

ACCOUNTABILITY AND LOCAL CONTROL: RESPONSE TO INCENTIVES WITH AND WITHOUT AUTHORITY OVER RESOURCE GENERATION AND ALLOCATION

Susanna Loeb

(corresponding author)
Stanford University
School of Education
CERAS Building, 5th Floor
Stanford, CA 94305-3084
sloeb@stanford.edu

Katharine Strunk

Stanford University
School of Education
CERAS Building, 5th Floor
Stanford, CA 94305-3084
kos@stanford.edu

Abstract

This article examines the interaction between school accountability and local control over revenue raising and resource allocation. In particular, it asks whether accountability policies are more or less effective at improving student outcomes in states with stronger local control. Local control is operationalized with multiple measures, including the percentage of education funding from categorical grants, state supreme court rulings overturning school finance systems, local entities' taxing authority, and principals' self-assessments of their ability to control school resources. Using the National Center for Education Statistics' Schools and Staffing Survey and National Assessment of Educational Progress, the article finds that accountability policies are more effective when there is greater local control.

1. INTRODUCTION

Accountability and high-stakes testing have been at the forefront of education policy discussions for the past two decades. These reforms tie student, school, and/or district performance on assessments (usually standardized tests) to various consequences, including refusal to promote “underperforming” students to the next grade, the withdrawal of funding from poorly performing schools and districts, or, conversely, rewarding schools or districts for meeting high performance standards. At the most extreme, districts and schools are subject to closure or reconstitution. The theory behind accountability posits that these consequences (and rewards) will motivate school and district actors to enact the necessary changes to enhance student performance.

As increasing numbers of states focused on high-stakes school accountability policies throughout the 1990s, researchers began studying the effects of accountability programs (Dee 2002; Figlio 2002; Figlio and Getzler 2002; Jacob 2002; Amrein and Berliner 2003; Carnoy and Loeb 2003). With the passage of the No Child Left Behind Act (NCLB) in 2001, the federal government began requiring all states to use assessment and accountability systems, placing accountability policies at center stage in national education policy debates. This federal action has spurred a proliferation of research on school assessment systems, in conjunction with discussions of the goals, equity, and consequences of such programs (Hanushek and Raymond 2002; Hanushek and Raymond 2004).

Although the theory behind accountability policies is fairly simple, these programs are based on a number of complex assumptions about the ability of district and school personnel to enact reforms. In order to improve student performance, districts and schools must have sufficient resources as well as the flexibility to control these resources. A key assumption behind accountability plans, then, is that local actors have sufficient amounts of control over their own resources and actions to improve schools. However, a number of factors constrain educational actors, causing wide variations in the amount of local autonomy enjoyed by schools and districts. Constraints include, but are not limited to, the amount of control districts have over the allocation of their resources, the degree to which local districts can raise revenues, and principals’ ability to control their school resources through, for example, spending allocation and teacher hiring decisions.

Although education policy analysts and researchers have explored how accountability policies affect students, teachers, and administrators, as well as the creation and implementation of such policies, there is only a small body of literature on the effects of local control on student outcomes, and no articles that we know of focus on the interaction between accountability policies and measures of local autonomy. This study focuses particularly on

the interactions between accountability and local control. We measure local control, or the lack thereof, by the percentage of state aid that comes from categorical grants, local citizen governance over revenue raising, whether or not a state has seen its school finance system overturned by a state supreme court ruling, and principals' autonomy over school policies. We first ask whether stronger accountability reforms were more likely to be implemented in states that had stronger or weaker local control, and whether accountability increased or reduced this control. The article then explores the effects of accountability and control on student outcomes and asks whether accountability systems are more or less effective in states with greater local control.

2. BACKGROUND

This section provides a brief review of the relevant literature on how accountability systems, local governance, and principals' autonomy affect student outcomes.

Accountability Policies

Accountability policies are premised on the theory that an important factor contributing to low student performance is a lack of effort on the part of teachers and administrators toward reaching specific state-set goals. This lack of effort may be caused by a misalignment between the goals of teachers and administrators and those of the community or state. To combat this misalignment, accountability reforms are structured to guide teachers' and administrators' goals toward alignment with those of the state. They reward or sanction schools for their students' performance on standardized tests, which are designed to measure how successfully students are reaching state goals. These consequences, both positive and negative, are intended to provide incentives to teachers and administrators to improve student performance, thereby achieving the set goals.

Whether high-stakes accountability policies work in the manner in which they were intended is not clear. For accountability policies to succeed, the underlying misalignment must significantly and negatively affect student outcomes, and the required tests must both effectively measure state goals and be implemented in such a way as to minimize gaming of the system by teachers and administrators.

Some researchers have found that accountability policies may not be the hoped-for solution. Their studies show that accountability programs are associated with decreases in student outcomes. However, many of these studies, such as Dee 2002, examine an earlier set of reforms, including those utilizing Course Graduation Requirements and Minimum Competency Tests. Because

these reforms are so different from the more recent accountability programs, their results may not be applicable to the recent policies. Amrein and Berliner (2003) study more recent reforms and conclude that these reforms also are associated with negative consequences for students. However, their study is not designed to estimate the causal effects of accountability reforms because they look at only those states with strong accountability policies and thus have no comparison group. Other research on recent accountability reforms has found positive effects (Carnoy and Loeb 2003; Hanushek and Raymond 2004). Using the National Assessment of Educational Progress (NAEP) mathematics tests, these studies find that states with strong accountability systems saw greater increases in fourth- and eighth-grade student performance during the 1990s.

Although the Carnoy and Loeb (2003) and Hanushek and Raymond (2004) studies find positive effects of accountability strength on student performance, there is still reason for concern. Researchers are finding evidence that teachers and administrators are “gaming” the system, so increased test scores may not accurately reflect progress toward state goals but rather manipulations by local actors. For example, there appear to be increased classification of students into special education programs, causing more students to be exempt from state exams (Figlio and Getzler 2002; Hanushek and Raymond 2004). Carnoy and Loeb (2003) and Hanushek and Raymond (2004) make accommodations for this possibility in their work. Jacob and Levitt (2003) find that some Chicago public school teachers cheated in order to boost their students’ test scores. Finally, some research points to increased student dropout rates, though Carnoy and Loeb (2003) do not find strong evidence of this behavior. Students may drop out either because the school encourages them to do so in order to increase overall test scores or because students themselves give up, thinking they cannot meet the requirements for a high school degree.

In addition to concerns about the unintended consequences of accountability reforms, there is evidence that accountability is more effective in some contexts than in others. While research shows that many strong-accountability states saw test score gains, it also finds that others did not (Carnoy and Loeb 2003). One factor that may affect the implementation of accountability policies is the amount of control the local community, school administrators, and teachers have over the generation and allocation of resources.¹

1. Local control is not the only factor that may affect the implementation of accountability policies. For example, it is likely difficult to enact school reforms without sufficient revenues. Existing research has examined the effect of revenues on student outcomes, and, as a body of literature, has found inconclusive results (Coleman 1966; Hanushek 1986; Card and Krueger 1992; Hanushek 1997; Loeb and Page 2000).

Governance of Revenue Raising and Centralization of School Finance Systems

Districts with the ability to raise funds and utilize them as they see fit may be more able to tailor their educational systems to the specific needs of the students in their communities and to use resources to meet state accountability standards. Without the flexibility to raise and use funds to fit their needs, local actors may be forced to allocate resources inefficiently, leading to decreased student performance. Figlio (1997) finds that schools in states facing property tax limits on revenue raising have higher student-teacher ratios and offer significantly lower starting salaries to entering teachers. What is more, students in states subject to limitations perform consistently worse on reading, science, math, and social studies exams than do students in states without revenue-raising limits. Figlio and Reuben (2001) find additional evidence that local tax limits reduce the average quality of education majors and of new public school teachers, which is a possible cause of reduced student performance.

While tax limitations are one form of constraining local revenue-raising activities, the centralization of school financing may also limit local actors' abilities to raise and spend funds in community-specific ways. The recent trend in education finance that has shifted revenue-raising authority away from localities to state governments has decreased the flexibility of local administrators. Research on the effects of school finance reform on student outcomes suggests that student outcomes have not been improved by the centralization of revenue-raising authority. Case studies of Kansas, Michigan, California, and Texas all point to the same conclusion: there has been little to no reduction in performance disparities after the enactment of centralizing reforms that restricted local control over revenue raising (Downes 1992; Sonstelie, Brunner, and Ardon 2000; Duncombe and Johnston 2002; Imazeki and Reschovsky 2004; Cullen and Loeb 2004).

This centralization of revenue-raising authority is often due to court rulings in which opponents of local school financing have successfully challenged state school finance systems, arguing that such locally funded systems are unconstitutional (Murray, Evans, and Schwab 1998). Before 1993, fourteen states had court rulings that overturned locally funded school finance systems in favor of more centralized funding. Murray, Evans, and Schwab (1998) use the variation in the timing of state supreme court decisions to investigate the effect of such reforms on the distribution of education resources. They find that these state supreme court rulings reduced within-state inequality in spending by 19 to 34 percent.

The trend toward centralization has led to an increased state share of total education revenue. As state governments provide an increasing share of the funding, many states are placing more restrictions on the money sent to districts. While unrestricted block grants make up the majority of state-provided

funds, a substantial portion comes through categorical grants, restricting local discretion over resource allocation (Sonstelie, Brunner, and Ardon 2000). When states give education dollars in the form of categorical aid, local actors are constrained by the specific requirements attached to the monies. The percentage of education funding that comes in the form of categorical funding, then, may serve as another measure of local control in schools and districts.

Tax limitations, court rulings overturning locally funded schools, centralized revenue raising, and state-imposed restrictions on resource allocation all serve to limit local actors' flexibility to raise and spend revenues. Yet districts and schools with flexibility in the raising and spending of revenues may be more responsive to the community-specific needs raised by accountability policies. As such, district personnel may be constrained in their abilities to successfully enact accountability-inspired reforms by their (lack of) ability to raise and spend revenues for their schools. Highly centralized state revenue-raising systems that remove the direct connection between local property taxes and local school revenues, and place restrictive tax limitations and heavy constraints on the manner in which districts and schools can spend revenues may not only fail to enhance student outcomes on their own but may also impede the ability of district personnel to raise sufficient funds or interest to enact necessary accountability-inspired reforms.²

To assess whether accountability systems are more effective in states with more local control, we use a measure of whether or not citizens or local boards have the authority to raise tax revenues for schools. We also consider which states have potentially restrictive tax limits and court rulings that overturn the localized financing of schools. Finally, we explore whether accountability systems are more effective when state aid is less restricted as measured by a lower percentage of categorical aid.

Principal Influence over School Policies

Once districts have (or do not have) resources, it may matter *who* has control over how to spend it, as well as over other important school-level decisions. The theory behind accountability policies assumes that school-level administrators will have the means to respond to incentives in order to increase student performance. However, the ability of school-level actors to successfully implement accountability policies may be hindered or helped by their degree of autonomy over important school-level decisions. If principals do not have the flexibility to enact reforms, accountability policies may be less effective.

2. Some research finds that spending equalization leads to improved test scores for certain student groups. For example, Card and Payne (2002) find spending equalization leads to a reduction in the test score gaps between students of different family backgrounds.

There is little literature regarding the extent of principal control or its association with state-level variables. Hannaway (1993) espouses a theoretical relationship between principal and teacher influence over school policies and revenues, parent and union influence over school policies, the proportion of minorities in a school, and the school's urbanicity. She posits that "political pressure" variables may moderate the abilities of principals and teachers to influence school outcomes. Hannaway examines how the level of perceived authority of school-level decision making is related to these "pressure" variables and concludes that principals in schools that are more dependent on state and federal aid, located in an urban center, and faced with stronger unions are less likely to have authority over school-level decision making. The findings point to the possibility of important interactions between local influence over school-level decisions and accountability policies. The work also highlights the need for more research on the interactions between principal and teacher autonomy and political context variables.

Summary

There is a good deal known and being discovered about the direct effects of accountability and local discretion over revenue raising on student and school outcomes. Less is understood about the effects of principal autonomy and local discretion over resource allocation more broadly, and almost no work exists on the interaction effects of all three variables together. In order to unravel the questions of whether, how, and why accountability policies enhance student outcomes, this article asks how state accountability, governance of revenue raising, and the degree of local control over budget and hiring decisions separately and together influence student outcomes.

3. ANALYTIC METHODS

This study has two parts. The first explores the relationship between accountability and measures of local control, asking whether stronger accountability reforms were more likely to be implemented in states that had stronger or weaker local control and whether accountability increased or decreased this control. The second assesses the effect of accountability and control on student outcomes and asks whether accountability is more or less effective in states with greater or lesser local control. We define greater local control as discretion over spending (fewer funds in categorical aid), decentralization of school finance systems (local ability to vote on education finance, the absence of tax limits for education, and court rulings overturning local finance structures), and greater principal autonomy over school-level decisions, specifically regarding hiring and spending.

Although these measures gauge aspects of local control, there are other, especially nonfinancial, facets of local control that are not captured by these measures. Decentralized states and states without tax limits may have more control over revenue raising, yet it is possible that districts in such states still do not have control over other important decisions such as curriculum. States may be making the curricular decisions even when the local districts are collecting the revenues. We focus here on the financial aspects of local control, leaving other aspects of control for later research.

The Relationship between Accountability and Control

This simple analysis first tests whether states that implemented stronger accountability systems had more or less local control. We then ask whether accountability changed local control, using a regression framework in which local control in 1999 is modeled as a function of local control in the early 1990s and accountability strength. We do not control for the strength of accountability policies in the early 1990s because in many states no accountability policies either existed or were carefully documented. These analyses are most effective for our measures of principal control because voting rights, tax/revenue-raising limits, and the status of court decisions did not change substantially over this time period. In addition, we only have data on categorical grants for one year and are therefore unable to assess whether or not accountability policies affected the percentage of these revenues.

Effects on Student Outcomes

In the second part of the study we estimate student achievement gains at the state level as a function of accountability, control, and other measures found to affect student outcomes and accountability (e.g., percentage of black and Hispanic students and population size). The estimates are based on the following equation:

$$\Delta Y_{t-(t-1)} = \gamma_0 + \gamma_1 Y_{s(t-1)} + \Delta X_{s(t-(t-1))} \gamma_2 + \gamma_3 A_s + L_s \gamma_4 + (A_s \times L_s) \gamma_5 + \varepsilon_{st}$$

where Y_{st} is an aggregate student outcome measure in state s at time t ; X_{st} represents a vector of time-varying measures found to be significant predictors of student achievement and to covary with accountability strength, such as student demographics and dollars of per pupil revenues;³ A_s is a measure of

3. We also ran models that included controls for the proportion of schools' funds coming from the state rather than local sources in 1995 and the percentage of the state that was in poverty in 1993. Neither of these variables was significant in predicting either accountability strength or outcomes.

accountability strength; and L_s is a vector of local control measures. ε_{st} is the residual error. Because of the small state-level sample with which we are able to work, we lose power by including all of the local control variables at once. Instead, we run separate regression analyses for the effect of accountability and different local measures on student outcomes, such that L_s signifies just one measure of local control in most specifications.

We run the analyses separately for white and black students and for fourth- and eighth-grade math performance. We also “stack” our data, expanding our sample from 51 states to 204 observations, with four observations for each state: one each for black fourth graders, black eighth graders, white fourth graders, and white eighth graders. This gives our regression models added power. However, because “stacking” the data generates four observations for each state, we cluster our standard errors at the state level to avoid possible misestimation of our standard errors.

A_s is the case with any state-level analysis of policy effects, this study is subject to a variety of challenges. Small sample size at the state level limits the number of controls we can include. Our regressions include only those additional variables found to affect the coefficients on the variables of interest in exploratory analyses and in Carnoy and Loeb’s earlier study. In addition, accountability policies vary substantially from state to state, and the accountability strength index we use is an imperfect measure of these differences. Our measures of local control are also subject to criticism, as they are clearly measured with error. However, this study is intended to provide a starting point for assessing a potentially important policy question. Further work is needed to improve these measures and to find alternative methods for assessing the interactions between accountability policies and local control, perhaps making use of variation across districts and schools.⁴

4. DATA

We combine eight sets of data for the analyses: the National Assessment of Education Progress (NAEP); Carnoy and Loeb’s accountability index; the Common Core of Data; data from the National Center for Education Statistics’ (NCES) *Public School Finance Programs of the United States and Canada: 1998–1999*; local taxing authority data from Randall Reback; David Figlio’s 1997 categorization of states that have imposed “tax revolt–era limits on school districts”; Murray, Evans, and Schwab’s (1998) classification of states’ supreme court

Due to the limited power of our sample, we left both control variables out of our final analyses. Analyses including these measures are available upon request from the authors.

4. We attempt to examine within-state variation through the use of the National Longitudinal School-Level State Assessment Database (www.schooldata.org). However, outcome data going back to the 1993 school year are only available for three states, and the data for those states are very limited.

rulings on the constitutionality of school finance systems; and the Schools and Staffing Surveys (SASS) from 1993–94 and 1999–2000.

Outcomes

For student data, we use the posted NAEP mathematics test scores by state. These data are collected approximately every four years in mathematics and reading in the fourth, eighth, and twelfth grades. Students are sampled with a multistage stratification design on a representative sample of student populations of interest.⁵ NAEP data have only been available by state since 1990, and states were not required to participate in the testing until the passage of the No Child Left Behind Act. Because the NAEP math test was given in both 1996 and 2000, it provides a measure of the effects of state accountability systems, many of which were enacted in the mid-1990s. We use the percentage of students that receive a “basic” or “proficient” score on the NAEP mathematics test. NAEP also reports the percentage of students reaching an advanced level, but these numbers are too small to be used effectively as separate outcomes for the analyses. The National Center for Education Statistics (NCES) provides the NAEP scores divided by ethnicity and race.

In addition to using the percentage achieving basic and the percentage achieving proficient, we use an average scale score adjusted for exclusion rates. McLaughlin (2001) has estimated an imputed set of fourth- and eighth-grade math NAEP scale scores for 1996 and 2000 by state, assuming that all excluded students had taken the test without accommodation. His imputations are made on the basis of information provided on whether or not each student in the sample was included or excluded from the tests. McLaughlin uses data taken from student and teacher survey responses, including questions about why the students are excluded from the tests. We use his imputed math scores to reestimate the regression equations. These three outcome measures can be roughly interpreted as an examination of students at the low (percentage passing at the basic level), mean (adjusted scale scores), and high points of the student performance distribution curve.

Table 1 shows that between 28 and 43 states have math test results in each NAEP year. The scores are slightly higher in 2000 than in 1996, and the scores for white students are higher than those for black students. The average gain from 1996 to 2000 in the percentage of students achieving at least the basic level was 5.5 percent. In 2000, a state average of 76.7 percent of white and 73.7 percent of black fourth-grade students and 74.4 percent of white and 33.1

5. See <http://nces.ed.gov/nationsreportcard/pubs/guide> (accessed July 2005).

Table 1. Descriptions of Outcome Variables

| Variables (SD) | Obs | Min | Max | Mean (SD) | Obs | Min | Max | Mean |
|---|--------------|-------|-------|--------------|--------------|-------|-------|------------|
| | Fourth Grade | | | | Eighth Grade | | | |
| <i>NAEP McLaughlin-Adjusted Level Scores</i> | | | | | | | | |
| 1996 Whites (6.9) | 43 | 218.5 | 238.4 | 227.1 (5.2) | 33 | 261.3 | 285.7 | 275.8 |
| 1996 Blacks (11.2) | 43 | 185.3 | 215.6 | 199.5 (6.8) | 33 | 227.0 | 275.0 | 243.6 |
| 2000 Whites (7.0) | 40 | 220.5 | 239.5 | 229.9 (5.2) | 37 | 263.0 | 290.0 | 278.4 |
| 2000 Blacks (11.0) | 41 | 191.1 | 229.2 | 204.2 (7.8) | 37 | 227.0 | 273.8 | 247.6 |
| <i>Change in McLaughlin-Adjusted Scores 1996-2000</i> | | | | | | | | |
| Whites | 35 | -3.6 | 10.1 | 3.4 (3.2) | 33 | -3.5 | 11.1 | 2.2 (3.2) |
| Blacks | 36 | -3.8 | 14.6 | 4.6 (5.1) | 33 | -9.7 | 23.0 | 3.8 (6.6) |
| <i>NAEP Math Levels¹</i> | | | | | | | | |
| 1996 % Basic—Whites (7.2) | 43 | 62 | 86 | 71.8 (6.5) | 40 | 55 | 81 | 70.5 |
| 1996 % Basic—Blacks (7.1) | 35 | 18 | 46 | 30.3 (6.5) | 30 | 15 | 43 | 27.3 |
| 2000 % Basic—Whites | 40 | 65 | 89 | 76.7 (6.2) | 39 | 59 | 85 | 74.4 (6.9) |
| 2000 % Basic—Blacks (8.7) | 31 | 21 | 61 | 37.7 (8.9) | 28 | 17 | 49 | 33.1 |
| 1996 % Proficient—Whites (6.8) | 43 | 13 | 38 | 23.8 (5.9) | 40 | 12 | 37 | 26.7 |
| 1996 % Proficient—Blacks (2.3) | 35 | 1 | 10 | 3.8 (1.8) | 30 | 1 | 11 | 4.1 |
| 2000 % Proficient—Whites (7.2) | 40 | 16 | 41 | 28.0 (6.5) | 39 | 14 | 43 | 29.7 |
| 2000 % Proficient—Blacks (2.5) | 31 | 2 | 11 | 5.3 (2.6) | 28 | 1 | 12 | 5.5 |
| <i>Change in NAEP Levels 1996-2000</i> | | | | | | | | |
| % Basic—Whites | 35 | -2 | 13 | 5.3 (3.9) | 33 | -1 | 14 | 4.5 (3.7) |
| % Basic—Blacks | 27 | -9 | 20 | 7.4 (6.8) | 23 | -8 | 18 | 5.1 (6.4) |
| % Proficient—Whites | 35 | -3 | 12 | 4.5 (3.6) | 33 | -2 | 13 | 3.7 (3.0) |
| % Proficient—Blacks | 27 | -5 | 7 | 1.7 (2.5) | 23 | -3 | 5 | 1.3 (2.2) |
| Stacked Fourth and Eighth Grades | | | | | | | | |
| <i>NAEP McLaughlin-Adjusted Level Scores</i> | | | | | | | | |
| 1996 | 152 | 185.3 | 285.7 | 233.5 (28.5) | | | | |
| 2000 | 155 | 191.1 | 290.0 | 238.9 (28.2) | | | | |
| <i>Change in McLaughlin-Adjusted Scores 1996-2000</i> | 137 | -9.7 | 23.0 | 3.5 (4.8) | | | | |
| <i>NAEP Math Levels</i> | | | | | | | | |
| 1996 % Basic | 148 | 15 | 86 | 52.6 (22.1) | | | | |
| 2000 % Basic | 138 | 17 | 89 | 58.4 (21.3) | | | | |
| 1996 % Proficient | 148 | 1 | 38 | 15.6 (11.7) | | | | |
| 2000 % Proficient | 138 | 1 | 43 | 18.8 (12.8) | | | | |
| <i>Change in NAEP Levels 1996-2000</i> | | | | | | | | |
| % Basic | 118 | -9 | 20 | 5.5 (5.2) | | | | |
| % Proficient | 118 | -5 | 13 | 3.0 (3.2) | | | | |

1. NAEP test score is taken from <http://nces.ed.gov/nationsreportcard/states>.

percent of black eighth-grade students reached this basic level. On average, only 28.0 percent of white and 5.3 percent of black fourth-grade students and 29.7 percent of white and 5.5 percent of black eighth-grade students achieved the proficient level in 2000.

Accountability

Our measure of accountability strength comes from a database developed by Carnoy and Loeb (2003) using 1999–2000 information from the Consortium

for Policy Research in Education (CPRE).⁶ The scale of accountability ranges from 0 to 5, based on year 2000 accountability conditions, with states such as Iowa and Nebraska, that did not have any state-level accountability requirements for schools or districts, coded zero, and states with the “maximum” state level demands on schools and that require a high school competency exam for graduation, such as Texas, North Carolina, New Jersey, and Florida coded five.

The zero-to-five scale captures the degree of state external pressure on schools to improve student achievement according to state-defined performance criteria. States receiving a zero did not test students statewide or did not set statewide standards for schools or districts. States that required state testing in the elementary and middle grades and the reporting of test results to the state, but no school (or district) sanctions or rewards (weak external pressure or none at all), score a one. States that tested at the elementary and middle school levels and had moderate accountability sanctions/rewards or, alternatively, a high school exit exam (that sanctions students but pressures schools to improve student performance) earn a two. States that tested at the lower and middle grades, have moderate accountability repercussions for schools and districts, and require an exit exam in high school receive a three. Those that tested and placed strong pressure on schools or districts to improve student achievement (threat of reconstitution, principal transfer, loss of students) but did not require a high school exit test score a four. States receiving a five tested students in primary and middle grades, strongly sanctioned and rewarded schools or districts based on improvement in student test scores, and required a high school minimum competency exit test for graduation. Table 2 shows that the states’ average index in 2000 was 2.12 on a scale of zero to five.

Local Control

We measure local control by (1) the percentage of total education revenue tied to categorical aid, (2) tax limitations, (3) whether or not a state’s supreme court has overturned the constitutionality of the state’s school finance system, (4) whether or not local voting for school revenues is allowed, and (5) principals’ reports of control over school-level decisions.⁷

6. See http://www.cpre.org/Publications/Publications_Accountability.htm (accessed July 2005).

7. Realizing that teachers’ unions may also play an important role in the amount of local control school or district personnel have over decision making and the effect of accountability policies on student outcomes, we also run analyses looking at the extent of unionization in the state. We use several measures of unionization: SASS’s designation of districts as having collective bargaining agreements, meet-and-confer agreements, or neither; whether or not a state allows collective bargaining; and whether or not a state is a right-to-work state. Although we find some evidence that unions are weaker in states with stronger accountability policies, our analyses show inconclusive and insignificant effects of the degree of unionization and its interaction with accountability policy

Table 2. Descriptions of Explanatory Variables (*n* = 50)

| Variables | 1993 | | 1999 | |
|---|-------|-------|-------|------|
| | Mean | SD | Mean | SD |
| Accountability ¹ | — | — | 2.12 | 1.44 |
| <i>State Education Finance</i> | | | | |
| Per pupil revenue (1000s) ² | 6.36 | 1.31 | 8.40 | 1.78 |
| Expenditures per pupil (1000s) | 5.67 | 1.29 | 7.33 | 1.57 |
| Instructional only per pupil (1000s) | 2.52 | 0.56 | 3.18 | 0.68 |
| <i>Measures of Teacher Quality³</i> | | | | |
| '93 district base salary (1000s) | 24.47 | 3.26 | 28.77 | 3.44 |
| '93 ln base salary | 10.10 | 0.13 | 10.26 | 0.12 |
| <i>Centralization of School Finance</i> | | | | |
| Just citizen vote | — | — | .25 | — |
| Just board vote | — | — | .20 | — |
| Both vote | — | — | .41 | — |
| Neither vote | — | — | .12 | — |
| States with tax limits | .43 | — | .43 | — |
| States with court decisions | .27 | — | — | — |
| Percent of funds—Categorical | — | — | .12 | .06 |
| % of ed. financing from state ⁴ | 0.41 | 0.18 | 0.53 | 0.16 |
| <i>Principal Influence Measures⁵</i> | | | | |
| <i>Principal</i> | | | | |
| Average | 4.25 | 0.17 | 4.32 | 0.10 |
| Hiring | 4.46 | 0.32 | 4.65 | 0.27 |
| Spending | 3.99 | 0.31 | 4.24 | 0.26 |
| Principal by district (avg.) | 1.49 | 0.19 | 1.43 | 0.15 |
| Principal by state (avg.) | 2.20 | 0.37 | 1.94 | 0.26 |
| <i>District by state</i> | | | | |
| Average | 1.68 | 0.24 | 1.49 | 0.21 |
| Hiring | 2.15 | 0.35 | 2.01 | 0.31 |
| Spending | 1.57 | 0.30 | 2.15 | 0.35 |
| <i>Demographic Variables</i> | | | | |
| Percent of students in poverty ⁶ | 0.45 | 0.12 | 0.36 | 0.11 |
| Percent black and Hispanic | 0.23 | 0.18 | 0.26 | 0.18 |
| 1995 state population (100,000s) | 52.44 | 57.59 | — | — |

1. http://www.cpre.org/Publications/Publications_Accountability.htm. The Accountability index ranges from 0 to 5, with 5 implying the strongest level of accountability.
2. Per pupil revenue, expenditures, and instructional expenditure data are taken from the Common Core of Data at <http://nces.ed.gov>. The “before” data are from 1996, and the “after” data are from 2001, chosen to best match the NAEP outcomes data.
3. District pay scale. All salary numbers are in unadjusted nominal dollars.
4. The percentage of total spending on education finance funded by the state government can be found at <http://nces.ed.gov>. It is important to note that the “1993” data are actually from 1963 and the “1999” data are from 1995. Again, these are the closest year matches to SASS feasible with these data.
5. These variables are taken from the SASS data set at <http://nces.ed.gov/SASS>.
6. This variable is from the SASS data set at <http://nces.ed.gov/SASS>.

We obtain the percentage of education revenues tied to categorical aid from the NCES *Public School Finance Programs of the United States and Canada: 1998–1999* (NCES 2001). There is a brief chapter on each state’s school finance

strength on student outcomes. This is possibly the result of our limited data sample and weak unionization indicators and is a topic for further research.

program, from which it is possible to generate an estimate of the percentage of categorical aid. We define categorical aid as the specific categorical aid programs in each state plus funds that are earmarked for technology, aid to private and alternative schools, and accountability and standards programs. It is important to note that we can calculate this measure as a percentage of total *reported* state and local revenues. We cannot account for private donations from parents and/or foundations. The mean percentage of categorical aid is 11.94%, with a wide range of categorical funding, from 0.44% (Oregon) to 31.61% (Washington, DC). Table 2 presents descriptive statistics of our explanatory variables, and table 3 gives the local control variables by state.

We measure the degree of centralized governance over education revenue raising using data from a survey gathered by Randall Reback at Columbia University. From this survey we are able to ascertain whether or not citizens directly vote on local taxes to support regular instruction in public schools, and whether or not elected boards such as school boards or town meeting representatives vote on local taxes to support regular instruction. It is difficult to determine which voting status—whether or not citizens alone, citizens and boards together, or just boards have the right to vote on local taxes to support instruction—indicates the lowest level of centralization. However, we believe that it is accurate to assume that states with neither citizen nor board votes have the highest level of centralization. As table 2 indicates, 12 percent of states allow neither citizen nor board voting for local revenue raising for schools (no vote), while 41 percent allow both types of votes (both vote). A quarter of states just allow citizens to vote (just citizen vote), and the remaining 20 percent allow just board voting (just board vote).

We also use a measure of whether or not a state has imposed a potentially restrictive tax limitation, taken from Figlio 1997. The states that he considers to have imposed “tax revolt-era limits on school districts” are outlined in table 3. Theoretically, it is important to be able to ascertain the effect of varying levels of district discretion over taxation on student outcomes. The state-by-state indices from which the dummy variables are created do so. Practically, while it would be preferable to be able to utilize a more in-depth measure, none exists at this time. Table 2 shows that in 1993, 43 percent of states were subject to restrictive tax limits.

In addition, we use a measure, taken from Murray, Evans, and Schwab 1998, of whether or not a state’s supreme court had overturned the state’s school finance system. Opponents of locally financed school systems had successfully challenged the state school finance system by 1993 in California, Connecticut, Kansas, Kentucky, Massachusetts, Montana, New Hampshire, New Jersey, Tennessee, Texas, Washington, West Virginia, Wisconsin, and

Table 3. Local Control Variables by State

| State | Percent Categorical Funding | Percent Total Rev from State | Voting Provisions | Tax Limits | Successful Court Challenge | Average Principal Control—'93 | Average Principal Control—'99 |
|---------------|-----------------------------------|------------------------------------|----------------------|---------------|----------------------------------|-------------------------------------|-------------------------------------|
| Alabama | 0.06 | 0.71 | Both | 0 | 0 | 4.08 | 4.33 |
| Alaska | 0.06 | 0.71 | Board | 0 | 0 | 4.26 | 4.18 |
| Arizona | 0.04 | 0.49 | Citizens | 1 | 0 | 4.29 | 4.40 |
| Arkansas | 0.01 | 0.67 | Both | 1 | 0 | 4.14 | 4.20 |
| California | 0.16 | 0.58 | Neither | 1 | 1 | 4.37 | 4.26 |
| Colorado | 0.04 | 0.45 | Citizens | 1 | 0 | 4.43 | 4.37 |
| Connecticut | 0.09 | 0.38 | Both | 0 | 1 | 4.26 | 4.34 |
| Delaware | 0.19 | 0.68 | Citizens | 0 | 0 | 4.23 | 4.26 |
| D.C. | 0.32 | — | Neither | — | 0 | 3.93 | 4.47 |
| Florida | 0.11 | 0.53 | Board | 0 | 0 | 4.29 | 4.37 |
| Georgia | 0.16 | 0.55 | Board | 0 | 0 | 4.21 | 4.37 |
| Hawaii | 0.15 | 0.97 | — | 0 | 0 | 4.58 | 4.51 |
| Idaho | 0.07 | 0.66 | Citizens | 1 | 0 | 4.34 | 4.45 |
| Illinois | 0.10 | 0.36 | Both | 1 | 0 | 4.32 | 4.31 |
| Indiana | 0.10 | 0.54 | Citizens | 1 | 0 | 4.24 | 4.33 |
| Iowa | 0.22 | 0.49 | Both | 0 | 0 | 4.38 | 4.29 |
| Kansas | 0.11 | 0.60 | Both | 1 | 1 | 4.30 | 4.32 |
| Kentucky | 0.17 | 0.69 | Board | 0 | 1 | 4.24 | 4.42 |
| Louisiana | 0.11 | 0.60 | Both | 1 | 0 | 4.13 | 4.40 |
| Maine | 0.10 | 0.48 | Both | 0 | 0 | 4.45 | 4.36 |
| Maryland | 0.10 | 0.39 | Neither | 0 | 0 | 3.93 | 4.31 |
| Massachusetts | 0.14 | 0.31 | Both | 1 | 1 | 4.37 | 4.28 |
| Michigan | 0.12 | 0.71 | Citizens | 1 | 0 | 4.19 | 4.32 |
| Minnesota | 0.15 | 0.54 | Citizens | 1 | 0 | 4.27 | 4.26 |
| Mississippi | 0.15 | 0.66 | Both | 1 | 0 | 4.26 | 4.37 |
| Missouri | 0.14 | 0.40 | Both | 1 | 0 | 4.29 | 4.19 |
| Montana | 0.05 | 0.55 | Citizens | 0 | 1 | 4.33 | 4.35 |
| Nebraska | 0.12 | 0.33 | Board | 0 | 0 | 4.34 | 4.33 |
| Nevada | 0.16 | 0.66 | Neither | 0 | 0 | 4.36 | 4.42 |
| New Hamp. | 0.02 | 0.07 | Citizens | 0 | 1 | 4.56 | 4.33 |
| New Jersey | 0.13 | 0.38 | Both | 1 | 1 | 4.30 | 4.39 |
| New Mexico | 0.25 | 0.84 | Neither | 1 | 0 | 4.28 | 4.39 |
| New York | 0.20 | 0.42 | Both | 0 | 0 | 4.11 | 4.35 |
| N. Carolina | 0.13 | 0.69 | Board | 0 | 0 | 4.20 | 4.36 |
| North Dakota | 0.06 | 0.45 | Both | 1 | 0 | 4.20 | 4.32 |
| Ohio | 0.12 | 0.41 | Both | 1 | 0 | 4.08 | 4.22 |
| Oklahoma | 0.19 | 0.65 | Citizens | 0 | 0 | 4.13 | 4.22 |
| Oregon | 0.00 | 0.50 | Citizens | 1 | 0 | 4.40 | 4.29 |
| Pennsylvania | 0.11 | 0.41 | Board | 0 | 0 | 4.10 | 4.21 |
| Rhode Island | 0.06 | 0.42 | Board | 0 | 0 | 3.95 | 4.14 |
| S. Carolina | 0.11 | 0.50 | Both | 0 | 0 | 4.40 | 4.35 |
| South Dakota | 0.06 | 0.29 | Both | 0 | 0 | 4.32 | 4.29 |
| Tennessee | 0.03 | 0.52 | Both | 0 | 1 | 3.94 | 4.32 |
| Texas | 0.16 | 0.46 | Board | 0 | 1 | 4.27 | 4.49 |
| Utah | 0.20 | 0.58 | Both | 1 | 0 | 4.42 | 4.41 |
| Vermont | 0.21 | 0.26 | Citizens | 0 | 0 | 4.43 | 4.27 |
| Virginia | 0.13 | 0.43 | Neither | 0 | 0 | 4.24 | 4.29 |
| Washington | 0.16 | 0.73 | Citizens | 1 | 1 | 4.33 | 4.41 |
| West Virginia | 0.05 | 0.69 | Citizens | 0 | 1 | 3.65 | 3.96 |
| Wisconsin | 0.06 | 0.46 | Both | 0 | 1 | 4.29 | 4.28 |
| Wyoming | 0.12 | 0.51 | Both | 0 | 1 | 4.46 | 4.22 |
| Mean | 0.12 | 0.53 | | 0.43 | 0.27 | 4.25 | 4.32 |

Wyoming. Murray, Evans, and Schwab (1998) provide a thorough explanation of these cases and their effects on the distribution of educational resources.

Our principal control variables come from the 1993–94 and the 1999–2000 NCES Schools and Staffing Survey (SASS). SASS is the largest cross-sectional sample survey of public, public charter, private, and Native American

elementary and secondary schools in the United States.⁸ We average principals' perceptions of their own control over school management issues such as curriculum development, spending, teacher hiring, standards, teacher evaluation, professional development, and student discipline policies. In this analysis we use both principals' perceptions of their own control and their perceptions of their control relative to their assessment of the influence of district and state Departments of Education. These variables are ranked by principals on a one-to-five scale, with one being the perception of the least amount of influence. When a relative measure is used, one (1) signifies that the principal believes that he or she has the same level of influence as the state or district, and less (more) than one (1) signifies that the principal believes that he or she has less (more) influence than the state or district.

Principals believe they have substantial control over most areas of schooling. In 1993, principals rank themselves as having an average control level of 4.25 on a scale from one to five across measures (see table 2). This is more than double the control they attribute to the state and one and a half times the control they attribute to the district. The average measure of principal control and principal control relative to the district stayed approximately the same over time, while principals' assessment of their authority relative to the state fell on average from 2.2 times as much control in 1993 to 1.94 times in 1999. It is not surprising that in a time period marked by the introduction of accountability policies that principals would feel that they lose authority relative to states. Table 2 indicates that principals feel that districts' influence relative to states dropped as well, decreasing from 1.68 in 1993 to 1.49 in 1999. Table 3 presents average principal authority by state for both 1993 and 1999.

While all our measures of local control capture some aspect of the underlying flexibility of schools and school districts, they are by no means the same measure. There is very little overlap between states with high levels of principal control and either low percentages of categorical grants or citizen control over voting. For example, the correlation between percentage of categorical grants and principal control is 0.16 and 0.12 for control over spending and hiring, respectively. Half of the states with no voting provision for revenue raising have low levels of control over allocation as measured by a high percentage of categorical grants, but the others do not. To ensure that we are not picking up effects due to high concentrations of poverty within a state, we also check correlations and interactions between the percentage of that state population in poverty in 1993 and our local control variables. We find no evidence of

8. We use only the public and public charter school surveys found at Schools and Staffing Survey Web site, via the National Center for Education Statistics Web site, at <http://nces.ed.gov/surveys/SASS> (accessed July 2005).

Table 4. 1993 Local Control Measures by Accountability Strength

| | Low Accountability | Medium Accountability | High Accountability |
|---------------------------------------|-------------------------------|----------------------------------|--------------------------------|
| Mean per pupil revenues | \$5,703 | \$5,473 | \$5,877 |
| Mean per pupil expenditures | \$5,133 | \$4,880 | \$5,259 |
| Mean per pupil instructional expenses | \$2,302 | \$2,197 | \$2,360 |
| % of states with no vote | 3.85% | 7.14% | 30.00% |
| % of funding from categorical aid | 11.05% | 10.23% | 14.69% |
| State % of state and local funding | 52.55% | 50.12% | 57.00% |
| % of states with tax limits | 0.36 | 0.64 | 0.30 |
| % of states with court decisions | 0.27 | 0.21 | 0.40 |
| Mean principal control | 4.30 | 4.23 | 2.03 |
| Mean principal by state control | 2.03 | 1.86 | 1.81 |
| Mean principal by district control | 1.45 | 1.40 | 1.41 |
| Mean district by state control | 1.55 | 1.44 | 1.38 |

significant correlation between poverty and any of the other variables. In addition, including poverty in our regression analyses has little effect on our results.

All of the models in this study control for the percentage of eighth-grade students in the state that are black or Hispanic in the 1995–96 school year, the yearly growth in the proportion of black or Hispanic students, the state population in 1996 (in hundred thousands), yearly population growth, and per pupil revenues in 1993. These measures were collected and utilized by Carnoy and Loeb (2003). We control for the proportion and growth of minority students in order to account for any possible omitted variables due to the effects of demographic background characteristics on student outcomes. We control for population size because it is a strong predictor of accountability strength. We include a control for per pupil revenue levels as a measure of states' capacity levels.

5. RESULTS

Initial Local Control Characteristics of States That Adopted Stronger Accountability Policies

This section addresses whether the level of local control in 1993 (before the implementation of most accountability policies) was systematically higher or lower in states that implemented stronger accountability policies. For this analysis, we divide states into three groups: the 26 states with an accountability rating of less than two (low); the 14 with an accountability rating of two to four (medium); and the 10 with an accountability rating of four or higher (high). Table 4 shows that there is little initial difference in revenues or expenditures

per pupil across the three accountability groups. Although 1993 per pupil revenue, expenditures per pupil, and instructional expenditures all appear to be slightly greater in high accountability states than in low or medium states, simple F statistics show that these differences are not statistically significant at conventional levels.

There are significant differences across accountability groups in some measures of local control, though this relationship does not hold for other measures. For example, 30 percent of high-accountability states had no local control over revenue raising, compared to only 4 percent of low-accountability and 7 percent of medium-accountability states. Since there are a manageable number of states to look at individually, we can assess the consistency of this finding that states with lower local revenue-raising control, on average, implemented stronger accountability. Not surprisingly, the finding does not hold for all states. Table 3 shows that, in particular, New Jersey had provisions for both board and citizen voting, which places it in the relatively high local control group, but implemented strong accountability.⁹ Florida, Texas, and North Carolina (other states that rank as fives on the accountability strength scale) had just board voting provisions, indicating that they had relatively low levels of local control over school financing.¹⁰

States that implemented strong accountability policies also had a higher proportion of their resources tied up in categorical grants. Approximately 15 percent of overall education funding was in the form of categorical aid for high-accountability states, whereas the states with the weakest accountability policies had 11 percent of overall education funding in the form of categorical aid, and medium-accountability states had 10 percent categorical aid. These differences are not statistically significant across the three accountability classifications, but the medium-accountability and high-accountability states are statistically different in the percentage of funding they receive from categorical grants.

We do not find significant differences in the prevalence of supreme court rulings overturning local school finance systems between accountability groups, though the trend is consistent with greater accountability in states with less local control. Twenty-seven percent of low-accountability states had

-
9. New York also appears to be an outlier having both strong local control and strong accountability; however, this accountability-strength rating is based on New York's 2001 standards and is one of the two state rankings that were revised upon further examination by Margaret Goertz of the University of Pennsylvania. Goertz's revised rankings do not significantly change the overall results but do show that New York fits into the overall pattern, leaving New Jersey as the sole outlier.
 10. Tax limits are another measure of district revenue-raising authority, though we find no relationship between tax limits and accountability. Districts in 36 percent of low-accountability states were subject to tax limitations in 1997. This compares with 64 percent for medium-accountability states and 30 for high-accountability states. These differences are not significantly different than zero.

such rulings, compared to 21 percent of medium-accountability states and 40 percent of high-accountability states. Table 4 also shows similar differences in 1993 principal-assessed control over school operations by accountability level. The only statistically significant difference across accountability groups is that principals in states that implemented stronger accountability perceived themselves as having somewhat less average influence relative to their district, but the trend is evident across all measures of principal control.

States that implemented stronger accountability appear to have had somewhat less local control over revenue raising and resource allocation. However, perhaps due to small sample sizes and imperfect measurement, not all of the differences are statistically significant. When we conduct a similar analysis but control for other state characteristics including the percentage of black and Hispanic students (levels and growth), population (levels and growth), and per pupil revenues, only average principal control is significantly related to accountability strength.¹¹

The Relationship between Accountability Policies and Changes in Local Control

This section evaluates whether accountability reforms were associated with changes in local control from 1993 to 1999. Unfortunately, we can only do this for local control as measured by principal assessments because for most other measures there was very little change in local control over this time period. While weak differences in principal control by accountability strength were evident in the early 1990s, accountability reforms appear to have changed principal autonomy. Table 5 shows systematic changes in control by accountability strength. Strikingly, although not unexpectedly, states gained more control over most aspects of schooling in states with stronger accountability policies, according to the principals surveyed. Less predictably, however, principals in higher-accountability states expressed greater gains in perceived control, especially in the areas of professional development, spending, and teacher hiring. Table 5 shows that principals in stronger accountability states gained significantly more perceived control over spending and hiring than did principals in lower-accountability states, and they gained more influence over hiring relative to the state, though they lose on curriculum, evaluation of teachers, and average authority relative to the state. Districts in stronger accountability states lost perceived authority over all categories (besides evaluation of teachers) relative to the state.

11. Results are available from the authors upon request.

Table 5. Linear Regression of Principal Control Variables on Accountability Index and Control Variables' 93 Value ($n = 50$)

| 1999 | Average | Curriculum | Discipline | Prof. Dev. | Spending | Evaluations | Hiring |
|--|----------------|-------------------|-------------------|-------------------|-----------------|--------------------|---------------|
| 1999 Principal Influence | | | | | | | |
| Accountability | 0.02** | 0.02 | -0.00 | 0.02* | 0.07** | 0.01 | 0.03** |
| 1993 Principal influence | 0.37** | 0.32** | 0.21 | 0.23** | 0.52** | 0.35** | 0.80** |
| Constant | 2.70** | 2.73** | 3.47** | 3.20** | 2.02** | 3.03** | 1.01** |
| 1999 Principal Influence Relative to State | | | | | | | |
| Accountability | -0.03** | -0.05** | -0.03 | -0.03 | -0.02 | -0.06* | 0.05** |
| 1993 Principal by state | 0.56** | 0.28** | 0.57** | 0.36** | 0.63** | 0.71** | 0.70** |
| Constant | 0.76** | 0.90** | 0.82** | 1.15** | 0.98** | 0.62** | 0.89** |
| 1999 Principal Influence Relative to District | | | | | | | |
| Accountability | 0.01 | 0.01 | 0.02 | -0.00 | 0.03 | 0.00 | 0.02 |
| 1993 Principal by district | 0.53** | 0.81** | 0.53** | 0.18 | 0.75** | 0.56** | 0.64** |
| Constant | 0.61** | 0.22 | 0.66** | 0.98** | 0.44* | 0.70** | 0.54** |
| 1999 District Influence Relative to State | | | | | | | |
| Accountability | -0.03** | -0.04** | -0.05** | -0.03* | -0.03 | 0.05** | -0.06** |
| 1993 District by state | 0.71** | 0.41** | 0.64** | 0.50** | 0.82** | 0.74** | 0.62** |
| Constant | 0.38** | 0.71** | 0.52** | 0.79** | 0.40** | 0.46** | 0.83** |

* $p < .10$, ** $p < .05$.

Effects on Student Outcomes

The analyses of the relationship between student outcomes and the interactions of accountability strength and local control are based on five sets of state-level data: joint data consisting of black and white fourth and eighth graders; separate black and white fourth-grade data; and separate black and white eighth-grade data. We cluster the standard errors at the state level when using the “stacked” data set.

Table 6 gives the average effects of accountability and local control on student outcome measures without including the interaction of these factors. The table reports the results for the change in the McLaughlin-adjusted scale scores, controlling for the 1996 score, the percentage of black and Hispanic students in the 1995–96 school year, the average yearly growth in the percentage of black and Hispanic students, the 1996 state population, and the average yearly population growth.¹² Most striking in this table is the finding that the

12. Similar results for the change in the percentage of students passing at the basic level and the change in the percentage of students passing at the proficient level as a

Table 6. Analysis of 1996–2000 Change in Adjusted NAEP Math Scores

| | Stacked Data (n = 129) | White 4th Grade (n = 33) | Black 4th Grade (n = 34) | White 8th Grade (n = 31) | Black 8th Grade (n = 31) |
|-----------------------------------|-----------------------------------|-------------------------------------|-------------------------------------|-------------------------------------|-------------------------------------|
| Accountability | 1.84** (0.67) | -0.16 (0.65) | -0.08 (0.99) | 1.32** (0.57) | 2.97** (0.99) |
| 1996 NAEP math score | -0.05 (0.06) | -0.12 (0.15) | 0.15 (0.17) | 0.00 (0.12) | -0.23 (0.14) |
| % of 8th grade black/Hispanic '96 | -5.28 (5.96) | 8.00 (7.53) | -5.63 (10.02) | 4.28 (5.60) | -27.27** (10.34) |
| Population in 1995 (100,000s) | 0.01 (0.01) | 0.01 (0.02) | 0.03 (0.02) | -0.01 (0.01) | -0.00 (0.02) |
| 1996 per pupil revenue (1,000s) | 0.36 (0.40) | 0.51 (0.81) | -0.81 (0.91) | 0.82 (0.70) | 1.47 (1.01) |
| Yearly growth % black/Hispanic | -45.69 (30.95) | -11.27 (33.48) | -55.29 (47.51) | -15.68 (34.09) | -89.93 (61.84) |
| Yearly population growth | 129.90 (84.61) | 5.14 (96.94) | 203.85 (133.63) | 53.63 (110.23) | 301.16 (187.77) |
| No vote | -0.55 (1.74) | -0.18 (2.18) | 0.12 (3.09) | -2.59 (1.94) | 1.20 (3.22) |
| Court legislation | -1.59 (1.13) | 0.02 (1.79) | 3.29 (2.24) | -0.33 (1.31) | -2.06 (2.27) |
| % of funding—categorical | -16.13 (11.31) | -15.93 (13.23) | -22.58 (18.38) | -27.30** (13.20) | -3.86 (22.85) |
| Average principal control—'93 | 0.80 (3.18) | 3.50 (4.90) | -4.49 (7.13) | 0.16 (4.38) | 6.39 (7.71) |
| White 8th grader (W8) | 2.49 (3.14) | — | — | — | — |
| Accountability × W8 | -0.81 (0.78) | — | — | — | — |
| Black 4th grader (B4) | 1.98 (3.93) | — | — | — | — |
| Accountability × B4 | -1.26 (0.78) | — | — | — | — |
| White 4th grader (W4) | 2.24 (2.55) | — | — | — | — |
| Accountability × W4 | -1.19* (0.71) | — | — | — | — |
| Constant | -1.19 (6.85) | 13.76 (37.50) | 0.79 (38.09) | -4.56 (29.49) | 21.91 (35.23) |
| Adjusted R ² | .22 | -.16 | -.05 | .15 | 0.23 |

* p < .10, ** p < .05. Standard errors in stacked model are clustered at the state level.

positive effect of accountability holds up at significant levels using all three outcome variables in both eighth-grade specifications and the stacked data. This suggests that accountability policy strength is an important predictor of student performance at all points of the distribution curve, even when controlling for the measures of local control, and especially so for students at the basic (lower) and mean levels. The results also show that accountability strength has

function of accountability and the local control variables are available from the authors upon request.

a larger positive effect on black than on white eighth graders on average and at the basic level. We do not find consistent evidence of direct effects of the local control measures on student outcomes, though the analysis is not designed to find these, since all of the local control measures are included simultaneously.

The next series of analyses address the interaction of accountability and local control. Note that the models include all measures in table 6 except for the local control measures; for simplicity, though, only the local control and the accountability interaction coefficients are provided. Table 7 shows that lack of local control over revenue, as measured by increases in the percentage of total education revenue from categorical grants, negatively interacts with accountability strength to affect student performance. This finding is significant in our clustered stacked regressions as well as in our separate analysis of white fourth graders and is consistently negative in the rest of the analyses. In addition, while the linear effect of accountability was evident for eighth graders but not fourth graders, this negative interaction is just as strong for the fourth grade once these interactions are included. This suggests that although students perform better in states with higher accountability strength on average, this positive effect is less strong when districts have less discretion over resource allocation. The stacked estimations predict that a point increase in accountability will increase the percentage of students achieving at the basic level by 4.4 points on average in states with no categorical aid. In states with 10 percent of revenues through categorical aid, the effect of accountability drops to approximately 3.0 points. In states with 20 percent categorical revenues, the effect of a one-unit increase in accountability drops to only 1.5 points. Examining the nonstacked predictions, we see that a point increase in accountability will increase the percentage of white fourth graders achieving at the basic level by 2.6 points in states with no categorical funding, whereas in states with the mean proportion of their overall funding in the form of categorical aid (12 percent), the effect of accountability drops to only one-third of a point.

Table 7 also looks at the effects of an interaction between voting rights and accountability strength. The importance of local control is evident here as well. Voting rights are measured by whether states have no local voting provisions (the weakest level of local control). The estimates show that accountability is less effective in states that do not allow local voting. This effect is particularly strong for white eighth-grade students, but the negative interaction between accountability and local control is evident in the point estimates for all groups. For example, in a state that allows local voting for revenue raising, a level increase in accountability strength is associated with a 2.8 percent increase in the likelihood that a student will pass the NAEP math exam at the basic level, and with a 1.0 percent increase in the likelihood that a student will pass the NAEP math exam at the proficient level. However, when voting is

Table 7. Analysis of 1996–2000 Change in NAEP Math Measures on Accountability and Percent Categorical and Local Citizen Voting

| | Stacked | W4 | B4 | W8 | B8 | Stacked | W4 | B4 | W8 | B8 |
|---|--------------------|--------------------|-------------------|------------------|-------------------|-------------------|------------------|-----------------|-------------------|-------------------|
| McLaughlin Adjusted Scale Scores | | | | | | | | | | |
| Accountability | 3.42** (0.81) | 1.30 (0.60) | 1.24 (1.62) | 2.57** (1.00) | 4.67** (1.72) | 1.70** (0.68) | -0.29 (0.57) | 0.25 (0.89) | 1.22** (0.50) | 2.61** (0.87) |
| % of funding— categorical | 13.89 (16.26) | 14.41 (17.51) | -5.49 (26.89) | -4.69 (20.83) | 48.61 (35.49) | 5.95 (4.24) | 4.29 (5.49) | 5.82 (8.34) | 10.25* (6.02) | 6.46 (10.74) |
| Accountability x % categorical | -13.46** (5.26) | -12.95* (7.10) | -8.87 (11.11) | -10.55 (7.21) | -18.35 (12.72) | -2.18* (1.30) | -1.65 (1.60) | -2.08 (2.43) | -4.12** (1.71) | -1.61 (3.05) |
| R ² | .24 | .05 | -.04 | .23 | 0.30 | .20 | -.09 | -.12 | .23 | .24 |
| Percent Passing at the Basic Level | | | | | | | | | | |
| Accountability | 4.43** (1.17) | 2.60** (1.23) | 3.91 (2.83) | 2.38* (1.19) | 4.52** (2.34) | 2.79** (0.56) | 0.33 (0.62) | 0.69 (1.31) | 1.59** (0.54) | 4.02** (0.79) |
| % of funding— categorical | 24.62 (17.04) | 34.14 (20.70) | 32.48 (48.13) | -1.24 (24.32) | 40.45 (55.60) | 13.39** (4.74) | 10.24 (6.13) | 9.01 (13.14) | 12.27* (6.46) | 22.91** (9.71) |
| Accountability x % categorical | -14.60** (7.27) | -18.80** (8.30) | -28.91 (20.49) | -6.77 (8.45) | -6.10 (18.16) | -4.57** (1.32) | -3.54* (1.77) | -3.80 (4.16) | -4.25** (1.83) | -6.45** (2.92) |
| Adjusted R ² | 0.29 | .16 | .02 | .19 | 0.38 | .35 | .13 | -.05 | .30 | .54 |

Table 7. Continued

| Percent Passing at the Proficient Level | | | | | | | | | | |
|---|--------------------|-------------------|------------------|------------------|------------------|-----------------------------|-----------------|-----------------|------------------|------------------|
| | Stacked | W4 | B4 | W8 | B8 | Stacked | W4 | B4 | W8 | B8 |
| Accountability | 2.19** (0.67) | 1.70 (1.17) | 1.55 (1.04) | 2.71** (0.93) | 0.77 (0.49) | Accountability | -0.14 (0.61) | 0.32 (0.43) | 1.40** (0.48) | 1.37** (0.22) |
| % of funding— categorical | 19.64** (9.54) | 33.09 (19.74) | 23.94 (17.31) | 3.61 (19.53) | -0.38 (11.12) | No vote | 4.71 (6.03) | -0.90 (4.84) | 7.18 (5.86) | 4.34 (2.64) |
| Accountability × % categorical | -10.55** (4.62) | -14.98* (7.87) | -9.76 (7.55) | -10.40 (6.73) | 4.13 (3.95) | Accountability * No vote | -1.93 (1.75) | -0.31 (1.49) | -2.86* (1.67) | -1.18 (0.80) |
| Adjusted R ² | 0.36 | .11 | -.02 | .24 | 0.74 | Adjusted R ² | .04 | -.06 | .18 | .74 |

* p < .10, ** p < .05. Models control for the 1996 test score or percentage passing, 1996 state population, percent of the eighth graders in the state who are black or Hispanic in 1996, yearly growth in proportion of black/Hispanic students, 1996 state population, and 1996 and yearly population growth. Stacked models include race/grade dummies and interactions and use clustering at the state level for standard error estimates.

Table 8. Analysis of 1996–2000 Change in Math Scores on Accountability and Court Legislation

| | Stacked Data | White 4th Grade | Black 4th Grade | White 8th Grade | Black 8th Grade |
|--|-------------------|-----------------|-----------------|------------------|------------------|
| McLaughlin Adjusted Scale Scores | | | | | |
| Accountability | 2.03** (0.62) | 0.03 (-0.11) | 0.83 (0.92) | 1.38** (0.61) | 2.84** (0.93) |
| Court legislation | 2.29 (2.22) | 3.86 (2.91) | 3.03 (4.03) | 3.88 (2.75) | -0.66 (4.17) |
| Accountability × court legislation | -1.67** (0.76) | -1.79 (1.05) | -2.45 (1.56) | -1.79 (1.07) | -0.84 (1.61) |
| R ² | .22 | -.02 | .00 | -.00 | 0.28 |
| Percent Passing at the Basic Level | | | | | |
| Accountability | 3.03** (0.63) | 0.50 (0.72) | 1.29 (1.25) | 1.65** (0.65) | 3.98** (1.05) |
| Court legislation | 2.03 (2.47) | 2.56 (3.28) | 5.28 (7.13) | 2.51 (2.88) | 0.06 (6.19) |
| Accountability × court legislation | -1.42 (1.09) | -1.39 (1.27) | -4.21 (2.61) | -1.21 (1.12) | -0.37 (2.18) |
| R ² | .28 | .02 | .14 | -.13 | 0.35 |
| Percent Passing at the Proficient Level | | | | | |
| Accountability | 1.15** (0.27) | 0.12 (0.67) | 0.58 (0.48) | 1.58** (0.55) | 1.19* (0.25) |
| Court legislation | 1.93 (1.30) | 2.50 (3.03) | 2.23 (2.88) | 3.36 (2.47) | -0.99 (1.47) |
| Accountability × court legislation | -0.95* (0.57) | -1.41 (1.18) | -1.20 (1.06) | -1.47 (0.96) | 0.49 (0.51) |
| R ² | .33 | .03 | -.04 | .08 | 0.71 |

* p < .10, ** p < .05. Models control for the 1996 test score or percentage passing, 1996 state population, percentage of the eighth graders in the state who are black or Hispanic in 1996, yearly growth in proportion of black/Hispanic students, 1996 state population, and 1996 and yearly population growth. Stacked models include race/grade dummies and interactions and use clustering at the state level for standard error estimates.

not authorized, accountability has a negative effect. A one-point increase in accountability strength results in a 1.8 percentage point *decrease* in the number of students who pass the NAEP exam at the basic level and in a 0.7 percentage point decrease in the percentage achieving proficiency. These findings are consistent with our hypothesis that local control over resources is an important factor in allowing schools and districts to respond to the incentives created by accountability reforms.¹³

Table 8 looks at the effects of accountability and whether or not states' supreme courts have overturned school finance systems that heavily rely on

13. We ran similar analyses for tax limitation but did not find any effect of the interaction of accountability and tax limits on student outcomes. The results are available upon request.

local financing. We see a consistently negative interaction between accountability strength and states' court legislation status, although this is only significant in our stacked regressions in the "adjusted level" and "percentage passing at the proficient level" analyses. (This effect is almost significant at the 10 percent level for both white scale score specifications.) This tells us that in a state with the mean level of accountability strength (accountability equals 2.12), students, on average, are likely to score 4.3 points higher on the NAEP math exam. However, in states where courts have overturned local school financing, students are only likely to perform 3.05 points higher on the NAEP math exam—1.25 points lower than students would in states without such court legislation. This indicates that accountability policies may be less effective as school funding becomes more centralized, and perhaps more so for average- and high-achieving students than for low-performing students.

Table 9 describes the effects of accountability and perceived principal control measures. The results continue to show more positive effects of accountability in states with greater local control. We see a consistently positive effect of the interaction between principal control and accountability strength. Specifically, stronger accountability reforms are associated with higher proportions of students scoring at or above the basic and proficient levels in states in which principals have greater control over spending. This finding is consistent with the theory that principals need control over resource allocation in order to effectively respond to accountability.

In summary, these analyses provide evidence that the estimates of accountability are robust to the inclusion of local control measures. Most of our measures of local control do not show a statistically significant direct effect on student outcomes, although this could easily be due to a combination of a lack of power, colinearity, and measurement error. There is, however, substantial evidence that accountability policies are more effective when there is more local control. This comes from the negative interactions of accountability with increases in the percentage of education funding from categorical grants, state supreme court legislation, and no voting provisions, as well as from the positive interaction of accountability and principal influence measures.¹⁴

14. One concern with the analysis is that because the sample size is so small, a single state might be driving the results. In order to check for this, we reran the analyses eliminating a different state each time. In general, the results were not affected by this. The elimination of individual states had no substantive effect on our estimate of the interaction between accountability and either percentage categorical or principal control. The elimination of Virginia from the estimate of the interaction between accountability and no local voting authority causes the point estimate to lose significance at all conventional levels, though it is still negative. We also use the `dfbeta` command in Stata to determine whether our results are sensitive to a single or a small number of observations. We find no evidence that this is the case.

Table 9. Analysis of 1996–2000 Change in Percentage Passing at the Basic and Proficient Levels on Accountability and Measures of Principal Authority

| | | Percent Passing at the Basic Level | | | | Percent Passing at the Proficient Level | | | | |
|--------------------------------------|-------------------|------------------------------------|-------------------|-------------------|-------------------|---|------------------|------------------|-------------------|-----------------|
| | Stacked | W4 | B4 | W8 | B8 | Stacked | W4 | B4 | W8 | B8 |
| Accountability | -11.89 (11.54) | -7.71 (15.47) | -20.30 (29.01) | -9.58 (13.77) | -12.46 (27.92) | Accountability | -9.21 (6.69) | -8.02 (11.50) | -9.58 (12.10) | 1.04 (6.21) |
| Average principal control '93 | -7.82 (3.62) | -4.05 (9.70) | -6.00 (18.24) | -7.24 (9.02) | -4.07 (26.65) | Average principal control '93 | -5.52 (4.30) | -1.44 (7.25) | -6.78 (7.87) | 2.35 (4.04) |
| Accountability × avg. prin. control | 3.53 (2.86) | 1.89 (3.71) | 5.12 (6.99) | 2.63 (3.31) | 4.05 (6.63) | Accountability × avg. prin. control | 2.46 (1.62) | 2.06 (2.76) | 2.60 (2.91) | 0.08 (1.50) |
| Adjusted R ² | 0.27 | -.02 | -.07 | .11 | 0.38 | Adjusted R ² | .33 | -.05 | .02 | .71 |
| Accountability | -6.13 (4.17) | -5.35 (7.87) | -17.68 (14.01) | -2.63 (6.98) | -2.95 (12.19) | Accountability | -4.35 (2.76) | -5.18 (4.94) | -4.69 (6.10) | 1.89 (2.92) |
| Principal control—hire '93 | -4.20 (3.41) | -3.02 (4.90) | -6.94 (9.05) | -2.29 (4.55) | -1.49 (10.30) | Principal control—hire '93 | -2.12 (1.80) | -0.07 (3.15) | -3.02 (3.93) | 1.34 (1.86) |
| Accountability × prin. control—hire | 2.06** (1.02) | 1.26 (1.80) | 4.32 (3.22) | 0.92 (1.60) | 1.65 (2.76) | Accountability × prin. control—hire | 1.24* (0.64) | 1.36 (1.13) | 1.37 (1.40) | -0.12 (0.67) |
| Adjusted R ² | 0.28 | -.01 | .02 | .10 | 0.38 | Adjusted R ² | .33 | .15 | .03 | .70 |
| Accountability | -9.23* (5.51) | -10.32 (9.05) | -10.98 (16.46) | -12.46* (7.09) | -13.77 (19.11) | Accountability | -8.99* (4.18) | -7.57 (6.03) | -12.15* (6.08) | 2.06 (4.01) |
| Principal control—spend '93 | -8.48* (4.68) | -6.01 (6.02) | -9.05 (11.32) | -7.99* (4.67) | -22.97 (19.71) | Principal control—spend '93 | -6.07* (2.88) | -5.33 (4.24) | -6.26 (3.99) | 0.52 (3.18) |
| Accountability × prin. control—spend | 3.15 (2.01) | 2.74 (2.36) | 2.98 (4.30) | 3.60* (1.84) | 4.45 (4.91) | Accountability × prin. control—spend | 2.61 (1.09) | 2.04 (1.56) | 3.49** (1.58) | -0.20 (1.04) |
| Adjusted R ² | 0.18 | .02 | -.08 | .22 | 0.42 | Adjusted R ² | .37 | -.03 | .20 | .68 |

* p < .10, ** p < .05. Models control for the 1996 test score or percentage passing, 1996 state population, percentage of the eighth graders in the state who are black or Hispanic in 1996, yearly growth in proportion of black/Hispanic students, 1996 state population, and 1996 and yearly population growth. Stacked models include race/grade dummies and interactions and standard errors are clustered at the state level.

6. CONCLUSIONS AND POLICY IMPLICATIONS

Our intent with this study is to begin assessment of the relationship between accountability and local control. We first find evidence that stronger accountability policies were implemented in states with weaker local control, as measured by local provisions for revenue raising and allocation. This finding seems intuitive but is relevant to the policy discussion surrounding accountability programs. If the ability for districts and schools to respond to accountability provisions depends on the amount of flexibility and control they have over their school- and district-based activities, then the states that are least able to respond to accountability programs are exactly those in which accountability policies are being instituted most strongly.

Next we look at changes in local control associated with the implementation of accountability reforms. We find little difference in principal control over school operations in the states that implemented different strength accountability policies. However, the implementation of stronger accountability corresponds to increases in principals' perceived control over school operations, through drops in principals' assessment of their power and their district's power relative to the state. This points to the possibility that principals do not feel disenfranchised by the implementation of strong accountability policies, as one might have guessed, but instead seemingly feel *more* in control with these policies.

The main analyses in this article address the differential effect of accountability reforms on student math achievement in states that had greater and lesser local control. The results are directly relevant to current policy discussions on accountability. First, we provide supporting evidence to Carnoy and Loeb's (2003) findings that accountability policies led to increases in student performance on NAEP math exams. Accountability policies may be tools for educational policy makers to use in the quest to improve student outcomes. However, our own and other research suggests that accountability policies do not exist in a vacuum: many factors constrain the implementation and effectiveness of accountability programs. Our analyses, while imperfect, point to the importance of local control over revenue raising and allocation in the successful implementation of accountability programs. We find that accountability policies were substantially more effective in states with stronger local control. It appears that without some local control, even well-thought-out accountability policies will be less effective, and sometimes ineffective and harmful. This problem is further augmented by the recent trend away from local and toward state control over education funding. As education finance continues to be centralized at the state level, citizens and districts lose control over revenue raising and allocation, potentially impeding the positive effects of state-implemented accountability policies.

We appreciate the helpful guidance from David Figlio and Jim Guthrie, as well as two anonymous reviewers. This work was partially funded by the American Education Research Association Dissertation Grant. Any remaining errors are the authors' own.

REFERENCES

- Amrein, Audrey L., and David C. Berliner. 2003. The effects of high-stakes testing on student motivation and learning. *Educational Leadership* 60:32–38.
- Card, David, and Alan B. Krueger. 1992. Does school quality matter? Returns to education and the characteristics of public schools in the United States. *Journal of Political Economy* 100:1–40.
- Card, David, and A. Abigail Payne. 2002. School finance reform, the distribution of school spending, and the distribution of SAT scores. *Journal of Public Economics* 83:49–82.
- Carnoy, Martin, and Susanna Loeb. 2003. Does external accountability affect student outcomes? A cross-state analysis. *Education Evaluation and Policy Analysis* 24 (4): 305–31.
- Coleman, James S. 1966. *Equality of educational opportunity*. Washington, DC: U.S. Government Printing Office.
- Cullen, Julie B., and Susanna Loeb. 2004. School finance reform in Michigan: Evaluating proposal A. In *Helping children left behind: State aid and the pursuit of educational equity*, edited by John Yinger, pp. 215–50. Cambridge, MA: MIT Press.
- Dee, Thomas S. 2002. *Standards and student outcomes: Lessons from the “first wave” of education reform*. Cambridge, MA: Kennedy School of Government.
- Downes, Thomas A. 1992. Evaluating the impact of school finance reform on the provision of public education: The California case. *National Tax Journal* 45 (4): 405–19.
- Duncombe, William, and Jocelyn M. Johnston. 2002. Is something better than nothing? An assessment of school finance reform in Kansas. Working paper, Syracuse University.
- Figlio, David N. 1997. Did the “tax revolt” reduce school performance? *Journal of Public Economics* (65):245–69.
- Figlio, David N. 2002. Funding and accountability: Some conceptual and technical issues in state aid reform. Paper presented at the Conference on State Aid to Education, Syracuse University, April.
- Figlio, David N., and Lawrence S. Getzler. 2002. Accountability, ability and disability: Gaming the system. NBER Working Paper No. 9307.
- Figlio, David N., and Kim S. Reuben. 2001. Tax limits and the qualifications of new teachers. *Journal of Public Economics* 80: 49–71.
- Hannaway, Jane. 1993. Political pressure and decentralization in institutional organizations: The case of school districts. *Sociology of Education* 66 (3): 147–63.
- Hanushek, Eric A. 1986. The economics of schooling: Production and efficiency in public schools. *Journal of Economic Literature* 24 (3): 1141–77.

Hanushek, Eric A. 1997. Assessing the effects of school resources on student performance: An update. *Educational Evaluation and Policy Analysis* 19 (2): 141–64.

Hanushek, Eric A., and Macke E. Raymond. 2002. Lessons and limits of state accountability systems. Paper presented at “Taking Account of Accountability: Assessing Policy and Politics” Conference, Harvard University, June.

Hanushek, Eric A., and Macke E. Raymond. 2004. Does school accountability lead to improved student performance? NBER Working Paper No. 10591.

Imazeki, Jennifer, and Andrew A. Reschovsky. 2004. School finance reform in Texas: A never ending story? In *Helping children left behind: State aid and the pursuit of educational equity*, edited by John Yinger, pp. 251–82. Cambridge, MA: MIT Press.

Jacob, Brian A. 2002. Test-based accountability and student achievement gains: Theory and evidence. Working paper, Harvard University.

Jacob, Brian A., and Steven D. Levitt. 2003. Rotten apples: An investigation of the prevalence and predictors of teacher cheating. *Quarterly Journal of Economics* 118 (3): 843–77.

Loeb, Susanna, and Marianne E. Page. 2000. Examining the link between teacher wages and student outcome: The importance of alternative labor market opportunities and non-pecuniary variation. *Review of Economics and Statistics* 82 (3): 393–408.

McLaughlin, Don. 2001. Exclusions and accommodations affect state NAEP gain statistics: Mathematics, 1996–2000. American Institutes for Research (mimeo).

Murray, Sheila E., William M. Evans, and Robert M. Schwab. 1998. Education finance reform and the distribution of resources. *American Economic Review* 88:789–812.

National Center for Education Statistics (NCES). 2001. Public school finance programs of the United States and Canada: 1998–99. Available at <http://nces.ed.gov/edfin/state.finance/StateFinancing.asp>. Accessed 11 July 2006.

Sonstelie, Jon, Eric Brunner, and Kenneth Ardon. 2000. *For better or for worse? School finance reform in California*. San Francisco: Public Policy Institute of California.