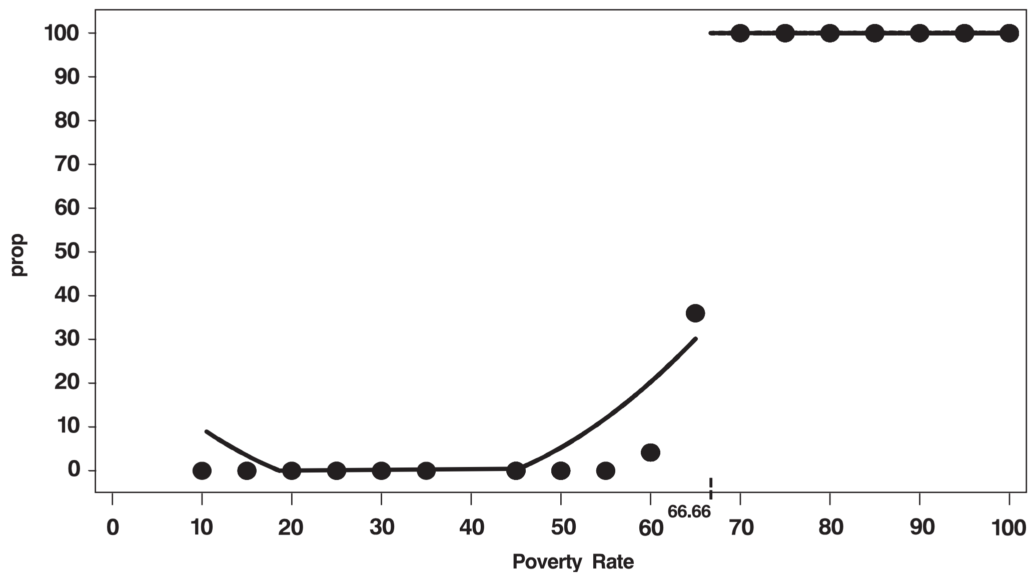

Fig. 5. Title I recipiency rate in 1997.

area from families receiving welfare (AFDC, TANF). The figures clearly show discontinuities at the poverty cutoff rate which itself increased from 62.23 in 1993 to 66.66 in 1997 and 68.28 in 2001. Notice that in 2001 none of the schools with poverty rates below the cutoff received Title I funding, reflecting the fact that the city had chosen to drop the hold-harmless provision by that time.

Table 2 present estimates of the impact of Title I status on average student outcomes for the three school years. The first column in each panel of the table presents simple OLS estimates, measuring the average difference in outcomes between Title I and non-Title I schools. Given the large differences between these two groups of schools in average student background and school characteristics, we expect these estimates to be biased. As discussed in the previous section, we can obtain more credible estimates by comparing schools with poverty rates close to the cutoff point. Columns 2 and 3 of each panel present Wald estimates based on discontinuity samples of schools within 3% and 5% points of the cutoff poverty rate in each year. Also shown

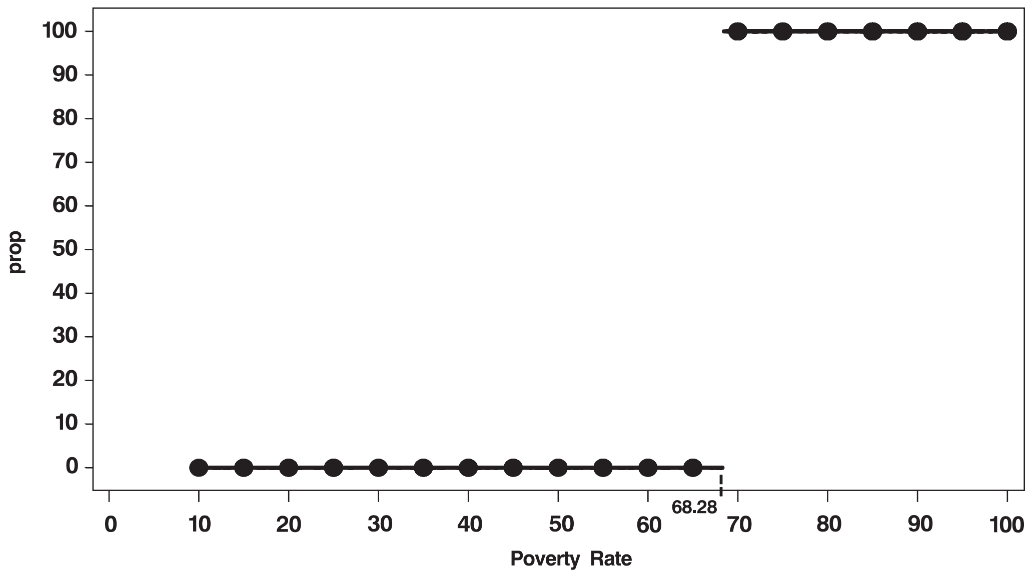

Fig. 6. Title I recipency rate in 2001.

in column 4 of each panel of the table are estimates based on the entire sample of schools obtained using the two-stage control function estimation procedure (referred to as CFE). Depending on the outcome variable, the optimal number of power functions J included in these regresssions varied from one to five, with the majority being greater or equal to three. The bottom part of each panel presents the corresponding first-stage estimates of the local Wald estimates and of γ, the discontinuity in the Title I award rate at $\mathscr{S}$ for the CFE estimator. These estimated discontinuities vary between 0.50 and 1.0 and are precisely estimated. In discussing the estimates I will focus first on the RD estimates, which will be followed by a comparison with the OLS estimates.

The first panel of Table 2 presents the effect estimates for the 1992–1993 school year. The first four outcome variables listed in the table measure the percentage of students who are held back a grade at the end of the year, the average proportion of days attended, the proportion of students who were not in school for the entire year, and the average number of days teachers were absent during the year. As many schools spend a small fraction of their Title I funds to pay for services of truant officers, social workers and counsellors, the program may have affected school attendance. The control function estimate indicates however that Title I status has led to slightly lower attendance rates, slightly higher grade repetition rates and higher rates at which students entered and left the school during the school year. Moreover the estimates imply a small positive effect on teacher absence rates. These findings, and those reported below, do not depend on the particular estimation method used. While the Wald estimates based on the 3% sample are less precise, as can be expected given the smaller samples on which they are based, they are qualitatively similar to the other two sets of estimates.

The second set of outcomes listed in first panel of Table 2 are based on standardized reading and mathematics tests results. Overall the estimates show little evidence that Title I funding leads to improved test results. In fact, most of the effect estimates imply a negative effect on performance, which in several cases is statistically significant. For example, the CFE estimates imply that being in a Title I school increases the probability that they score in the bottom two quartiles nationally by 3.55% and 5.26% points, and reduces the probability that students in grades 3 through 8 can read above the state standard by 8.25% points. Fig. 7 presents visual evidence of the latter negative effect. It shows both the raw data, as averages across 5 point (non-overlapping) intervals, and estimates of a regression of the outcome variable on a fifth degree polynomial in s and the discontinuity indicator $1\{s>\mathscr{S}\}$.[14] Similar results are found for mathematics test scores and both

[14]A piecewice quadratic specification produced an almost identical fit. Fig. 7 is representative of similar figures (not shown) for most other outcomes.

Table 2
Impact on student achievement

Outcome variable	OLS		RD Estimates					
			Local Wald 3%		Local Wald 5%		CFE	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
(A) Results for 1993								
% retained[a]	1.42	(0.21)	1.01	(0.67)	1.11	(0.51)	1.49	(0.54)
% attendance[b]	−3.94	(0.25)	−1.56	(0.94)	−2.05	(0.76)	−2.43	(0.65)
Mobility rate[c]	8.72	(0.50)	5.18	(2.71)	5.69	(1.88)	4.17	(1.30)
Teacher absence[d]	1.02	(0.11)	0.93	(0.56)	0.85	(0.42)	0.54	(0.23)
Reading scores[e]								
% gr3 > stnd	−26.22	(1.18)	−8.78	(5.05)	−6.48	(3.93)	−1.43	(2.09)
% gr6 > stnd	−23.46	(1.68)	−8.46	(5.86)	−8.71	(5.20)	−0.22	(2.88)
% gr8 > stnd	−16.28	(1.78)	−11.32	(5.51)	−12.51	(4.87)	−11.07	(3.58)
% all gr > stnd	−28.85	(0.97)	−13.14	(5.09)	−11.71	(3.72)	−8.25	(2.93)
% Q1 national	22.49	(0.97)	10.35	(3.76)	8.97	(2.92)	5.26	(2.35)
% Q2 national	5.54	(0.42)	2.77	(2.11)	2.73	(1.57)	3.55	(0.82)
Math scores[f]								
% gr3 > stnd	−17.29	(1.00)	−9.75	(4.29)	−7.26	(3.23)	−5.76	(2.92)
% gr6 > stnd	−16.41	(1.39)	−6.99	(5.17)	−7.04	(4.45)	−3.15	(2.46)
% Q1 national	20.10	(0.88)	9.24	(4.06)	8.36	(2.96)	4.30	(1.60)
% Q2 national	8.38	(0.38)	4.42	(2.12)	2.43	(1.53)	2.08	(0.79)
Reading gains[g]								
% Q1 < 0	1.87	(0.49)	0.57	(2.16)	1.15	(1.52)	1.80	(1.54)
% Q1 [0,4]	2.07	(0.38)	1.75	(1.14)	1.30	(0.92)	1.05	(0.74)
% Q1 [5,9]	2.48	(0.53)	2.41	(2.29)	1.61	(1.78)	−0.94	(1.44)
% Q2 < 0	5.07	(0.62)	3.80	(2.64)	2.87	(2.10)	2.45	(1.30)
% Q2 [0,4]	2.77	(0.37)	1.90	(1.43)	0.32	(1.39)	0.92	(0.71)
% Q2 [5,9]	3.19	(0.60)	2.36	(2.54)	−0.05	(2.01)	−2.90	(1.70)
% Q3 < 0	6.65	(0.83)	4.42	(3.29)	4.07	(2.69)	5.89	(2.28)
% Q4 < 0	5.18	(1.15)	5.89	(4.36)	4.17	(3.35)	5.77	(2.77)
% all < 0	−0.69	(0.58)	0.94	(2.68)	0.70	(2.00)	3.71	(1.68)
% all [0,4]	0.20	(0.32)	−0.00	(0.86)	−0.20	(0.73)	−0.31	(0.48)
% all [5,9]	0.26	(0.49)	1.34	(2.28)	0.47	(1.76)	−1.65	(1.67)
First stage $1\{s_i \geqslant \mathscr{S}\}$			0.76	(0.09)	0.83	(0.06)	0.86	(0.02)
Obs	815		46		70		815	
(B) Results for 1997								
Suspensions[h]	1.26	(0.37)	2.40	(2.14)	0.81	(1.33)	2.03	(0.79)
Incidents[i]	0.90	(0.20)	2.03	(1.56)	1.31	(0.78)	1.25	(0.38)
% attendance[j]	−3.50	(0.22)	−2.74	(1.55)	−1.63	(0.96)	−0.84	(0.46)
Mobility rate[k]	4.61	(0.77)	−1.03	(1.53)	−0.78	(0.94)	−2.39	(1.31)
Reading scores								
% above state level[l]	−27.79	(1.09)	−19.47	(9.62)	−11.11	(5.30)	−4.73	(2.22)
% gr 3 > stnd[m]	−24.45	(1.30)	−21.25	(9.47)	−13.35	(5.43)	−2.20	(2.51)
% gr 6 > stnd	−21.33	(1.73)	−31.14	(32.70)	−12.14	(11.06)	−8.18	(3.32)
% gr 8 > stnd	−14.31	(1.58)	0.18	(19.26)	−1.64	(5.74)	−5.46	(2.75)
Math scores								
% above state level[l]	−27.98	(1.22)	−13.53	(10.48)	−9.12	(5.90)	−5.76	(2.49)
% gr3 > stnd[n]	−10.95	(1.22)	−20.21	(18.63)	−7.11	(6.18)	−4.95	(2.35)
First stage $1\{s_i \geqslant \mathscr{S}\}$			0.50	(0.10)	0.64	(0.08)	0.70	(0.02)
Obs	809		43		64		809	
(C) Results for 2001								
Suspensions[p]	1.67	(0.38)	−0.12	(1.17)	−0.65	(0.90)	−0.99	(0.66)
Incidents[o]	0.20	(0.06)	0.01	(0.14)	0.16	(0.10)	−0.08	(0.08)
Teacher absence[q]	−0.37	(0.21)	−0.27	(0.92)	−0.83	(0.67)	−0.99	(0.41)

Table 2 (*continued*)

Outcome variable	OLS		RD Estimates					
			Local Wald 3%		Local Wald 5%		CFE	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
Reading scores[r]								
% well below stnd	14.25	(0.75)	4.43	(1.74)	2.84	(1.60)	0.40	(1.37)
% meet stnd	−29.10	(1.11)	−7.21	(3.56)	−5.49	(2.83)	0.13	(2.76)
% gr4 well below	12.45	(0.86)	4.64	(1.59)	1.17	(2.13)	−0.65	(1.78)
% gr4 meet stnd	−29.14	(1.32)	−4.38	(3.70)	−2.67	(3.28)	−4.38	(3.22)
% gr8 well below	15.01	(1.72)	1.23	(3.09)	1.07	(2.74)	1.73	(3.31)
% gr8 meet stnd	−27.21	(2.30)	−6.47	(6.61)	−8.10	(4.81)	−6.77	(6.77)
Math scores[s]								
% well below stnd	19.30	(1.18)	5.17	(3.61)	4.48	(2.74)	1.59	(2.38)
% meet stnd	−28.75	(1.26)	−6.63	(4.96)	−5.84	(3.57)	−0.22	(3.39)
% gr4 well below	11.31	(0.92)	3.13	(2.04)	1.30	(2.34)	−0.64	(1.87)
% gr4 meet stnd	−27.43	(1.47)	−4.51	(4.61)	−3.73	(3.94)	3.85	(3.62)
% gr8 well below	23.80	(2.55)	−3.47	(5.64)	−1.24	(4.38)	0.50	(4.45)
% gr8 meet stnd	−25.17	(2.17)	−3.55	(8.59)	−5.22	(6.38)	−8.67	(5.26)
First stage $1\{s_i \geqslant \mathscr{S}\}$			1.00	(0.00)	1.00	(0.00)	1.00	(0.00)
Obs	851		51		102		851	

Heteroskedasticity consistent standard errors in parentheses. The two-stage control function estimates (CFE) have been corrected for generated regressors.

[a]Percentage of general education students retained on grade at the end of the school year.

[b]Aggregate daily attendance divided by the aggregate daily register.

[c]Percentage of students who were not in the same school for the entire year.

[d]Average number of days pedagogical staff was absent during the year.

[e]Measures represent the percentage of all tested students in given grade who scored at or above the state reference point on the reading (Degree of Reading Power) test, and the percentage of all tested students in all grades who scored in given quantiles of the national DRP reading test.

[f]Measures represent the percentage of all tested students in given grade who scored at or above state reference point on the PEP (Pupil Evaluation Program) mathematics test, and the percentage of all tested students in all grades who scored in given quantiles of the national California Achievement Test (CAT) in mathematics.

[g]Percentages of all tested students in all grades in 1992 and 1993, by quantiles in 1992 who had declines, gains of 0–4 NCEs (0–3 for middle schools) or gains of 5–9 NCEs (4–6 or middle schools) on the national DRP reading test.

[h]Number of suspensions per 100 students.

[i]Number of incidents per 100 students.

[j]Aggregate daily attendance divided by the aggregate daily register.

[k]Percentage of students who were not in the same school for the entire year.

[l]Percentage of students meeting NY State Minimum Standards on Reading (DRP) and Mathematics (PEP) tests.

[m]Percentage of all tested students in grade who score at or above grade level on the Citywide (CTB) Reading Test.

[n]Percentage of all tested students in grade who score at or above grade level on the Citywide (CAT) Mathematics Test.

[o]Number of suspensions per 100 students.

[p]Number of students per 100 involved in police department incidents during school year.

[q]Average number of days pedagogical staff was absent during the year.

[r]Measures represent percentage of students meeting the standards in tested grades (3–8) in City (CTB) and State (ELA) reading tests, as well as percentage of students who score far below the standard in tested grades, showing minimal understanding of written and oral text.

[s]Measures represent percentage of students meeting the standards in tested grades (3–8) in City and State mathematics tests, as well as percentage of students who score far below the standard in tested grades, showing minimal understanding of key math ideas.

findings are robust with respect to the estimation method and sample used, with the CFE estimates typically being the smallest of the three RD estimates.

Perhaps a more informative measure of educational progress is the reading gain score. The 1993 panel of Table 2 presents estimates of the effect of Title I on proportions of students, by quartile of pre-test results, who either make negative gains or only small or moderate gains in norm-curve equivalents (NCEs)

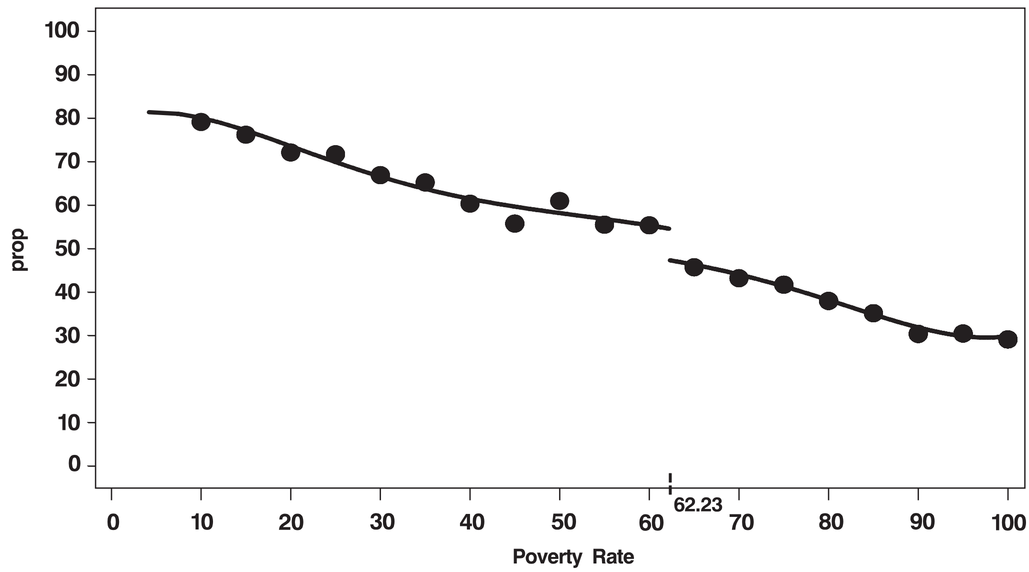

Fig. 7. Proportion reading above standard in 1993.

during the year.[15] Even though most of the estimated effects are not statistically significant at conventional levels, they suggest that in all quantiles, students in Title I schools were more likely to experience negative gains. There is no evidence that low-achieving students performed better as a result of Title I funding, nor is there evidence of positive spill-over effects for students performing in the third or fourth quartiles. The estimates appear to indicate that the adverse effect of the program were smaller for students in the bottom quartile, who are most likely to have been selected for compensatory instruction, than for those in higher quantiles. However, given the relatively large standard errors and so-called floor-and-ceiling effects, where those scoring at the bottom (top) cannot go any lower (higher), this observation should be interpreted with caution.

A comparison with the OLS estimates reveals that a simple comparison of Title I and non-Title I school would have produced large biases where some of the negative consequences associated with higher concentrations of students from disadvantaged backgrounds would have been attributed to Title I receipt.

Overall these findings suggest that in 1993 eligibility to Title I funds produced small to moderately large declines in average student performance. The evidence for the 1996–1997 school year, presented in the second panel of Table 2 reveals very similar patterns showing little evidence that Title I funds led to better student outcomes. In fact, we again find some evidence of small to moderately sized deteriorations in attendance rates, suspensions, incidents and reading and mathematics test scores. The only apparent improvement found is with respect to the student mobility rate, which falls slightly. The OLS estimates again point to the importance of controlling for the non-random assignment of Title I status, indicating that ignoring it would generally cause one to conclude that the program had much larger deleterious effects than found here.

The third panel of Table 2 presents estimates for the 2000–2001 school year. The results indicate that Title I funding did not lead to better student outcomes, whether measured by reading test scores, mathematics test scores or other outcome measures. However, when compared to the two earlier school years, their is less evidence in 2001 that the program had a negative effect on student performance. This suggests that changes in the program implemented since the 1994 reauthorization, which included the removal of an important disincentive by which improving schools were faced with reductions in funding, and a greater targeting of funds, may have helped overcome some of the negative outcomes associated with receiving Title I funding.

[15]Norm-curve equivalents are simply normalized test scores, with a mean of 50 and a standard deviation of 21.06. They range from 1.0, which is the NCE that corresponds to a percentile of 1.0 to 99.0, which is the NCE that corresponds to a percentile of 99.0.

It may also reflect the increased prevalence of schoolwide programs relative to targeted Title I programs.[16] In summary, the evidence indicates that schools that were eligible for Title I funds in these school years did not generate better student outcomes, and in fact appear to have led to performance declines at least in the 1993 and 1997 school years, with the estimates for 2001 showing less evidence of adverse effects.

It is possible that instead of affecting learning per se, a school's Title I status may affect who enters and leaves a school. The estimated effects on average student achievement could then be due to changes in the composition of students attending a school and taking tests. While this would constitute a true causal effect of Title I receipt on average school outcomes, it obviously would have very different policy implications and could lead to very different conclusions regarding the effectiveness of compensatory education programs.

While it is impossible to dismiss this possibility completely, there are several reasons to expect the role of student mobility in explaining our results to be small. First, limited evidence based on 1993 gain scores suggests that declines also occurred among students present both at the beginning and end of the year, and therefore cannot be entirely due to a change in the composition of students taking the test at the beginning and end of the school year. Second, while we found mobility rates in Title I schools to be higher in 1993 (but not in 1997), these increases were relatively small. Third, if Title I services actually lead to improved educational outcomes, one would expect attrition rates (the main component of the mobility rates) to be lower instead of higher in Title I schools. On the other hand, if Title I programs had a negative effect, one would expect parents of children most likely to receive those services (those performing below average) to be more likely to transfer their child, which would cause average achievement gains to be over-rather than underestimated. Finally, the sensitivity results presented in the next section show that the estimates are insensitive to the inclusion of average student characteristics, suggesting that at least with respect to these observed characteristics, there is little evidence that the results reported in the paper are due to selective mobility of students to non-Title I schools during these years.

5. Sensitivity analysis and interpretation

As discussed in Section 3, the validity of the RD approach relies on continuity assumptions which require schools with poverty counts just below and above the cutoff to have comparable average potential outcomes under identical Title I funding experiences. While it is not possible to test these assumptions directly, one can test whether the two groups of schools on average have similar observed characteristics. While a failure to reject equality does not guarantee that the two groups are comparable in terms of unmeasured characteristics, it may increase our confidence in the validity of RD approach. At the same time, a rejection of equality of average characteristics does not necessarily represent a violation of the continuity assumptions. For example, finding that the two groups of schools differ in terms of average total enrollment is only of concern if school size is actually related to student performance.[17] An arguably more direct way to test for the comparability of the two groups, which accounts for both the possible existence of differences in average school characteristics as well as their relevance in explaining student achievement, would be to control for such differences in observed characteristics in estimation.

Accordingly, linear controls for school characteristics were included in the outcome equation (1), which was then estimated on the two discontinuity samples with two-stage least squares, using the indicator $1\{s_i \geqslant \mathscr{S}\}$ and controls as instruments.[18] TSLS estimates obtained this way for both discontinuity samples for each year, and corresponding t-values for a test of equality between the local Wald and local TSLS estimates calculated using bootstrap methods, indicate that for 1993 in only 4 out of 50 cases the difference between the two estimates is

[16]Given that the hold-harmless provision was dropped before 2001, the 2001 effect estimate also applies to a different subpopulation of schools than the 1993 and 1997 estimates (see Section 3). In addition, as will be discussed in the next section, the correct interpretation of the estimated effect for 2001 is somewhat different.

[17]It may also be the case that some of the gaps in average current school characteristics may actually represent causal responses (by students, parents or school officials) to differences in past Title I receipts, in which case the associated differences in outcomes will be part of Title I's overall effect.

[18]The local Wald estimator defined in (4) represents a special case of this TSLS approach, where the control variables are excluded from the first-stage and outcome equations. The controls included were: school enrollment, percentage of students who are white, black, applied for free lunch, receive special education services, and an indicator for whether the school was a middle or elementary school.

statistically significant, and in none of these cases does the estimate change sign.[19] Similar results for the 1997 and 2001 school years indicate that accounting for possible differences in characteristics between groups just below and above the poverty cutoff generally does not have a significant effect on our impact estimates.[20]

While the RD approach is often interpreted and implemented as an instrumental variable approach (as exemplified by the use of the local Wald estimates above), more generally non-parametric identification and estimation of treatment effects in the case of an RD design is based on the estimation of four limits of conditional mean functions which comprise the numerator and denominator in (3). Alternative estimation approaches rely on different assumptions regarding these conditional mean functions, but the main goal of each proposed method is to obtain consistent estimates of the discontinuities in the functions $E[y|s]$ and $E[d|s]$ at $\mathscr{S}$. While the non-parametric estimators discussed by Hahn et al. (2001) and Porter (2003) allow the functional forms of each of these functions to be different on either side of the cutoff, as well as to differ from each other, more parametric approaches including the local Wald estimator as implemented here and instrumental variable approaches such as those used by Angrist and Lavy (1999) often assume the functional forms to be the same. The reliability of estimates based on these latter approaches is therefore tied to the validity of the parametric restrictions. In the case of an RD design these assumptions are likely to be particularly important, as it applies to situations where treatment is assigned based in part on a variable which itself is suspected to be directly related to the outcome of interest.

In general it is therefore important to analyze the sensitivity of estimates to variations in parametric specifications. In case of the local Wald estimator, which assumes that average student outcomes are unrelated to the school's poverty rate in a small neighborhood left and right of the cutoff, this was done by varying the size of this neighborhood. In case of the two-stage control function approach, which allows different specifications for $k(s)$ in the outcome and $g(s)$ in the first-stage Title I award equations, this is done by adopting a series approximation of $k(s)$. Analyzing the sensitivity of the estimate to changes in these specifications, I found the estimates to be very insensitive to the first-stage specification of the Title I award function. To evaluate the sensitivity to the specification of the control function $k(s)$, I also considered the case in which both $k(s)$ and $g(s)$ were specified to be continuous piecewise quadratic functions in s. In this case the two-stage estimation approach is equivalent to two-stage least squares. These TSLS estimates were generally found to be very similar to the CFE estimates, but to have slightly larger standard errors.

6. Longer term impacts and impacts on school finance

Based on the findings of this sensitivity analysis, including the robustness of the estimates to several alternative estimation methods and model specifications, we can expect the RD approach to provide credible estimates of Title I's impact on student outcomes. However, before continuing with an analysis of the impact of Title I funding on school expenditures, it would be useful to characterize in some more detail the treatment effect estimated here. As discussed in Section 3, the precise interpretation of the local average treatment effect depends on the source and nature of the variation in a school's Title I status over time. To analyze this further, we can compare past Title I recipiency among schools just right and schools just left of the cutoff poverty rate. Restricting the sample of 1997 schools to be within 5%, 3% and 1.5% points of the 1997 poverty cutoff rate, we find the 1993 recipiency rate of the latter to exceed that of the former by, respectively, 39%, 24% and −4% points. In a similar comparison for the 2001 sample we find that Title 1997 recipiency among schools just right of the cutoff in 2001 exceed those of schools just below the cutoff by 48%, 42% and 46% points, respectively. Therefore, while the 1993 recipiency rate appears to be continuous in the 1997 poverty rate, this does not appear to be the case for the 1997 recipiency rate in 2001. As discussed in Section 3 (and shown in Fig. 3) this implies that the estimated effects for that year represent a weighted average effect of newly gaining Title I status for some schools, and the cumulative effects of repeated years of funding for a significant proportion of other schools.

[19]These results and those of several additional sensitivity tests discussed below are reported in the appendix of the working paper version of this article, which is available as a website download at http://www.newyorkfed.org/research/economists/van%20der%20klaauw/

[20]When the school poverty rate was added as additional control, the TSLS estimates were found to be very similar, and the difference between the Wald and TSLS estimates was found to be statistically significant for none of the outcomes and years.

Thus while we found schools on either side of the poverty cutoff to be comparable in terms of relevant observed characteristics, they differ in their past Title I recipiency rate. The estimated effects therefore not only reflect the immediate effects of current Title I receipt, but the cumulative effects of different histories of Title I receipt. A discontinuity in the lagged recipiency rate at the cutoff, suggests the existence of a behavioral response to earlier Title I receipt as well as an ability by at least some schools to influence their current eligibility status. This would for example occur if principals who have hired a Title I teacher in a previous year spend extra effort making sure that eligible children sign up for the free lunch program to avoid having to let go of the teacher.[21] If such behavioral responses are related to school characteristics, or to expected student achievement gains, this could affect not only the interpretation but also the validity of the RD estimates. For example, if larger schools near the poverty cutoff find it easier to predict and manipulate the school's poverty rate in order to retain Title I funding, then schools left and right of the cutoff may differ in average total enrollment. However, as shown earlier, in our discontinuity samples we found Title I and non-Title schools to have comparable school characteristics. Moreover, if a greater probability of retaining Title I status among past Title I recipients is associated with higher benefits (better student performance) from receiving such funding, then this would imply that the estimated effects are overestimates of the true average impacts of Title I receipt, suggesting that the program may have had even larger deleterious effects than estimated here.

In order to differentiate between the short term and cumulative effect of funding receipt, I repeated the earlier analysis by adding controls for lagged Title I receipt. The first panel of Table 3 provides estimates of current Title I status on student achievement, conditional on Title I receipt four years earlier. Controlling for funding receipt in 1993 causes many of the 1997 estimates to change, with several of the differences being statistically significant. We find the same to be true when conditioning on 1997 status in the 2001 analysis. In all cases where these changes were significant we find the short run effects (while still not showing beneficial effects) to be less negative than the estimates reported earlier.[22] This implies that the negative effects associated with Title I receipt appear to grow with exposure.

Another way to investigate the longer term effects of Title I receipt is to analyze the impact of current Title I receipt on outcomes four years later. The first two panels of Table 4 consider, respectively, the effect of Title I receipt in 1993 on school outcomes in 1997, and of receipt in 1997 on outcomes in 2001. The estimates are similar to those reported earlier in terms of the presence and magnitude of deleterious effects, and show little evidence of any beneficial effects. Sensitivity tests, similar to those presented in Section 5 and reported in the appendix of the working paper version of this article, indicate that the results are robust to the inclusion of controls for various school characteristics and to adopting a continuous piecewise quadratic control function instead of estimating it using a polynomial series estimator.

While these results are consistent with the overall lack of evidence of a positive Title I impact in the literature, the question remains why schools eligible for Title I and receiving up to \$1000 (in 1997) per eligible pupil do not produce improvements, and appear in fact to be performing worse. To examine this issue further, Table 5 assesses the impact of Title I status on school budget and expenditure patterns in the three school years. An immediately striking finding in all three panels of the table is that gaining Title I eligibility does not necessarily lead to a statistically significant increase in average per pupil expenditures. While Title I eligibility is associated with a significant increase in Title I funding of \$300–\$400 dollars per student, overall expenditures per student show changes based on the two-stage regression estimator that vary from an increases of \$108 and \$448 in 1993 and 1997, to a decrease of \$325 in 2001. Moreover, none of these estimates are statistically significant. The estimates also indicate that Title I funding translates into very small, if any, increases in the share of total expenditures spent on teachers in the three years. While Title I receipt is associated with a 2% point increase in this share in 1993, it is associated with a 1% point drop in 1997 and no change in the share in 2001. Results from a sensitivity analysis further indicate that these findings are robust with respect to the inclusion of controls for school characteristics and to alternative RD estimation methods.

[21]It also requires an ability to accurately forecast future poverty cutoff rates.

[22]An exception is the impact on teacher absence rates where the estimate implies a much stronger deleterious effect in the short-run.

Table 3
Estimates with controls for lagged Title I receipt

Outcome variable	TSLS 3% interval		TSLS 5% interval	
	Estimate	t-Stat	Estimate	t-Stat
(A) Results for 1997				
Suspensions	2.19	0.43	0.47	1.75
Incidents	1.70	0.64	0.87	3.23
% attendance	−2.47	−1.56	−1.14	−3.89
Mobility rate	−0.84	−0.52	−0.68	−0.56
Reading scores				
% above state level	−19.59	−1.04	−10.10	−2.18
% gr 3 > stnd	−20.48	−0.31	−12.08	−1.31
% gr 6 > stnd	−25.11	−0.69	−1.56	−2.36
% gr 8 > stnd	−1.48	−0.14	4.55	−3.23
Math scores				
% above state level	−12.06	−1.41	−6.64	−3.51
% gr3 > stnd	−17.06	−0.69	−3.16	−1.72
Title I receipt 1993	0.24	(0.15)	0.39	(0.12)
Obs	42		63	
(B) Results for 2001				
Suspensions	−1.00	4.60	−1.48	7.42
Incidents	0.03	−0.52	0.14	−0.24
Teacher absence	1.13	−7.10	0.41	−11.22
Reading scores				
% well below stnd	2.74	2.51	1.67	2.66
% meet stnd	−2.57	−3.83	−2.67	−4.87
% gr4 well below	3.70	2.01	1.23	−2.10
% gr4 meet stnd	−0.01	−5.80	−0.30	−4.34
% gr8 well below	0.76	−0.43	0.42	0.43
% gr8 meet stnd	−3.32	0.75	−5.78	1.00
Math scores				
% well below stnd	0.77	3.82	0.37	6.83
% meet stnd	0.21	−4.63	−0.63	−7.61
% gr4 well below	2.39	1.56	0.74	−0.28
% gr4 meet stnd	−2.04	−2.99	−1.65	−2.29
% gr8 well below	−3.65	−0.16	−1.53	0.04
% gr8 meet stnd	0.14	0.01	−2.85	0.39
Title I receipt 1997	0.42	(0.15)	0.48	(0.11)
Obs	51		102	

Columns 3 and 5 report test statistics for the equality of Wald estimates (which do not control for lagged Title I receipt) and the TSLS estimates reported in columns 2 and 4 (which do control for past Title I status) for each discontinuity sample. The entries for the last row of either panel (reported in parentheses) are heteroskedasticy consistent standard errors. See Table 2 for variable definitions.

The absence of large and significant gains in per pupil expenditures is partly due to the fact that Title I funding on average only represents 5% of a school's total budget. But the estimates for 1997 and 2001 also suggest that Title I schools on average receive smaller non-Title I allocations.[23] Protest letters during this period indicates that there also exists anecdotal evidence that the city and the State appear to have shifted some of their own funding from Title I schools to non-Title I schools (EPP, 1999 April Letter). The 1997 and 2001 panels of Table 5 show in fact that a redistribution of PCEN funds (funds for the city's own, and much smaller compensatory education program), makes up for between 25% and 40% of the increase in funding for remedial education funded by Title I. Evidence of states and cities substituting away some of their own funding to Title I schools has also been found at the national level (Gordon, 2004).

[23]However, the impacts on non-Title 1 allocations and not statistically significant.

Table 4
Longer-term effects of Title I receipt

Outcome variable	OLS		RD Estimates					
			Local Wald 3%		Local Wald 5%		CFE	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
(A) Effect of Title I status in 1993 on 1997 outcomes								
Suspensions	1.08	(0.33)	−0.13	(1.45)	0.01	(1.02)	0.82	(0.65)
Incidents	0.85	(0.19)	−0.18	(0.68)	0.33	(0.46)	0.80	(0.32)
% attendance	−3.41	(0.21)	−1.30	(0.82)	−1.79	(0.61)	−2.03	(0.55)
Mobility rate	−3.14	(0.25)	1.50	(1.10)	1.97	(0.86)	0.35	(0.48)
Reading scores								
% above state level	−26.22	(1.06)	−10.17	(5.55)	−7.80	(4.01)	−6.49	(2.90)
% gr 3 > stnd	−23.41	(1.25)	−14.66	(6.85)	−9.44	(4.62)	−4.80	(2.28)
% gr 6 > stnd	−20.44	(1.63)	−10.65	(6.62)	−6.71	(6.26)	−9.14	(2.90)
% gr 8 > stnd	−13.50	(1.46)	−8.26	(4.82)	−9.43	(4.18)	−6.40	(2.02)
Math scores								
% above state level	−26.68	(1.18)	−13.04	(5.93)	−8.69	(4.39)	−8.26	(2.27)
% gr3 > stnd	−10.45	(1.10)	−7.03	(4.08)	−1.94	(4.16)	−3.08	(2.81)
Obs	767		46		70		767	
(B) Effect of Title I status in 1997 on 2001 outcomes								
Suspensions	1.49	(0.38)	2.11	(2.25)	0.35	(1.70)	1.21	(0.83)
Incidents	0.13	(0.05)	−0.04	(0.19)	−0.10	(0.15)	0.12	(0.09)
Teacher absence	−0.49	(0.21)	−3.06	(2.09)	−2.19	(1.20)	−1.11	(0.45)
Reading scores								
% well below stnd	12.87	(0.76)	7.67	(5.34)	4.34	(2.92)	2.20	(2.15)
% meet stnd	−26.44	(1.15)	−16.63	(10.45)	−7.59	(5.52)	−2.26	(2.42)
% gr4 well below	11.63	(0.88)	1.35	(6.49)	2.02	(3.81)	−2.20	(1.64)
% gr4 meet stnd	−27.58	(1.37)	−11.10	(11.42)	−7.76	(6.73)	1.79	(2.65)
% gr8 well below	13.01	(1.90)	3.88	(8.53)	2.89	(4.54)	2.80	(4.51)
% gr8 meet stnd	−21.09	(2.72)	−9.83	(21.81)	1.22	(9.66)	−5.47	(5.62)
Math scores								
% well below stnd	17.05	(1.17)	14.60	(8.90)	7.34	(4.82)	7.85	(4.13)
% meet stnd	−26.19	(1.30)	−23.76	(13.01)	−11.26	(6.82)	−8.43	(5.28)
% gr4 well below	10.40	(0.92)	2.01	(6.84)	1.14	(3.76)	−1.80	(1.80)
% gr4 meet stnd	−25.69	(1.49)	−17.38	(13.38)	−7.78	(7.25)	2.06	(3.03)
% gr8 well below	20.53	(2.72)	12.07	(17.20)	8.02	(8.10)	8.81	(4.95)
% gr8 meet stnd	−20.25	(2.62)	−13.77	(22.08)	−5.23	(10.69)	−7.00	(5.78)
Obs	745		39		60		745	

Heteroskedasticity consistent standard errors in parentheses. The two-stage control function estimates (CFE) have been corrected for generated regressors. See Table 2 for variable definitions.

Given the absence of a significant impact on per pupil expenditures, it is perhaps less surprising to find that Title I has been unable to fulfil its initial goal of eliminating the achievement gap. However, even if being a Title I school does not imply a significant increase in average per pupil expenditures, the Title I program itself could still have an effect, as it puts restrictions on the minimum amount of resources to be spent on low-achieving students, on the way it is to be spent, and also makes the school accountable for its students' achievements. In fact Table 5 does provide some evidence that Title I schools do indeed spend somewhat more per student on remedial education and classroom instruction in general (although this is not the case in 2001), and have slightly lower pupil–teacher ratios. While these overall increases may appear small, in terms of the remedial services provided to low-achieving students, and as a result of the increased accountability requirements, Title I could potentially have made a significant difference.

Table 5
Impact on school expenditures

Outcome variable	OLS		RD Estimates					
			Local Wald 3%		Local Wald 5%		CFE	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
(A) Results for 1993								
Per pupil total expenditure[a]	499	(71)	50	(265)	72	(233)	108	(120)
Expenditure shares								
% teachers	−2.53	(0.35)	1.70	(1.34)	1.15	(1.11)	1.84	(0.93)
% paraprofessionals	3.00	(0.24)	0.47	(0.92)	0.24	(0.70)	−0.08	(0.70)
% remedial teachers	3.63	(0.28)	1.92	(0.90)	1.16	(0.72)	1.83	(0.46)
% remedial paraprof	1.91	(0.15)	0.80	(0.52)	−0.04	(0.51)	−0.06	(0.39)
% remedial education	6.38	(0.40)	3.26	(1.11)	1.59	(0.94)	1.56	(0.86)
Average class size K-3	−0.62	(0.24)	−0.17	(1.05)	−0.12	(0.96)	−0.16	(0.54)
Average class size 4+	−1.53	(0.28)	−0.21	(1.15)	0.50	(1.23)	−0.47	(0.60)
Obs	815		46		70		815	
(B) Results for 1997								
Per pupil total expenditure	439	(142)	839	(1776)	514	(957)	448	(363)
Per pupil expenditures[b]								
Direct services	767	(39)	619	(399)	510	(254)	348	(90)
Instruction	312	(68)	496	(1210)	220	(630)	260	(190)
Teachers	118	(60)	269	(1050)	93	(553)	269	(168)
Paraprofessional	152	(14)	123	(158)	102	(84)	60	(54)
Per pupil funding by source								
Title 1	574	(16)	397	(164)	377	(90)	348	(62)
PCEN	40	(11)	−96	(89)	−124	(52)	−139	(22)
Expenditure shares[c]								
% teachers	−0.98	(0.38)	−4.42	(4.50)	−2.83	(2.50)	−1.33	(1.72)
% paraprofessionals	1.73	(0.14)	0.46	(0.99)	0.80	(0.63)	0.45	(0.46)
Pupils per teacher	−1.85	(0.20)	−1.07	(1.96)	−0.53	(1.14)	−0.68	(0.45)
Obs	809		43		64		809	
(C) Results for 2001								
Per pupil total expenditure	196	173	−90	(348)	−271	(477)	−325	(396)
Per pupil expenditures[b]								
Direct services	210	(170)	−75	(336)	−257	(467)	−306	(389)
Instruction	−14	(69)	−74	(143)	−135	(201)	−195	(170)
Teachers	−168	(60)	−72	(127)	−144	(175)	−174	(148)
Paraprofessionals	45	(12)	51	(37)	17	(41)	0	(37)
Per pupil funding by source								
Title 1	621	(19)	438	(28)	443	(31)	433	(26)
PCEN	52	(15)	−93	(70)	−97	(44)	−90	(40)
Expenditure shares[c]								
% teachers	−2.61	(0.41)	−0.62	(1.31)	−0.45	(1.14)	−0.08	(1.04)
% paraprofessionals	0.52	(0.12)	0.65	(0.42)	0.40	(0.36)	0.22	(0.35)
Pupils per teacher	−2.58	(0.23)	0.29	(0.75)	−0.03	(0.54)	−0.04	(0.41)
Obs	851		51		102		851	

Heteroskedasticity consistent standard errors in parentheses. The two-stage control function estimates (CFE) have been corrected for generated regressors. All expenditure amounts are measured in 2001 dollars.
[a]Expenditure shares represent proportions of total budget spent on teachers, paraprofessionals, compensatory education teachers and paraprofessionals, and the total fraction spent on remedial education.
[b]Per pupil amounts spent on direct school services, classroom instruction, teachers, paraprofessionals, as well as per pupil amounts received from Title 1, and PCEN (New York City's Pupils with Compensatory Education Needs) program.
[c]Expenditure shares represent proportions of total budget spent on teachers and paraprofessionals.

Although I cannot directly investigate these issues with my data, a survey of the literature on Title I funded remedial education programs point to several potential reasons for the ineffectiveness, and even adverse effects of the program on student performance. First, as indicated by Table 5, overall Title I funds comprise only a small proportion of a school's total budget. The popularity of the Title I program, despite its poor results, stems in fact largely from the fact that almost all school districts in the country are eligible for at least some Title I funds. As a result funds are spread out thinly across a very large number of schools, with many schools in affluent school districts receiving funds.

A second likely reason for its ineffectiveness, is that what since the early 1970s has been the most popular delivery model for compensatory education, the pull-out program, is by many educators considered to be an ineffective instruction method (Jendryka, 1993). It has been found to have a stigmatizing effect on students (Peterson, 1987), potentially leading to adverse outcomes. Pull-out programs have been found to add little extra instruction time (on average less than 30 min a day) and predominantly use drill and practice exercises involving basic thinking skills (Millsap et al., 1993). The additional time Title I students receive in reading and mathematics instruction, replaces the class time that regular students usually receive in more advanced subjects, such as science and social studies (LeTendre, 1991). Thus it is not clear that Title I students enjoy much of a net gain in total instruction.

Two other issues frequently reported in the education literature is that remedial classes, especially those in high-poverty schools, are often taught by inexperienced teacher aides, the majority of whom do not have college degrees (Millsap et al., 1993; Jendryka, 1993). There also have been complaints in the past about a lack of coordination between Title I teachers and regular classroom teachers (Peterson, 1987).

7. Conclusions

Almost 40 years after it was introduced, Title I remains the largest federal program for K-12 education. While the objective of the program, to eliminate the educational disadvantage associated with poverty, remains popular, a lack of evidence that the program actually has had a measurable effect on educational performance is likely to lead to further reforms.

Using school-level data from New York City covering the 1993–2001 period and a more credible evaluation methodology than those used in earlier evaluations, the estimates indicate that Title I has not led to better student outcomes. Instead results for the 1993 and 1997 school years indicate that the program may have caused students to fall further behind. But less evidence of a negative effect is found for the 2001 school year, suggesting that recent changes in the funding and operation of the program may have been beneficial. While various aspects of Title I funded remedial education programs are discussed as possible reasons for their ineffectiveness, an important finding of this study is that Title I only represents on average about 5% of a school's total budget. Moreover, the actual increase in average per-pupil expenditures associated with Title I eligibility may be even lower because state and local authorities appear to reduce their own funding for Title I schools relative to non-Title I schools. A more detailed analysis of the interrelationships between various school funding sources, in studying the effectiveness of various school programs, including those funded by Title I in school districts other than New York City represents an important area for future research.

Acknowledgments

I have benefited from helpful comments from Nicole Fortin, Howard Bloom, Guido Imbens, Thomas Lemieux, Maia Mukherjee, two anonymous referees, participants at the 2003 Regression-Discontinuity Conference in Banff and seminar participants at NYU. Research support from the University of North Carolina provided through a University Research Council grant is gratefully acknowledged. I would like to thank the Institute for Education and Social Policy at New York University for providing me with the data, and Kevin Muller for excellent research assistance.

Appendix

Table A1 presents a cross tabulation of Title I recipiency status across the three years.

Table A1
Changes in Title I recipiency status

Status 1993	Status 1997	Status 2001			Total
		Non-Title I	Title1	Missing	
Non-Title I	Non-Title 1	181	18	2	201
	Title 1	5	21	0	26
	Missing	0	0	4	4
Title I	Non-Title I	6	0	0	6
	Title 1	12	479	52	543
	Missing	0	7	28	35
Missing	Non-Title I	0	0	1	1
	Title 1	1	22	9	32
	Missing	11	88	0	99
Total		216	635	96	947

References

Abt Associates, Inc., 1997. Prospects: The Congressionally Mandated Study of Educational Growth and Opportunity.

Anderson, J.I., 1991. Using the Norm-Referenced Model to Evaluate Chapter 1. ERIC Document ED350315.

Angrist, J.D., Lavy, V., 1999. Using maimonides' rule to estimate the effect of class size on scholastic achievement. Quarterly Journal of Economics 114, 533–575.

Borman, G.D., D'Agostino, J.V., 1996. Title 1 and student achievement: a meta-analysis of federal evaluation results. Educational Evaluation and Policy Analysis 18 (4), 309–326.

Carter, L., 1984. The sustaining effects study of compensatory and elementary education. Educational Researcher 13 (7), 4–13.

Davis, A., 1991. Upping the stakes: using gain scores to judge local program effectiveness in Chapter 1. Educational Evaluation and Policy Analysis 13, 380–388.

Educational Priorities Panel (EPP), 1999. Letter on the Way Pupils with Compensatory Education Needs (PCEN) Funds are Allocated. ⟨http://www.edpriorities.org/Pubs/Opinion/Letters99/Let99_PCEN.4.9.html⟩.

Gordon, N., 2004. Do federal funds boost school spending? Evidence from Title I. Journal of Public Economics 88 (9,10), 1771–1792.

Hahn, J., Todd, P., Van der Klaauw, W., 2001. Identification and estimation of treatment effects with a regression-discontinuity design. Econometrica 69, 201–209.

Hanushek, E.A., 1998. Conclusions and controversies about the effectiveness of school resources. Economic Policy Review, vol. 4(1), Federal Reserve Bank of New York, New York, pp. 11–27.

Imbens, G., Angrist, J., 1994. Identification and estimation of local average treatment effects. Econometrica 62 (2), 467–476.

Jaeger, R.M., 1979. The effects of test selection on Title 1 project impact. Educational Evaluation and Policy Analysis 1 (2), 33–40.

Jendryka, B., 1993. Failing grade for federal aid. Policy Review 66, 77–81.

Kennedy, M.M., Birman, B.F., Demaline, R.E., 1986. The Effectiveness of Chapter 1 Services: Second Interim Report from the National Assessment of Chapter 1. US Government Printing Office, Washington, DC.

Krueger, A.B., 1998. Reassessing the View that American schools are broken. Economic Policy Review, vol. 4(1), Federal Reserve Bank of New York, New York, pp. 29–43.

LeTendre, M.J., 1991. The Continuing Evolution of a Federal Role in Compensatory Education. Educational Evaluation and Policy Analysis 13, 328–334.

Linn, R.L., 1979. Validity of inferences based on the proposed Title 1 evaluation models. Educational Evaluation and Policy Analysis 1 (2), 23–32.

Linn, R.L., 1981. Measuring pretest-posttest performance changes. In: Berk, R.A. (Ed.), Education Evaluation Methodology: The State of the Art. The Johns Hopkins University Press, Baltimore, MD.

Millsap, M.A., Moss, M., Gamse, B., 1993. Chapter 1 in public schools. The Chapter 1 implementation study. Final Report. US Department of Education, Office of Policy and Planning.

Mullin, S.P., Summers, A.A., 1983. Is more better? The effectiveness of spending on compensatory education. Phi Delta Kappan 64, 339–347.

Myers, D.E., 1986. An analysis of the impact of Title 1 on reading and math achievement of elementary school aged children. Revised, Evaluative Report, ERIC ED293956.

Peterson, P., 1987. The evolution of the compensatory education program. In: Doyle, D.P., Michie, J.S., Williams, B.I. (Eds.), Policy Options for the Future of Compensatory Education: Conference Papers. Research and Evaluation, Washington, DC, pp. 29–54.

Porter, J., 2003. Estimation in the Regression Discontinuity Model, unpublished manuscript, Harvard University.

Powers, S., Slaughter, H., Helmick, C., 1983. A test of the equipercentile hypothesis of the TIERS norm-referenced model. Journal of Educational Measurement 20 (3), 299–302.

Puma, M.J., Jones, C.C., Rock, D., Fernandez, R., 1993. Prospects: the congressionally mandated study of educational growth and opportunity. Interim Report, Abt Associates.

Sonnenberg, W.C., 2004. Federal support for education—FY 1980 to FY 2003. National Center for Education Statistics, US Department of Education.

Tallmadge, G.K., 1982. An empirical assessment of norm-referenced evaluation methodology. Journal of Educational Measurement 19 (2), 97–113.

Thistlethwaite, D., Campbell, D., 1960. Regression-discontinuity analysis: an alternative to the ex post facto experiment. Journal of Educational Psychology 51, 309–317.

US Department of Agriculture, 2006. Budget Outlays (Expenditures) from Appropriations, Loan Authorization, and Corporations and Other Revolving Funds, Fiscal Years 1987 Through 2004 and Estimates for 2005 and 2006, ⟨http://www.usda.gov/agency/obpa/Budget-Summary/2006/outlaytables.pdf⟩.

US Department of Education, 2006. Budget History, ⟨http://www.ed.gov/about/overview/budget/history/edhistory.pdf⟩.

van der Klaauw, W., 2002. Estimating the effect of financial aid offers on college enrollment: a regression-discontinuity approach. International Economic Review 43 (4), 1249–1287.