sanctions and rewards for students, teachers, schools, and even entire districts. Other systems, sometimes called report card systems, simply report test results to the public, without attaching explicit consequences for anyone.

If hard accountability systems establish fixed rewards and sanctions, report card systems rely on external forces to generate performance incentives. Report card systems depend on either of two mechanisms to encourage improvements in student test scores: the market or the ballot box. The market works because school quality is capitalized in property values, which in turn influence school budgets. Because land prices reflect the quality of local schools and because school budgets are tied to local property taxes, districts have (admittedly weak) incentives to boost student achievement, especially when test results are widely distributed. The ballot box, on the other hand, establishes a more immediate link between school board members’ electoral fortunes and student performance. Voters may reward board members in successful districts with reelection, while they send members in faltering districts packing, or so the theory goes. Although a good deal of empirical evidence confirms that property values do, in fact, fluctuate with school quality, almost nothing is known about whether voters actually hold school board members accountable for student learning. Indeed, we are aware of no existing research that tests the basic proposition that school board members’ electoral fortunes are tied to student performance.

This chapter examines whether average voters hold school board members accountable for the performance of their schools. Specifically, it assesses whether voters punish or reward incumbent school board members on the basis of changes in student learning (as measured by test scores) in local and district schools. It also scrutinizes candidate behaviors, assessing the impact of student learning trends on incumbents’ decisions to seek reelection and potential challengers’ decisions to contest them. The findings from the 2000 elections in South Carolina are striking. From the initial decision to run to the final vote tallies, we observed robust relationships between student learning and incumbents’ electoral fortunes. During the 2002 election, however, when turnout dropped by roughly half of that observed in 2000, we found scant evidence that voters hold members accountable for changes in test scores. Whether voters hold school board members accountable for recent changes in test scores, we suggest, depends critically on who shows up on election day.

This chapter first reviews the relevant literatures on accountability and voting behavior and identifies forces that contribute to and detract from accountability in school board elections. It then introduces new data from school board elections in South Carolina and tests whether changes in student test scores systematically affected the probability that incumbent school board members would seek reelection and the probability that they would face competitors if they did.
Literature Review

This chapter draws from two bodies of literature, one on accountability systems, which resides within the disciplines of education and economics, and the other on retrospective voting, which is firmly enmeshed in political science. We review each in turn.

Accountability

Although performance-based accountability in education remains in its infancy, a significant scholarly literature on the subject is emerging. Considerable attention has been devoted to the optimal design of accountability systems, and questions regarding top billing include how to appropriately measure student achievement, whom to hold accountable, and which rewards and sanctions are most effective. Other authors have provided case studies of a particular state’s or district’s experience with accountability reforms: for example, Florida (Goldhaber and Hannaway 2004), California (Bets and Dunnemberg 2003), and Chicago (Bryk 2003; Jacobs 2003). And, of course, a flood of ink has been spilled over the anticipated effects of NCLB (Peterson and West 2003).

Although the empirical evidence on school accountability systems is thin, the initial findings are encouraging. Hanushek and Raymond, for example, found that students in states with either hard accountability or report card systems registered larger gains on the National Assessment of Educational Progress than did students in states lacking accountability systems. It is important to note that although students in hard accountability states outperformed those in report card states, the difference between the two systems was not statistically significant. Thus, early results suggest that accountability, of either variety, pushes school test scores in a positive direction.

That hard accountability and report card systems produce comparable results raises questions about whether explicit sanctions and rewards are necessary to produce improvements in student learning. Hanushek and Rivkin found that the mere disclosure of student achievement on the Texas Assessment of Academic Skills generated the competition needed to produce gains in average scores. In his assessments of accountability in California, Florida, and Texas, Martin Carnoy argues that even without the imposition of sanctions, public designation of failing schools led to substantial subsequent improvements, through a mechanism

1. For important recent contributions, see Ladd (1996), Eves and Walberg (2002), and Peterson and West (2003).
2. Finn (2003); Isen and Eves (2002).
3. Hanushek and Raymond (2003). The performance difference between either a report card system or a hard accountability system and no accountability system was not statistically significant.
4. Hanushek and Rivkin (2003). In their paper, the authors analyzed the Texas accountability system before direct consequences were linked to performance.

Retrospective Voting

For decades, scholars have examined the ways in which incumbents’ electoral fortunes rise and fall with their constituents’ material well-being. By placing a minimal informational burden on voters while also recognizing incumbents’ attempts to advance a record of success, retrospective voting poses a simple and powerful voting heuristic: voters support incumbents whose outcomes are marked by improvements in the state of the world, and they oppose those who have overseen declines. The electorate thereby assumes the “role [of] an appraiser of past events, past performance, and past actions.”

7. This is a special case of the well-known Tiebout model, which suggests that residential choice leads to efficiency-enhancing competition among local governments.
Over the past thirty years, scholars have amassed a voluminous body of
empirical research on retrospective voting. Considerable evidence supports the
naive hypothesis—put simply, that citizens vote strictly in accordance with
recent economic fluctuations. Arguments persist, however, over the salience of
different economic indicators in voting decisions. Should incumbents worry
most about unemployment figures, inflation, earnings, or the gross domestic
product? And to the extent that trade-offs among them are unavoidable, should
incumbents seek some optimal balance? Lewis-Beck and Stegmuier conclude
that “the savvy modeler, given the choice of only one predictor [of elections],
would do well to select an economic variable. Which one? The answer varies
from country to country. It could be unemployment, inflation, or growth.”

Having selected a feature of the domestic economy, it remains unclear
whether voters reflect on their own condition or society’s more generally when
evaluating an incumbent. In a growing economy, will citizens who have become
unemployed (or who have had their wages cut or have witnessed their stock
portfolio decline in value) nonetheless support the incumbent? Or will citizens
who meet personal misfortune instinctively punish the incumbent? To the
extent that the issues on the matter, most scholars suggest that collective
(or sociotropic) considerations dominate pocketbook (or egotropic) concerns.
The issue, however, remains far from settled, as other scholars have detected evi-
dence of pocketbook voting.

Finally, arguments endure about whether the electoral fortunes of incum-
bents in different branches of government facing their constituencies in differ-
ent elections depend on economic improvements in equal measure. While much
of the empirical literature on retrospective voting began by focusing on Con-
gress, supportive evidence appears especially robust in presidential elections,
where turnout is much higher. By extension, scholars have tended to observe
more robust relations between economic developments and voting behavior
during on-year elections than during off-year elections.

Although the retrospective voting literature offers a sophisticated theoretical
apparatus and a set of testable hypotheses, it has yet to examine local elections,

12. Lewis-Beck and Stegmuier (2000, p. 211). A handful of retrospective voting studies exam-
ined noneconomic indicators; see, for example, Fiorina, Achen, and Pope (2003).
13. Kiewiet (1983); Kinder, Adams, and Geerke (1989); Mackie (1992); Alvarez and Nagler
15. Tofte (1978); Weatherford (1978); Jacobson and Kessel (1983); Boes (1986); Alstott and
Rosemael (1989); Edelson (1990).
16. Fiorina (1981); Kinder and Kessel (1981). The 2000 presidential election, for which the fore-
casting models regularly predicted a strong Gore victory, appears exceptional in this regard. See the
symposium in the March 2001 issue of the Journal of PS.
Board turnover may reflect members' disinterest and poor working conditions more than a weak致力electorate's lambasting incompeinces any time student learning declines. Ultimately, willingness to serve may be the only real prerequisite for joining a school board. If true, then there really is no politics to speak of in school board elections. Elections are more formalities, whereby willing servants declare their candidacy and, more often than not, a grateful electorate elects them into office.

Finally and perhaps most obviously, student learning (at least that part of student learning captured by test scores) may not be at the forefront of citizens' minds when they enter the voting booth and choose from a slate of school board candidates. Voters instead may be preoccupied with safety issues, the football team's record, the convenience of the busing system, or the attractiveness of the buildings. In principle, the function of schools is to promote and enhance student learning. Voters, however, may not hold school boards strictly to that charge. And if not, then school board elections may be just as complex as the national and state races that have been the focus of the recent retrospective voting literature.

Ultimately, these are empirical questions. If, in fact, board elections are competitive and voters hold members responsible for the academic performance of schools, then changes in student test scores should correlate positively with incumbents' reelection. On the other hand, if unions rig election outcomes, if most races are noncompetitive, or if voters do not pay attention to student learning, then null findings should emerge. We now investigate these possibilities, combining school- and district-level trends on standardized exams with school board election returns in South Carolina.

South Carolina

We analyzed data from South Carolina, which is, to our knowledge, the only state to collect precinct-level election data for local races. Precinct-level data for all local elections are available from the South Carolina Election Commission, while in all other states local election data must be collected from individual counties. Furthermore, South Carolina recently instituted a statewide standardized student achievement test, making school-level data publicly available. This combination of readily available electoral and achievement data make South
Carolina an ideal, and (temporarily) unique, testing ground for theories of democratic accountability in education.

**Election and Student Achievement Data**

South Carolina is divided into eighty-five school districts. More than 90 percent of school boards have between five and nine members, while the largest board (in Beaufort) has eleven. Of these districts, thirty-nine held school board elections in 2000. We collected precinct-level election returns for all school board races and then computed the vote share, by precinct, for each incumbent running in a competitive election (more on this below). That our units of observation are unique incumbent-by-precinct combinations. Because each incumbent runs in more than one precinct, and because each precinct may host more than one school board race, we have multiple observations of most incumbents and precincts. Specifically, in 2000 we had sixty-seven incumbents running in 306 precincts, for a total of 960 observations on incumbents’ vote share.

Student achievement data were obtained through the South Carolina Department of Education. Since 1999, South Carolina has administered the Palmetto Achievement Challenge Tests (PACT) to students in grades 3 through 8. These tests, based on the South Carolina Curriculum Standards, are given in English and math. We averaged the English and math scores to arrive at a composite score for each school and then computed district-level and precinct-level average composite scores. The precinct-level percentile scores indicate the performance of the schools near the polling place and hence those schools most likely to be attended by a voter’s children or those of a neighbor; district-level scores indicate the overall performance of all schools in the district. To test claims about egotropic and isocentric voting, we estimated models with both precinct-level and district-level scores.

**Model Specifications and Results**

Our analysis focuses first on the 2000 South Carolina school board elections.

In that year, sixty-seven incumbents from thirty-seven school boards ran for reelection in competitive races. Of the sixty-seven incumbents, fifty were reelected, and the median vote share for all incumbents was 58 percent. We estimated simple least squares regressions that posited incumbent vote share as a function of test score trends and some basic controls, each of which is explained below. Because observations for the same incumbent across precincts and for multiple incumbents in the same school district are not independent, we allowed for clustering of the standard errors by school district.

Table 7-1 presents the results. The first column shows the simplest model, with only precinct-level and district-level test score changes on the right side of the equation and 2000 vote shares on the left. As discussed above, precinct-level test scores represent the performance of voters’ nearby schools, while district-level scores reflect the performance of all schools in the district. Only precinct-level test score change is significant in this model, with the expected positive coefficient indicating that incumbents won more votes where test scores showed improvements. That district-level scores were not significant suggests that voters were behaving egotropically—or perhaps more accurately, “egocentrically”—focusing on the performance of their own local schools more than on that of the broader district.

The second column of table 7-1 shows the results from models that control for lagged incumbent vote share (that is, the 2000 incumbent’s vote returns in 2000...
the same precinct in 1996), test score levels, and measures of school expenditures. The lagged vote share should capture unobserved aspects of an incumbent’s profile, such as name recognition, experience, endorsements, and fundraising capacity. Not surprisingly, its coefficient is highly significant and positive, indicating that candidates who did well in 1996 also garnered more votes in 2000. Levels of test scores, on the other hand, were nowhere near significant, consistent with the prediction from the retrospective voting literature that rational citizens will base their assessment of incumbents on changes during their tenure rather than the absolute level of performance. Finally, to account for the possibility that races are more competitive in higher-spending districts and that votes may punish board members for marginal increases in their taxes, we controlled for levels and changes in per-pupil expenditures. Neither of these variables, however, has significant effects on incumbent vote shares.26

Importantly, test score changes are robust to the inclusion of these additional variables. The coefficients reported in column 2, our preferred specification, indicate that a movement from the sample’s 25th to 75th percentile of test score change—that is, moving from a loss of 4 percentile points to a gain of 3.8 percentile points between 1999 and 2000—is associated with an increase of 3 percentage points in an incumbent’s vote share. Similarly, a movement from the sample’s 10th to 90th percentile of test score change is associated with an increase of 4.8 percentage points in an incumbent’s vote share. With an average incumbent vote share of 58 percent, these estimates suggest that a major swing in test scores can erode as much as two-thirds of an incumbent’s margin of victory in a two-way race.27

The results reported in table 7-1 reflect the experience of incumbents running in competitive elections. Many incumbents, however, either did not run for reelection or ran unopposed. Specifically, of the 157 incumbent board members in thirty-nine school districts who were up for election in 2000, 112 sought reelection and forty-five of them did not face a challenger. As a result, the sixty-seven incumbents reflected in the results of table 7-1 represent less than half of the incumbents whose seats were in play in 2000. Because these candidates presumably were not randomly selected into competitive elections, test scores may have influenced electoral outcomes beyond the observed vote shares. Indeed, if board members anticipate citizens’ voting behavior, then incumbents in districts with declining test scores should be less likely to seek reelection and more likely to face competition when they do run. If either of these effects is present, then

26. Models that account for test score changes and levels as a function of dollars spent on student resources are virtually identical to those presented below.

27. Virtually identical results for all coefficients are obtained when estimating the probability that an incumbent won a majority of the votes in each precinct in a competitive race, rather than the margin by which an incumbent wins. In addition, weighting the observations by the number of votes cast in the precinct yields nearly identical results.

<table>
<thead>
<tr>
<th>Table 7-2. Seeking Redefinition and Facing Competitors, 2000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Change in total score, 1999–2000 (district)</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Average percentile score, 2000 (district)</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Incumbent vote share, 1996</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Per-pupil expenditures, 2000</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Percent change in per-pupil expenditures, 1999-2000</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Dummy - 1 if position pays no salary</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Constant</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

Summary statistic:

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>152</td>
</tr>
<tr>
<td>Pseudo R²</td>
<td>0.09</td>
</tr>
</tbody>
</table>

Source: Authors’ calculations.

a. * p < .50, two-tailed test; ** p < .05.

the results shown in table 7-1 underestimate the effect of test score change on incumbents’ electoral prospects. We therefore ran two logistic regressions, first to estimate the effect of test scores on the incumbents’ decision to run and then to estimate the probability that those who did run would face competition. In contrast to the vote share models, we do not have precinct-level observations here because when candidates run, they run in all precincts in the district. Thus our unit of observation is the incumbent, and we have only one observation per incumbent. For this reason, we used only district-level rather than precinct-level test scores and lagged vote share on the right side. Again we allowed for clustering of standard errors within school districts.

Column 1 of table 7-2 presents the results of the first logistic regression. As in table 7-1, the model controls for test score levels, incumbents’ lagged vote share, and measures of per-pupil expenditures. In addition, we included a dummy variable indicating whether board members received remuneration for their service. In our sample, approximately 20 percent of officials
received no compensation, while the remainder received a salary, per diem payments, or reimbursement for expenses.\textsuperscript{28} As shown in table 7-2, incumbents were significantly less likely to seek reelection when they were not compensated for their service. In addition, incumbents were less likely to seek reelection in areas where per-pupil expenditures had increased, perhaps because they anticipated a tax revolt at the polls. Neither test score levels nor lagged vote shares were significant, indicating that incumbents in higher-performing districts and incumbents who did especially well in past elections were no more likely to run for office. Most relevant, however, effects of test score changes continue to attain statistical significance and remain in the expected direction: incumbents appear disinclined to seek reelection when their district's test scores drop. This result may indicate that incumbents bow out in anticipation of voter reprimand for poor performance or that serving in a declining district is less rewarding for board members. The point estimates in column 1 indicate that a movement from the 75th to the 25th percentile in the sample's test score change is associated with a drop of 13 percentage points in the probability that the incumbent would seek reelection (from 84 to 71 percent, holding other variables at their median). A movement from the 90th to the 10th percentile is associated with a more than 30 percentage point drop in the probability of seeking reelection (from 90 to 59 percent).\textsuperscript{29}

If declining test scores discourage incumbents from seeking reelection, the retrospective voting literature suggests just the opposite in the case of competition: falling test scores should bring out more challengers. Of the 112 incumbents who sought reelection in 2000, forty-five ran unopposed. To test this hypothesis, we ran a logistic regression where the dependent variable is coded 1 if the incumbent faced at least one challenger and zero if he or she ran unopposed. In addition to the three variables introduced in column 1, we added the number of registered voters in the district in the expectation that because larger districts have a bigger pool of potential candidates, they should be more likely to host contested elections.\textsuperscript{30}

Test score changes, once again, are highly significant, and this time they are negatively associated with the probability of competition. The point estimates from column 2 suggest that a movement from the 75th to the 25th percentile in test score change is associated with an 18 percentage point increase in the probability of facing a challenger (from 44 to 62 percent, holding other variables at their median). A movement from the 90th to the 10th percentile in test score change is associated with a whopping 42 percentage point increase in the probability of facing opposition (from 32 to 74 percent). These results suggest that incumbents running in districts where test scores have taken a nose dive are almost certain to face a challenger.\textsuperscript{31} All of the control variables, meanwhile, appear insignificant with the exception of whether a position is paid. Curiously, challengers were more willing to take on an incumbent board member when victory did not ensure some kind of financial remuneration.

In summary, we identified three major effects of test score change in South Carolina school board elections.\textsuperscript{32} First, incumbents were significantly less likely to seek reelection when test scores declined on their watch, because of either anticipated voter retaliation or frustration with serving in a faltering district. Second, those incumbents who did run were significantly more likely to be challenged at the polls if they presided over a test score drop. Third, those incumbents who ran in competitive elections received a significantly lower share of the vote where test scores had fallen. Because test scores influenced the chances that an incumbent ran in a competitive election—essentially deterring the worst performers from running and rewarding the best with an uncontested seat—the results reported in table 7-1 should be considered the lower bound estimates of the effects of test scores on incumbents' electoral fortunes.\textsuperscript{33}

31. We extended these analyses by examining the possibility that declines in test scores have greater impact than do test score gains. Specifically, we use regression splines on test score change, setting a single knot at zero. Asymptotic effects are apparent in the competitive model. While positive improvements in test scores appear unrelated to the probability that an incumbent faces a challenger, test score declines have a substantial impact on the willingness of challengers to enter the race. This finding suggests that the coefficients shown in table 7-2 underestimate the effect of falling test scores on the probability of an incumbent facing competition. The point estimates for negative test scores, in fact, is about three times larger under the spline specification. We did not, however, find evidence of asymmetric effects in models of an incumbent's decision to seek reelection or in models of vote share.

32. In further analyses (not shown) we address the concern that test score change may be a proxy for other conditions that influence voters' assessments of incumbents. If test score change is correlated with broader shifts in the overall health of the local community, then the effects we observe in the preceding analysis may be spurious, merely indicating that all incumbents faced greater electoral pressures in deteriorating communities. To test for this possibility, we modeled the effects of test scores on county council elections. If test scores provide unique information about the performance of the school board, then we expect them to have negligible effects on county council races. Remarkably, we found no systematic effects of test scores on county council elections.

33. Several considerations dissuade us from attempting to estimate a Heckman-type selection model in this case. First, we would require at least one identifying variable that strongly affirms the chances of observing an incumbent in a competitive election but that is unrelated to incumbents
The 2002 School Board Elections

Here, we replicate the main school board models for the 2002 elections. As previously noted, our theoretical expectations for this midterm election were less clear. Roughly 53 percent of registered voters turned out in 2000; in 2002, meanwhile, just 26 percent did. Unfortunately, available demographic data on the voting population—race, age, and gender—do not provide much of a basis for probing the differences between the two electorates in any depth. It seems fair to assume, however, that voters in 2002 were significantly more educated than their counterparts of two years prior. If Zaller (2004) is correct that retrospective voting occurs principally among low-information voters while high-information voters pay more attention to candidates' policy positions, then test score changes may not impact school board members' electoral prospects during the off-year election.

Consistent with Zaller's argument, the results from 2002 differed markedly from those observed in 2000. As table 7-3 shows, changes in test scores did not affect the probability that an incumbent would seek reelection, the probability that he or she faced competition, or his or her final vote share. These null findings, what is more, do not appear to be a statistical artifact. In models not presented, we added administrative data from teacher, parent, and student ratings of local schools: we experimented with two- and three-year changes in test scores, rather than one-year changes; we looked at changes in the percentage of students who received failing scores on the PACT and we replicated the spline models for 2002 in an attempt to uncover asymptotic effects. None of these alternative approaches produced any evidence of retrospective voting in the 2002 elections.

Differences between the 2000 and 2002 models also were observed with respect to the control variables. Lagged incumbent vote share, for instance, is only marginally significant in the 2002 vote share model, though it stood out as the most highly significant variable in the 2000 model. This finding, however, is

<table>
<thead>
<tr>
<th>Variable</th>
<th>Did incumbent run for reelection? (Logit)</th>
<th>Incumbent vote share, 2002 (linear)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change in test score, 2001-02 (district)</td>
<td>0.001 (0.010)</td>
<td>-0.004 (0.007)</td>
</tr>
<tr>
<td>Change in test score, 2001-02 (precinct)</td>
<td>-0.004 (0.003)</td>
<td>-0.001 (0.001)</td>
</tr>
<tr>
<td>Percent change in per-pupil expenditures, 2001-02</td>
<td>0.595*** (0.246)</td>
<td>-0.264 (0.218)</td>
</tr>
<tr>
<td>Pre-pupil expenditures, 2002</td>
<td>0.913 (0.054)</td>
<td>0.013 (0.025)</td>
</tr>
<tr>
<td>Incumbent vote share, 1998</td>
<td>0.149 (0.101)</td>
<td>0.030 (0.111)</td>
</tr>
<tr>
<td>Dummy = 1 if position pays no salary</td>
<td>0.132 (0.126)</td>
<td>0.119 (0.090)</td>
</tr>
<tr>
<td>Total number registered voters</td>
<td>0.066*** (0.002)</td>
<td>-0.274 (0.235)**</td>
</tr>
<tr>
<td>Constant</td>
<td>0.588* (0.322)</td>
<td>0.235** (0.363)</td>
</tr>
</tbody>
</table>

Summary statistic:
- N = 184
- R² (pseud R² for logit) = 0.03
- Unit of analysis: Incumbent

Source: Author's calculations.

a. Robust standard errors in parentheses, with clustering by school district. * p < .10, two-tailed test; ** p < .05, two-tailed test. 

The results are not altogether surprising. Again, because voters in this off-year election were likely more educated, they may have been less influenced by a candidate's name recognition or campaign spending, the kinds of characteristics that lagged vote shares mean to capture. While high-information voters in off-year local elections may not vote retrospectively, they probably do care about board members' policy positions. Which policy positions? As Moe demonstrates in chapter 11 of this volume, low-turnout school board elections attract a disproportionate share of high-interest stakeholders, such as teachers and administrators. It seems likely that these stakeholders.
stakeholders evaluated board members less on the basis of test scores than on their demonstrated commitment to the stakeholders' employment interests. That change in per-pupil expenditures correlated positively with the probability that incumbents ran for office in 2002 (recall, they registered negative and significant impacts in 2000) certainly is consistent with the claim that vested interests exerted a disproportionate level of influence during this off-year election.

Unfortunately, without individual-level data, we cannot further examine Zaller's and Moore's claims about high- and low-information voters and the composition of school board electorates in on- and off-year elections. We leave it to future research to extend their arguments to different kinds of local elections with variable turnout rates. At a minimum, though, we note that retrospective voting is not a forgone conclusion in all local elections, and, more specifically, that the draw of a presidential election appears to improve the chances that electorates hold board members accountable for the performance of schools.

Still, it is worth recognizing that the 2000 and 2002 races in South Carolina present a tough test for retrospective voting and hence for democratic accountability. First, we are working with aggregate voting data from less than 100 different districts. From a purely statistical standpoint, individual-level data collected from a wider variety of races may produce more pronounced effects. Second, South Carolina's accountability system is relatively new. Over time, as voters grow accustomed to the testing regime and learn more about the performance of their schools, retrospective voting may become more common. And finally, unlike with local elections that revolve around issues of public safety or property taxes, large proportions of the electorate have little reason to be informed about the performance of public schools.

Conclusion

The 2000 findings presented in this chapter suggest that explicit sanctions and rewards associated with hard accountability systems may supplement rather than stand in for accountability systems already built into public education. Those charged with governing public schools currently have strong electoral incentives to promote policies and develop practices that enhance student learning. When test scores dropped in South Carolina schools, incumbents were less likely to run, they were more likely to face competition, and they won (if they won) by smaller margins. As long as the electorate has the information needed to evaluate student learning—test scores providing one such source—many voters appear willing to punish and reward those who govern public schools according to their performance in office.

An analysis of NCLB, the most important accountability system currently in place, illustrates the specific ways in which formal accountability systems complement—and in some cases conflict with—long-standing forms of democratic accountability. In three ways, NCLB strengthens existing forms of democratic accountability. First, under NCLB, schools and districts are evaluated every single year, while school board elections are held only every two years—and, as the 2002 elections suggest, voters do not always hold board members responsible for changes in student test scores. Second, to avoid penalties under NCLB, public schools must demonstrate annual improvements in test scores; it is not enough simply to maintain last year's performance levels. Retrospective voters, meanwhile, appear more likely to rise up and punish their incumbent board members only when test scores decline. Third, NCLB also mandates that schools demonstrate annual improvements for various subgroups; students for whom English is not the primary language, ethnic and racial minorities, and special education students. Without individual-level data, it is extremely difficult to discern whether voters pay attention to the test scores of all students or those within specific subpopulations. If citizens vote primarily on the basis of their own child's test scores or those of their own ethnic group more generally, then minorities cannot hope to affect the composition of a school board. NCLB, however, provides the needed backup, establishing incentives for schools and school boards to improve the test scores of students who come from weak voting blocs in local elections.

In other ways, however, state mandates and democratic accountability systems find themselves in tension with one another. When evaluating schools, NCLB does not make any allowances for temporary setbacks in student learning. For every year that a school fails to make the prescribed gains in test scores, preestablished penalties automatically apply. Retrospective voters, meanwhile, may rightfully see success in constant ten scores, especially at schools that confront especially disadvantaged student populations. When NCLB classifies such schools as underperforming and mandates the reallocation of Title I funding, retrospective voters may respond by rallying behind incumbent school board members in defiance of a meddling federal government.

36. Individual-level data may permit additional empirical explorations. Evidence of retrospective voting may derive from either a small percentage of high-information voters who vote strictly on the basis of changes in test scores or from a partially informed electorate that pays causal attention to test scores. Similarly, we do not know whether citizens vote on the basis of test scores or changes in student learning more generally. It would be that citizens obtain direct information about how students score on standardized tests and vote accordingly; alternatively, those results might derive from parents observing students and schools—but not test scores—provided that changes in test scores correlate positively with changes in student learning. Third, more data are needed to distinguish these various causal pathways.

37. PACT scores in South Carolina have become more widely publicized. The state now sends all parents in the state notification of their school's performance, and districts are required to report school test scores in local newspapers.

38. We do not know whether voters in South Carolina are responding to test scores per se or whether they are formulating independent evaluations of student achievement for which test score trends are acting as proxies.
Still, as a practical matter, both state mandates and democratic accountability systems may fail to achieve their intended objectives. Test scores may increase because students are learning more, because students have grown accustomed to taking standardized tests, because teachers and students are cheating on tests, or because schools experience an influx of higher-performing students. When doling out punishments and rewards, NCLD does not distinguish among these possibilities. Unfortunately, the data from South Carolina do not allow us to determine whether retrospective voters attempt to assess the causes of test score changes. Given the difficulties of assessing school quality, however, there is good reason to believe that even the most adept retrospective voters have a hard time doing so. Hence, while state mandates and democratic accountability systems hold schools and school boards accountable, it is not clear that either hold them accountable for their contributions to student learning per se.

It may be years before it is known whether the latest round of accountability reforms have had their intended effects. But it is clear, even now, that accountability is nothing new. NCLD did not introduce accountability into American education, nor did any state's accountability system suddenly make local educators answerable to its citizens. Ironically, even those critics who point to the poor relative performance of American students on international tests as justification for new accountability systems seem oblivious to the fact that American schools are perhaps the most accountable in the world. With nearly 15,000 school districts governed by more than 80,000 popularly elected officials, American public education may be the most democratically accountable institution in this nation or any other. Analysts of the current accountability movement would do well to focus on the specific content of the reform proposals and ignore the rhetoric of accountability. NCLD and other performance-based reforms may bring new standards, new tests, and new incentives to American education. What they will not introduce are systems of accountability. They are already here.

Appendix

Calculation of vote share for multiple-seat elections

School board members are elected as large in thirty-two districts, by "constituent district" (ward) in thirty-eight, and by using a combination of both methods in fifteen districts. Some of the at-large and combination districts allow for "multiple-winner" elections, in which voters are allowed to vote for more than one candidate. For multiple-winner contests, we computed vote share as the number of votes the incumbent received divided by the maximum number of votes he or she could have received. The maximum number of votes that any candidate can receive is equal to the number of voters (a voter cannot vote for the same candidate more than once). For example, consider an election in which 100 voters turn out, each of whom votes for three of five candidates running. If a candidate in this election receives 60 votes, then we compute his or her vote share to be 60 percent, meaning that 60 percent of the voters who could have voted for this candidate actually did so. Put differently, with a total of 300 votes cast by 100 voters, 60 votes is 20 percent of total votes, but 60 percent of possible votes. It is the latter number that we use as our measure of vote share in multiple-seat elections.

Calculation of district- and precinct-level changes in test score

Palmetto Achievement Challenge Tests results are reported both as "scale scores" and as "performance levels." We chose to use the scale scores, which provide more detail to distinguish performance among schools. Scale scores are determined independently for each grade and subject. The South Carolina Department of Education set the average scale score at 100 times the grade level, with a range of 128 scale score points around this average—for example, the initial average score for grade 3 is 300 and the range is 236 to 364.

In developing our performance measure, we began by computing a school-level scale score. Because scores for each grade were on different scales, we first converted them into within-grade statewide percentile rankings. For each school, we then averaged these percentile scores across grades, weighting by enrollment in each grade. These steps were conducted separately for English and math scores, and then the two were averaged to produce a composite score. The result is a school-level percentile score. Based on these school-level scores, we then calculated a score representing the average of nearby schools for each voting precinct. Specifically, for each precinct, we computed the average score of all the schools in the same zip code as the polling place. 39 For zip codes with only one school, that school's performance measures served as the zip code values. For zip codes in which more than one school was located, we averaged across the schools, weighting by enrollment. For zip codes where no schools were in operation, we imputed the achievement of schools in the nearest zip code with at least one school. 40 The result is a precinct-level percentile score. The district-level test score was based on the same underlying school-level percentile scores, 39. While zip codes are not perfect mappers of "neighborhoods," they are the best possible approximations given our data. Attempts to match schools and precincts based on geography did not prove more successful than zip code matching because a large number (nearly half) of the addresses in our database could not be geocoded to a level beneath the zip code.

40. Distance between zip codes was calculated using the "Great Circle" formulas, measuring distance between zip code latitude-longitude centroids. The coordinates of the county centroid were used when the zip code could not be located in the Census Bureau's TIGER (Topologically Integrated Geographic Encoding and Referencing) database (www.census.gov/geo/www/tiger/tigerindex.html [December 3, 2004]).
and we simply averaged over all the schools in the district, weighted by enrollment. Test score change for both precursors and districts was computed as the difference between 2000 and 1999 percentile scores. Thus our test score change variable measures the increase or decrease in statewide percentile ranking over the year preceding the election.

References


Minority Incorporation and Local School Boards

MELISSA J. MARSHALL

Since passage of the Voting Rights Act in 1965, social scientists have periodically examined the state of minority representation in U.S. politics. Although the vast majority of that research has focused on patterns of minority office holding at the local level, nearly all of it has examined African American representation in municipal-level offices. To date, considerably less attention has been paid to black and especially Hispanic representation on local school boards or in other local elected positions. Numbering approximately 15,000 across the United States, school boards are not only the most prevalent form of government in this country but also the most common point of entry into public office among those seeking political careers. In addition, schooling is not simply the most expensive local service provided by state and local governments but also unquestionably one of the most important public services that government provides. There is much at stake in local school politics, and consequently minority incorporation in this policy arena warrants greater scholarly attention.

In this chapter, I address the issue of minority incorporation by analyzing each of its dimensions—descriptive and substantive representation. I begin with descriptive representation, which refers to the extent of minority representation

1. See, for example, Soss and Pinnear (2000), Bullock and MacManus (1987), and Engstrom and McDonald (1984).