Institute for Research on Poverty
Discussion Paper no. 1142-97

School Finance Reforms, Tax Limits, and Student Performance: Do Reforms Level Up or Dumb Down?

Thomas A. Downes
Department of Economics
Tufts University
tdownes@tufts.edu

David N. Figlio
Department of Economics
University of Oregon
dfiglio@oregon.uoregon.edu

September 1997

We are indebted to the National Center for Education Statistics of the U.S. Department of Education for making available the data for this research. Our thanks to Helen Connolly for her invaluable assistance and to Bill Fischel, Bill Duncombe, Linda Loury, Kim Rueben, Joe Stone, Ann Velenchik, and participants in the annual meeting of the American Education Finance Association for their helpful comments and suggestions. Blame for all remaining errors lies with us.

IRP publications (discussion papers, special reports, and the newsletter *Focus*) are now available electronically. The IRP Web Site can be accessed at the following address: http://www.ssc.wisc.edu/irp/

Abstract

During the late 1970s and early 1980s, a majority of states substantially changed the ways in which schools were funded, either directly through court- or legislatively mandated school finance reform, or indirectly through tax and expenditure limits. To date, there have been few academic attempts to gauge the effects of these policy changes on actual outcomes of education. This paper is an attempt to fill this gap in the literature. We find compelling evidence that the imposition of tax or expenditure limits on local governments in a state results in a significant reduction in the mean for that state of student performance on standardized tests of mathematics skills. We also find that finance reforms in response to court mandates do not result in significant changes in either the mean level or the distribution of student performance on standardized tests of reading and mathematics. In addition, substantial finance reforms that are not legislative responses to explicit court mandates generally result in increases in mean student performance. Further, in those states that have implemented finance reforms of this type, the test performance of students residing in localities in which local revenues formed smaller shares of total revenue prior to the reforms improve relative to others after the reforms are implemented.

School Finance Reforms, Tax Limits, and Student Performance:
Do Reforms Level Up or Dumb Down?

1. INTRODUCTION

Historically, local control of public education has meant that per pupil funding and school quality were closely related to the wealth in the locality and to the demographic characteristics of the locality. In the past two decades, policy changes of two distinct kinds have weakened this relationship. First, partly in response to the perception that property tax burdens have become increasingly onerous, property tax limits have been passed, either by legislation or voter initiative, in twenty states. Second, state supreme courts in a number of states, most recently Ohio, Vermont, and New Jersey, have invalidated the existing system of school finance in decisions that have highlighted the fact that equalization aid has not accomplished the goal of equalizing educational opportunities. While these policy changes have lessened the reliance on property taxes as a source of school funding and have mitigated inequalities in per pupil spending, the changes have also been accompanied by reduced local control over school funding, a potentially deleterious outcome. For example, observers have argued that the experience in California with strict limits on local discretion over the level of public school expenditures, and with the accompanying reduction in the ability of districts to tailor their programs to local needs and preferences, raises troubling questions about the effects of these policy changes on all districts in a state, particularly those districts with low wealth.

A number of authors (Silva and Sonstelie 1995; Downes and Shah 1996; Evans, Murray, and Schwab 1995; Manwaring and Sheffrin 1995; Hoxby 1996) have attempted to determine if finance reforms of the type implemented in California in the late 1970s succeeded in improving the absolute position of low-wealth districts. The results of this research are equivocal, apparently because the effect of these reforms on state-level support for public education is indeterminate. However, by focusing on per-pupil spending instead of on student performance, all of this research fails to address the basic question facing policy makers: Is the problem of inadequate educational opportunities in low-wealth, high-cost districts best addressed by shifting control of school financing to the state, or can this problem be better addressed within a system that allows substantive local control over educational finance and policy?

More work has been done examining the effects of tax limits on student achievement, but no conclusive estimate of these effects exists. To our knowledge, the only study using national-level data on the impact of tax limits on student outcomes is Figlio 1997a. However, since the estimates in this paper are derived from cross-sectional variation, the results are subject to the criticism that the effects of limits have not been disentangled from unobserved factors related both to student performance in a state and to the likelihood that a limit will be passed. Further, cross-sectional analysis does not provide any indication of how the pre-limit status of districts is related to the post-limit performance of students in those districts.

Our goal in this paper is to begin to rectify these shortcomings of the current literature. In particular, we seek to determine both how the finance reforms and tax limits of the late 1970s and early 1980s affected the distribution of student performance in reform states and how student performance has changed in states in which these policies have been implemented relative to student performance in states in which there have been no major policy changes.

To implement this empirical strategy, we use two national data sets, the National Longitudinal Study of the High School Class of 1972 (NLS-72) and the 1992 (senior year) wave of the National Education Longitudinal Survey (NELS), that provide much of the necessary student- and school-level information. We link these data to the Census of Governments for 1972 and the Common Core of Data for 1991 and to the Census of Population and Housing for 1970 and 1990 to get the necessary financial and demographic information. The NELS data were collected sufficiently far from the passage of most tax limits and school finance reforms to allow us to quantify the long-run effects of these policy changes by analyzing changes in the distributions of student performance between the NLS-72 cross section and the NELS cross section.

We find that the imposition of tax or expenditure limits on local governments in a state results in a significant reduction in the mean for that state of student performance on standardized tests of mathematics skills. Further, we find significant heterogeneity in the effects of tax limits. Specifically, these reductions are larger in localities with lower levels of per pupil spending prior to the imposition of limits. The effects of finance reforms on the level and distribution of student performance are less clear. The estimates we find most persuasive support the conclusion that finance reforms in response to court decisions do not result in significant changes in either the mean level or the distribution of student performance on standardized tests of reading and mathematics. We also find that substantial finance reforms that are not legislative responses to explicit court mandates generally result in increases in mean student performance and that, in those states that have implemented finance reforms of this latter type, the test performance of students residing in localities in which local revenues formed smaller shares of total revenue prior to the reforms improve relative to others after the reforms are implemented. All of these results are robust to modeling the implementation of these policies as endogenous.

The next section of the paper reviews the existing theoretical and empirical literature on the relationship between the distribution of student performance in a state and the imposition of either tax limits or finance reforms in that state. Section 3 provides an overview of the methods we use to estimate the effects of these policy changes. After discussing the data in Section 4, we present our results in Section 5. We close the paper with concluding remarks and suggestions for future research.

2. SCHOOL FINANCE REFORMS, TAX LIMITS, AND STUDENT PERFORMANCE: BACKGROUND

Much of the popular discussion of the effects of tax limits and school finance reforms on the level and distribution of student performance starts with the California case. For example, the author of a recent article in the *New York Times* discussing the efforts in California to reduce class size suggests that the combination of declining per pupil spending and a growing limited English proficient (LEP) population have led to declines in student performance (*New York Times* 1996).[1] However, no systematic evidence exists on the effects of the policy changes of the late 1970s on the average performance of California students on standardized tests.[2] The findings of Downes (1992) actually appear to contradict the impression left by the *New York Times* article; he found that the average performance on the California Assessment Program (CAP) test of sixth-graders in the unit (K–12) districts rose in the period from 1976–77 to 1985–86. However, since the CAP test changed over time, cross-time comparability of the CAP scores is a matter of debate.

A small but growing body of theoretical work suggests that the absence of a clear picture of the post-policy changes in California may in part be attributable to the fact that, depending on the environment and the structure of the education production function, mean student outcomes could be either higher or lower in the post-policy equilibrium. For example, while his focus is on the general equilibrium effects of the introduction of private school vouchers, Nechyba (1996) provides evidence

[1]The article in question also notes that the average score of fourth-graders from California on the reading portion of the 1992 National Assessment of Education Progress (NAEP) test exceeded only the average scores for Mississippi. In addition, the average scores of both fourth- and eighth-graders from California on the mathematics portion of the 1992 NAEP were substantially below the national and regional averages. However, since no evidence on trends in NAEP scores exists, it is impossible to conclude on the basis of these results that student performance in California has declined in the aftermath of the policy changes of the late 1970s.

[2]This statement is not quite true; our data provide a partial picture of changes in student performance in California relative to the changes elsewhere in the country. We will elaborate on this comparison below.

that both spending and perceived school quality could be lower under a state-financed system than under either a foundation aid system or a district power equalization system.

Like Nechyba, Bénabou (1996) builds a model in which individuals are heterogeneous and community formation is endogenous. He uses this model to show that the impact on student performance of a move from locally financed schools to state-financed schools depends critically on the importance of both peer effects and purchased inputs in production and on the extent of cross-community migration that the move to state financing induces. Both increases in and decreases in the mean level of educational achievement are possible in the model. In addition, Bénabou concludes that, if no restrictions are placed on the type of communities that can form, a system of locally financed schools will be inefficient in the sense that aggregate surplus is not maximized. Fernandez and Rogerson (1997) reach a similar conclusion. In particular, using a two-period overlapping generations model, they argue that moving from local to state financing can raise aggregate welfare by 10 percent by generating a more equal distribution of income in future generations. However, this conclusion is contingent upon the assumed form of the educational production process.

As the discussion above makes clear, part of the reason why there is no clear prediction of the impact of school finance reforms on the level and the distribution of student performance is that the effects of these reforms on spending are unclear (Leyden 1992; Silva and Sonstelie 1995). However, even if school finance reforms level-up spending, as Evans, Murray, and Schwab (1995) argue, the conclusion that performance will also be leveled up does not necessarily follow. Similarly, even if, as the empirical literature indicates (Figlio 1997a, 1997b; Dye and McGuire 1997), tax limits reduce both the level and the growth of per pupil spending, we cannot presume that declines in student performance will be the result. As Hanushek's 1986 review and subsequent updates, such as Hanushek 1996, make clear, the central problem is that researchers have been unable to consistently document a link between inputs to the schooling process and student performance.[3] In fact, the logic of proponents of the budget-maximizing bureaucrat model could be used to support the argument that some or all of the post-finance reform increases in state aid received by low-wealth districts will be wasted. As a result, if there is any redistribution from high-wealth to low-wealth districts, mean performance could decline, even if mean per pupil spending in the state increased. Similarly, in the aftermath of a tax limit, performance gains could result if the limits forced school district administrators to reduce waste both by limiting available resources and by increasing voter attention on how these resources are used.[4]

While possible, these seemingly perverse outcomes are unlikely for several reasons. First, the recent empirical work of Krueger (1997), Ferguson (1991), Ferguson and Ladd (1996), and Figlio (1996) has again raised the possibility that additional dollars are productive, though Figlio asserts that these productivity gains, though statistically significant, are still very modest. Second, questions about the validity of Hanushek's (1986) assessment of the literature that dollars do not matter have been broached in more recent literature reviews by Hedges and Greenwald (1996) and Dewey, Husted, and Kenny (1997). Third, recent work by Ehrenberg and Brewer (1994) and others presents compelling evidence that while traditionally used measured school inputs do not affect student achievement much, attributes reflecting teacher quality do.[5]

In addition, in the case of tax limits, even if schools in the districts affected by these limits are capable of reducing waste and thus maintaining the pre-limitation level of student performance, the limits typically provide no explicit incentive to administrators of these districts to eliminate waste. In fact, if school district administrators are budget maximizers and if technical inefficiency in the schools persists because insiders in school systems have more information than outsiders about how resources

[3]Similarly, Betts 1996 shows that studies linking student inputs and labor market outcomes are tenuous.

[4]See Downes 1996 for a formal model linking the imposition of limits and student performance.

[5]Figlio and Rueben (1997) find evidence that tax limits systematically reduce the levels of teacher quality attributes shown in the literature to affect student academic performance.

can be used productively, then administrators may have an incentive to allow student performance to decline and to use this decline to argue for additional resources.[6] Similarly, in those districts that have more to spend in the aftermath of finance reforms, budget-maximizing administrators have an incentive to use some, if not all, of the additional resources effectively. Otherwise, future requests for additional resources are likely to fall on deaf ears.

The main lesson from this brief review of the theoretical literature is that establishing the relationship between tax limits, finance reforms, and student performance is an empirical task. To accomplish this task, we build primarily on a small set of empirical papers that have examined the effects of finance reforms and tax limits on student performance. The most direct antecedent in this line of research is Downes 1992, in which the extensive school finance reforms in California in the late 1970s were analyzed. This work indicated that greater equality across school districts in per pupil spending was not accompanied by greater equality in measured student performance. However, because this research focused on the possibly unique California case, the generalizability of the conclusions are debatable.

Hoxby 1996 represents the first attempt to use national-level data to examine the effects of finance reforms on student performance. Although Hoxby is primarily interested in determining how finance reforms change the incentives facing local districts and, thus, per pupil spending, she also considers how these changes are related to dropout rates. She finds that, on average, dropout rates increase about 8 percent in states that adopt state-level financing of the public schools. In addition, while Hoxby's work does not explicitly address the effect of equalization on the within-state distribution of student performance, it seems likely that much of the growth in dropout rates occurred in those districts with relatively high dropout rates prior to equalization. In other words, these results imply that equalization could adversely affect both the level and the distribution of student performance.

---

[6]Figlio and O'Sullivan (1997) provide evidence suggesting that such a situation occurred with regard to police and fire protection in the wake of tax revolt-era limits.

While Hoxby raises an important point, it is possible that her approach misses key features of school finance reforms. Because she does not explicitly account for the imposition of tax or expenditure limits and because the passage of such limits is, in many cases, roughly contemporaneous with school finance reforms, it is unclear whether the changes in performance observed by Hoxby are attributable to school finance reforms or to the imposition of tax or expenditure limits. Furthermore, Hoxby's method does not explicitly account for changes in direct state support of public schools, instead focusing on local incentives.

While the dropout rate is an outcome measure of considerable interest, analyses of the quality of public education in the United States tend to focus on standardized test scores and other measures of student performance that provide some indication of how the general student population is faring. Recent work of Husted and Kenny (1996) suggests that equalization may detrimentally affect student achievement. Using data on 37 states from 1987–88 to 1992–93, they find that the mean SAT score is higher for those states with greater intrastate spending variation. However, like Hoxby, Husted and Kenny fail to control for the imposition of tax or expenditure limits. Also, the period they consider postdates the imposition of tax or expenditure limits. Thus, the data do not permit direct examination of the effects of policy changes. In addition, because they use state-level data, Husted and Kenny cannot examine the intrastate impact of equalization. Finally, since only a select group of students take the SAT, Husted and Kenny are not able to consider how equalization affects the performance of all students in a state.

Two recent papers, which focus on the effects on student performance of revenue and expenditure limits and not finance reforms, provide the model for the empirical work in this paper. Using a cross section of student-level data from the NELS, Figlio (1997a) finds that, ceteris paribus, the performance of tenth-graders on mathematics, reading, science, and social studies tests is significantly lower in those states in which local school districts face either revenue or expenditure limits. Downes, Dye, and McGuire (1997) use the recent imposition of property tax limits on school districts in the Chicago suburbs to argue that it is not possible to conclude in the short term that these limits have translated into slower growth in student performance. For the current research, what is relevant about these two papers is not their seemingly contradictory answers but rather the methodology they suggest. In particular, this research implies that evaluating the effects of finance reforms requires not only before-and-after data on students in districts that have implemented reforms but also a control group of students from states in which no significant finance reforms have been enacted. In addition, both of these papers seem to confirm an above-noted prediction of theoretical work: the effects of a finance reform may not be uniform within a state because these effects depend critically on a district's initial conditions and demographics. This is a potential reality to which we are sensitive in our empirical work.

3. METHODS

Previous research makes clear that further analysis of the relationship between tax limits and student performance and school finance reforms and student performance is needed. However, earlier work does not provide an obvious method for exploring these links. For example, the earliest work, which focused on California, used an event analysis, treating the policy changes of the late 1970s as exogenous. While the reasonableness of treating these changes as exogenous can be questioned, the arguments for assuming exogeneity in the California case are relatively strong.[7] In the case of other states, the exogeneity assumption is probably not tenable (Fischel 1996; Figlio 1997a). Further, these policy changes are not similar events (Hoxby 1996; Figlio 1997a). The bottom line is that an event analysis approach will provide what is, at best, an imperfect understanding of the links between tax limits and student performance and between finance reforms and student performance.

---

[7] See Downes and Greenstein 1996 for further discussion of these issues.

These criticisms of the standard event analysis approach suggest potential modifications that could mitigate the problems. Figlio (1997a) and Rueben (1995) use information on the variation in the language of each state's constitution to instrument for the imposition of a tax or expenditure limit. Other authors (e.g., McUsic 1991) suggest that variation in constitutional language may also provide reasonable instruments for a state supreme court decision mandating equalization of school finance. In both cases, these arguments make use of the fact that state constitutional language predates, by many years, the policy changes of the late 1970s and early 1980s. As a result, unobservable factors that may be correlated both with the imposition of a tax limit or finance reform and with changes in student performance in a state are unlikely also to be correlated with the language of a state's constitution.

Coping with the heterogeneity of these policy changes is a more difficult task. Recent work (Downes and Shah 1995; Hoxby 1996; Evans, Murray, and Schwab 1997) has emphasized the tremendous diversity of school finance reforms. Any attempt to classify finance reforms will be imperfect. Nevertheless, there is general consensus that the key elements of a finance reform are the effects of the reform on local discretion, the effects of the reform on local incentives, and the change in state level responsibilities in the aftermath of reform (Hoxby 1996; Courant and Loeb 1997). Downes and Shah (1995) suggest that one crude method for distinguishing between reforms is to treat as different those reforms implemented in response to court mandates. This distinction between court-ordered and legislative reforms makes use of the fact that reforms of the former type tend to impose more substantive constraints on local discretion and to shift more responsibility to the state.

Hoxby (1996) correctly criticizes the use of court mandates as the basis for classification, arguing that this method ignores the fact that certain legislative reforms have dramatically limited local discretion while some of the court-ordered reforms have made it less costly, from the local perspective, to increase spending by one dollar. Hoxby goes on to contend that the defining characteristic of finance reforms is the effect these reforms have on local tax prices. In particular, reforms that dramatically increase tax prices are anti-spending in the sense that they reduce the local incentive to increase spending, while reforms that have the opposite effect on tax prices are pro-spending.

Our data do not permit us to construct tax prices in 1972. As a result, we cannot pursue Hoxby's strategy of using the tax price as a policy variable. In addition, we are not comfortable with the strategy of assuming that all of the effects of finance reforms work through the tax price.[8] As noted above, such a strategy ignores the fact that state-level decisions can have dramatic effects on overall spending while leaving tax prices unchanged. Further, the environment in which these decisions are made is likely to change dramatically in the aftermath of finance reforms. Thus, we have chosen to explore two distinct classifications of reforms. First, we divide reforms into court-ordered or legislative, as was done in Downes and Shah 1995. Second, following Hoxby 1996, we classify some reforms as pro-spending and others as anti-spending.[9]

Like finance reforms, tax limits vary in the extent to which they limit local discretion. We follow Figlio 1997a in the rules we use to determine whether school districts in a state are subject to a revenue or expenditure limit. The publication *Tax and Expenditure Limits on Local Governments* (Advisory Commission on Intergovernmental Relations 1995) provides the information used to determine if a limit is present. Although general expenditure and revenue limits are potentially different from property tax levy limits, and both of these types of limits are different from limits on the nominal tax rate together with limits on assessment growth, we do not differentiate between these types of limits. Nor do we account for whether districts subject to tax limits find these limits binding.

[8] We are also unwilling to assume that the valuation of property in their jurisdictions is a choice variable for school district administrators. For those states that use a foundation aid system of financing, such an assumption is needed to produce tax prices different from one, and within-state variation in tax prices.

[9] Reforms that have little or no impact on tax prices are not included in either of these classifications.

To estimate the effects of school finance reforms and tax limits, we use several variants of the following basic specification of the natural log of the performance on a standardized test (O_{ijt}) of student i residing in district j in year t:

$$O_{ijt} = \alpha + B_{ijt}\beta + F_{jt}\gamma + D_{jt}\rho + F_{j0}\theta + C_{j1}\lambda + C_{j1}F_{j0}\phi + \delta Y_1 + u_{ijt} , \qquad \textbf{(1)}$$

where t = 0 is 1972 and t = 1 is 1992. Here, vector B_{ijt} consists of characteristics of the student and his or her family, vectors F_{jt} and D_{jt} of financial and demographic characteristics, respectively, of the school district in which the student resides, vector C_{j1} of indicators of the imposition in the period from 1972 to 1992 of tax limits or finance reforms. In addition, Y_1 is a dummy indicating the year is 1992 and u_{ijt} is a white-noise error.

The starting point for our first approach to ascertaining the effects of policy changes is the imposition of the constraint $\theta = \lambda = \phi = \delta = 0$. We then estimate the remaining parameters separately for each year. Using these estimates as a benchmark, we then determine how the performance of each student in our sample in 1972 would have changed if the district in which that student resided had its 1992, instead of its 1972, financial characteristics. We do the same exercise for each student in our sample in 1992. In each case, by comparing predicted performance when the district has 1972 financial characteristics to predicted performance when the district has 1992 financial characteristics, we determine the change in student performance attributable to the change in district characteristics. Then, by comparing the mean performance changes for students residing in states with tax limits or finance reforms to the mean changes for those residing in states without these policy changes, we can compute one possible measure of the impact of each policy change on performance.

Measures of policy effects calculated using this first methodology have two obvious flaws. First, these measures can only provide correct answers if all of the effects of the policies work through changes in the observed district financial characteristics. In addition, the answers can only be correct if all of the changes in district characteristics are attributable to the policies. Neither of these conditions are likely to be satisfied. Thus, we also use a more traditional event-analysis approach to measure these effects. In particular, we start by imposing the constraint $\gamma = \phi = 0$. In this case, the strength of the effects of the policy changes is reflected in the estimated value of λ.

Again, if we take seriously the argument made in both the theoretical and the empirical literatures that a district's initial conditions will influence the effect of tax limits or finance reforms on that district, this standard approach will also provide an incomplete picture of the effects of these policy changes. Thus, the final variant of (1) that we consider imposes only the constraint that $\gamma = 0$. Policy effects are thus allowed to vary as the district's initial financial characteristics vary.

4. DATA

The data for this analysis are drawn primarily from two National Center for Education Statistics (NCES) administered surveys of high school seniors, the National Longitudinal Survey of the High School Class of 1972 (NLS-72) and the National Education Longitudinal Survey (NELS). To insure comparability to the NLS-72, we use the second follow-up to the NELS that was administered in 1992. The second follow-up is a representative sample of the high school class of 1992, and therefore provides a natural comparison to the high school class of 1972.

Each of these surveys provides detailed information on the sampled students, their parents, and the schools they attend. In addition, since we have access to the restricted-use version of the NELS, we are able to identify in both cross sections the school district in which each school is located.[10] Using this information, we were able to match to each student financial and demographic data for the district in

[10]Both the NLS-72 and the NELS are stratified samples with the school as the basic sampling unit. Thus, the possibility exists that, in large districts, the sampled schools are unrepresentative of the district in which they are located. While this would only create problems for our analysis if there was some relationship between the extent to which a school is atypical of its district and the implementation of a policy change, we used the NCES's Common Core of Data to compare characteristics of the sampled schools in the NELS to the overall characteristics of the schools in the sampled schools' districts. No statistically significant differences existed.

which the student attended school. In particular, for the NLS-72 we use financial data from the 1972 Census of Governments and demographic data from the 1970 Census of Population and Housing. For districts in the NELS, we draw financial data from the NCES's Common Core of Data and demographic data from the 1990 Census of Population and Housing.

Limitations of the earlier data sources restrict the financial and demographic data available for use. For each district, we are able to construct measures of total expenditures per pupil, the full-time equivalent (FTE) teaching staff per pupil, the FTE administrative staff per pupil, and other FTE employees per pupil. In addition, since we have information on each district's revenue sources, we can calculate the percentage of revenues raised locally.[11] Districts with more taxable resources tend to raise larger shares of their revenues locally. Thus, finance reforms that have as their goal a reduction in the link between spending and taxable resources are likely to have larger absolute effects on the performance of students residing in districts with larger shares of revenue raised locally. Similar reasoning leads us to expect that, all else equal, districts with low initial levels of spending should benefit disproportionately from finance reforms. Finally, we also expect that districts with lower initial levels of spending will, all else equal, tend to find the adjustments to a post-tax limit regime more difficult to make.

For each district in each cross section, we are able to calculate a relatively rich set of demographic variables. In particular, we can compute the percentage of adults residing in the district who are high school and college graduates, the percentage of families in the district with incomes below the poverty line, the mean income of families in the district, and the fraction of the residents of the

[11]For fiscally dependent districts, the 1972 Census of Governments does not give enrollment numbers and does not provide a decomposition of school district revenues independent of the decomposition of the revenues of the parent government. To insure consistency between observations in the NLS-72 and to avoid omitting students residing in fiscally dependent districts, for all districts we use the district enrollment data supplied in the NLS-72. For dependent districts, we use the decomposition of revenue sources for the parent government to calculate the percentage of revenues raised locally. As a check, we reestimated our basic specifications omitting the dependent school districts and using the enrollment figures supplied in the Census of Governments. There were no substantive changes in the results.

district who are African American or Hispanic. These controls provide us with proxies both for unobserved family inputs into education and for the characteristics of the student's peers.

We control for a number of individual-specific variables that are available in both the NELS and NLS-72. Specifically, we account for whether the student is female, African American, or Hispanic. We control for whether the student considers him- or herself to be Protestant, Catholic, Jewish, or unaffiliated with a religion (other religions are the omitted category). We have limited parental education data, but we can account for whether these data are present in the database and whether the student's father and/or mother has a bachelor's degree. We control for whether the student's family's income is at least $10,000 (in 1990 dollars), at least $25,000, at least $35,000, and at least $50,000, and for whether these income data are omitted from the database.[12] We account for the number of siblings that the student has, and since the NELS truncates this variable at six, for whether the student has at least six siblings. We control for whether the student reports that he or she speaks English well, and for whether the student's family subscribes to a newspaper. Finally, we account for the number of mathematics Carnegie units that the student has taken (reported from transcript data).[13] We opt not to directly include measured school inputs (e.g., student-teacher ratio) in our achievement equations, as they are likely to be functions of the finance reforms and tax limits.[14]

---

[12]In both data sets, we only have income ranges, rather than exact income amounts. We chose these particular income ranges because they are the only ones for which we have a virtually exact range match, after adjusting for inflation, amongst the possible income ranges provided by each database.

[13]It may seem odd that we choose to use the number of mathematics units even in reading test equations. However, while we believe that the number of mathematics classes taken affects student outcomes (Levine and Zimmerman 1995), we also contend that this may be an indicator of unmeasured student (or parent) motivation and ability. This is borne out by the fact that the number of English units has significant effects on reading achievement when included independently, but loses its significance when the number of mathematics units enters the reading equations. On the other hand, the t-statistics of the coefficient on mathematics units in reading equations generally exceed 40.

[14]As a robustness check, we have estimated all of the specifications presented below with the district-level spending measures replaced with corresponding input measures (e.g., pupil-teacher ratio in 1972). None of the conclusions change.

The means of the variables used in our analysis are given in Table 1. Observations were omitted if information on any of the variables was missing. In addition, because in the 1972 Census of Governments the financial data for North Carolina cities with dependent school districts were merged with the financial data for the surrounding counties with separately operating dependent school districts, students attending schools in districts in these cities and counties were omitted. Finally, since 1972 per pupil expenditures could not be calculated for any dependent school district in the NELS that was not also in the NLS-72, students in the NELS who attended school in such districts were omitted. These omitted observations account for about 6 percent of the total number of observations with otherwise complete data.

To avoid any sample selection concerns, we used as our measures of student outcomes the scores of students on the reading and mathematics tests given to all students in the NLS-72 and to all students in the NELS. Since the scale of the NELS test scores differed from the scale of the NLS-72 test scores, for ease of direct comparison we renormed the NELS reading and mathematics scores to create easily comparable outcome measures. The students in each sample with ACT scores were used to benchmark the renormed scores.[15]

Among the districts in each sample, 189 are represented in both of the samples. This panel data set has the obvious advantage of allowing us to control for unobserved district-level heterogeneity. Unfortunately, the students in this panel of districts differed significantly (usually at the 1 percent level or better) in most dimensions from the remaining students in the two cross sections, the only exceptions being whether the student expressed no religious affiliation and whether the student’s family received a daily newspaper. That the panel is not representative is not surprising, since the districts in the panel are far more likely to be located in central cities. Nevertheless, the panel provides us with a check on the

[15]We also used those students in each sample with SAT scores to benchmark the renormed NELS scores. Using the test scores generated using this alternative renorming fails to alter any of the conclusions presented below.

TABLE 1
Means of Variables Used in Analysis

Variable	Full Sample Mean	Full Sample Std. Dev. (continuous variables)	NLS-72 Sample Mean	NELS Sample Mean
Log of math test	2.53	0.59	2.44	2.65
Log of reading test	2.23	0.64	2.15	2.35
Female	0.51		0.50	0.52
Black	0.09		0.09	0.08
Hispanic	0.07		0.04	0.10
Protestant religion	0.44		0.42	0.46
Catholic religion	0.26		0.26	0.26
Jewish religion	0.02		0.02	0.01
No religious affiliation	0.07		0.05	0.09
Father has bachelor's	0.22		0.18	0.27
Mother has bachelor's	0.16		0.11	0.23
No father education info	0.12		0.11	0.13
No mother education info	0.11		0.10	0.11
Income over $10,000*	0.83		0.76	0.93
Income over $25,000*	0.64		0.58	0.72
Income over $35,000*	0.47		0.40	0.56
Income over $50,000*	0.26		0.20	0.35
No income data	0.12		0.19	0.04
Number of siblings	2.10	1.62	2.05	2.17
Has more than 6 siblings	0.06		0.06	0.06
Gets daily newspaper	0.83		0.88	0.76
Speaks English well	0.95		0.92	0.99
Math units in HS	2.44	1.18	1.91	3.19
Per pupil expenditure in 1972 ($1000)	3.01	0.93	3.04	2.97
Percent local revs in 1972	0.51	0.19	0.51	0.51
Per pupil exp/state average in 1972	1.09	0.25	1.09	1.08
Percent HS grads in community	0.58	0.17	0.49	0.74
Percent college grads in community	0.11	0.10	0.06	0.20
Percent of households below poverty line	0.13	0.09	0.12	0.13
Number of observations	14,726	14,726	8,672	6,054

*Income values are in 1990 dollars.

validity of results generated using the pooled cross section. Specifically, we estimated our basic specifications using the panel and including district-specific fixed effects. We then reestimated the model using the panel but treating the panel as a repeated cross section. In this reestimation, we replaced the fixed effects with the interaction of state-specific dummy variables with urbanicity dummy variables created using the seven urban statuses defined by the Census, our method of choice for capturing unobserved heterogeneity in our pooled cross-section estimates.[16] These two sets of results were substantively the same, indicating that these state-specific urbanicity measures effectively control for any unobserved, temporally stable effects that would necessitate the use of fixed effects.

5. RESULTS

Table 2 contains the results of our first approach for ascertaining the effects of tax limits and school finance reforms on student performance. In the first two columns, the constraints described above are imposed, and the remaining parameters are estimated using the observations in the NLS-72 sample. In the remaining two columns, the NELS sample is used to generate the results.

If the effects of the policy changes in question worked only through changes in per pupil expenditures and in the percentage of revenues raised locally, and if the changes in these variables were fully attributable to these policy changes, then the results in Table 2 support the conclusion that mean student performance in states in which a tax limit was imposed declined relative to mean student performance in states with no limit. The story for states with court-mandated or legislative finance reforms is similar; mean student test scores in reading and mathematics decline relative to these means in states without finance reforms. Of these types of reforms, legislative reforms appear to have slightly less

---

[16]These are large central city, mid-sized central city, suburb of large central city, suburb of mid-sized central city, large town, small town, and rural.

TABLE 2
Estimated Effects of School Finance Reforms and Tax Limits:
Suggestive Cross-Section Evidence Using Changes in Finance Variables

	Using NLS-72		Using NELS	
Policy Variable	Reading Test	Math Test	Reading Test	Math Test
Tax limit	-0.0076	-0.0084	-0.0104	-0.0058
	(0.0018)	(0.0013)	(0.0026)	(0.0036)
Court-mandated reform	-0.0081	-0.0104	-0.0125	-0.0122
	(0.0018)	(0.0013)	(0.0026)	(0.0036)
Legislative reform	-0.0039	-0.0043	-0.0054	-0.0032
	(0.0018)	(0.0013)	(0.0026)	(0.0036)
Anti-spending reform	-0.0126	-0.0162	-0.0193	-0.0192
(Hoxby 1996)	(0.0018)	(0.0013)	(0.0026)	(0.0036)
Pro-spending reform	0.0130	0.0138	0.0173	0.0079
(Hoxby 1996)	(0.0018)	(0.0013)	(0.0026)	(0.0036)

Note: This table presents the estimated percentage change associated with comparing 1990 values of per pupil expenditures and percent local revenues to the 1972 values. In the NELS specifications, the figures represent the estimated change in student performance using actual values of financial variables relative to using 1972 financial values. In the NLS-72 specifications, the figures represent the estimated change in student performance using 1990 values of financial variables relative to using actual financial values. Standard errors are in parentheses beneath the estimates.

deleterious effects. Finally, as expected in light of the analysis of Hoxby 1996, mean student performance exhibits relative decline in states that implemented finance reforms that can be termed anti-spending and relative improvement in states that implemented pro-spending reforms.

The results in Table 2 would warrant more thorough discussion if these results provided a conclusive picture of the effects of the policy changes in question. However, since implausible assumptions must be made if these results are to measure the true magnitude of the performance changes attributable to tax limits or finance reforms, we view these results as suggestive. Table 3 contains results we find more compelling. These results are generated using a standard event-analysis methodology. The effects of tax limits and finance reforms are assumed to be the same in all districts affected by these policy changes.

Whether observations from California are included or excluded,[17] the changes in the implied effects from Table 2 to Table 3 are striking. We begin our discussion of the specific results in Table 3 with the results that are generated when the observations from California are excluded.

For states that implemented a tax limit in the 1970s or 1980s, there was relative decline in the mean performance on the mathematics test on the order of 5.5 to 6 percent. Performance in reading exhibits no significant improvement or decline. Although the directions of these effects are basically the same as those in Table 2, for mathematics the magnitude of the effect is not the same. In fact, the relative decline in mathematics is very close to the estimated tax limit effect found by Figlio (1997a) in his cross-sectional analysis.

Although for tax limits the suggestive estimates in Table 2 are roughly similar to the estimated impacts in Table 3, that is not the case for school finance reforms. When observations from California are excluded, we find that, while the mean performances in reading and mathematics in states

[17]Several authors, notably Evans, Murray, and Schwab (1997), have argued that in California the effects have been atypical. In response to these arguments, we present all of the remaining results with observations for California included and excluded.

TABLE 3
Differences-in-Differences Results: Estimated Effect of School Finance Policies on Student Math and Science Performance

Policy Variable	Reading Test	Math Test	Reading Test	Math Test
I. Including California Observations				
Tax limit	-0.0005	-0.0574**	-0.0058	-0.0562**
	(0.0242)	(0.0204)	(0.0242)	(0.0210)
Court-mandated reform	0.0333	0.0936**		
	(0.0260)	(0.0212)		
Legislative reform	0.0178	0.0448**		
	(0.0215)	(0.0178)		
Anti-spending reform			0.0516	0.1534**
(Hoxby 1996)			(0.0315)	(0.0260)
Pro-spending reform			-0.0394	-0.0539**
(Hoxby 1996)			(0.0290)	(0.0251)
II. Excluding California Observations				
Tax limit	0.0003	-0.0554**	-0.0053	-0.0601**
	(0.0242)	(0.0204)	(0.0242)	(0.0208)
Court-mandated reform	0.0188	0.0335		
	(0.0292)	(0.0240)		
Legislative reform	0.0109	0.0141		
	(0.0222)	(0.0185)		
Anti-spending reform			0.0340	0.0454
(Hoxby 1996)			(0.0506)	(0.0439)
Pro-spending reform			-0.0373	-0.0512**
(Hoxby 1996)			(0.0291)	(0.0252)

Note: The regressions also include all variables reported in Table 1, as well as a time dummy and *state-specific* variables reflecting seven urban status possibilities (large central city, mid-sized central city, suburb of large central city, suburb of mid-sized central city, large town, small town, and rural, as defined by the Census Bureau).
Robust standard errors are in parentheses beneath parameter estimates.
Parameter estimates marked ** are significant at the 5 percent level; those marked * are significant at the 10 percent level.

implementing either court-mandated or legislative finance reforms improve relative to the mean performances on these tests in states with no finance reforms, there is no statistically significant effect in the case of either type of reform. Obviously, these results contrast sharply with the suggestive negative effects in Table 2. Further, these results are consistent with the argument that states have responded to court mandates by leveling-up spending (Evans, Murray, and Schwab 1995).

The contrasts between Table 2 and Table 3 are even more striking when we switch from classifying school finance reforms as either court-mandated or legislative to classifying reforms as either anti-spending or pro-spending. In states with anti-spending reforms, performance in reading and mathematics exhibit relative improvements, but these improvements are statistically insignificant. Nevertheless, these results suggest that, when state governments implement finance reforms that dramatically reduce the incentive for districts to raise revenue locally, other changes are made that increase statewide support for education and encourage more productive use of existing resources. We also find that, in states that implement reforms that increase the incentives for districts to raise revenue locally, there is relative decline in student performance in reading and mathematics and that the decline in mathematics is statistically significant.[18]

The results in the top half of Table 3 indicate that including those observations from California fails to alter any of our substantive conclusions, though some estimated effects that were previously insignificant become significant. These results support the conclusion that, all else equal, mean student performance in mathematics in California increased relative to the rest of the nation.[19] Nevertheless, we

[18]These results for pro- and anti-spending reforms are surprising, particularly in light of Hoxby's findings on the relationship between per pupil spending and changes in local incentives. However, the explanation for this seeming contradiction is not that our sample is unrepresentative. In particular, if we replicate Hoxby's analysis, using per pupil expenditure as our dependent variable and replacing her tax price variable with our pro- and anti-spending reform variables, we find that, as expected, the coefficient on the pro-spending reform variable is positive and significant, and the coefficient on the anti-spending reform variable is negative and significant.

[19]If no controls are included, we also find that mean mathematics performance in California improved relative to the rest of the nation.

do not feel that the differences between the top and the bottom of Table 3 are of sufficient magnitude necessarily to support the conclusion that California is unique.

The lingering question, then, is how to reconcile the disparate conclusions implied by Tables 2 and 3. The obvious answer is that school finance reforms and, to a lesser extent, tax limits result in changes that are not reflected solely in changes in financial variables like per pupil spending or the percentage of revenues raised locally. What these other changes are is not obvious. However, other research suggests possible explanations. In the case of tax limits, Figlio and O'Sullivan (1997) provide evidence of strategic behavior of municipalities and townships in response to the imposition of a limit. One goal of this behavior seems to be to encourage voters to override the limit. For school districts, one potential strategic response is to allow scores to decline by more than is dictated by changes in financial circumstances. Local governments may have similar incentives in the case of school finance reforms, particularly anti-spending reforms, if there is the chance that the state government could "bail out" the school district. A second potential explanation for the larger effects in Table 3 is offered by Figlio and Rueben (1997), who find that tax limits may deter individuals with more ability from entering teaching to a greater extent than any changes in starting salaries or other circumstances would lead one to expect.

As for the changes between Tables 2 and 3 in the estimated effects of school finance reforms, Downes (1992) argues that one potential explanation for the failure of finance reforms in California to alter the dispersion across districts in student performance is that the high-wealth districts most affected by the reforms have found methods of circumventing the limits imposed. Brunner and Sonstelie (1996) provide evidence on the magnitude of one such response. A second explanation for the changes is suggested by Downes and Schoeman (1997) and Kenny and Husted (1996), who find that the share of students attending private school increases in the aftermath of court-mandated finance reforms. If the students departing for private schools are of slightly below-average ability, then this movement could explain some of the difference between Tables 2 and 3.

As was noted above, the effects of tax limits and finance reforms may vary across districts within a state. Tables 4A and 4B contain results generated when the observations from California are excluded from the analysis and when the effects of the policy changes are allowed to vary with the initial financial conditions of each district affected. Similar results are given in Tables 4C and 4D for the case when the observations from California are included. Finally, in Table 5 we present the implied policy effects when the estimates in these four tables are used to determine the mathematics performance of a student in a state without a policy change relative to a student in a district with the mean financial characteristics, in a district with the relevant financial characteristic one standard deviation below the mean, and with the relevant financial characteristic one standard deviation above the mean.[20] As is apparent from Tables 4C and 4D, the results with California observations included are qualitatively similar to those with California excluded. Thus, the discussion that follows refers exclusively to Tables 4A and 4B and to the accompanying results in Table 5.

The estimates in Table 4A indicate that mean test performance in states with court-mandated school finance reforms is relatively larger by a statistically insignificant amount and that, within a state with this type of reform, the impact of that reform is independent of the initial financial conditions of the affected district. This latter result, which is confirmed in Table 5, matches at the national level Downes' 1992 findings for California. In addition, the positive, though insignificant, effect of court-mandated reforms on student performance in all districts is consistent with the leveling-up of spending observed by Evans, Murray, and Schwab (1995).

A comparison of Tables 3 and 4A reveals that the effects of legislative reforms on student performance are obscured when we fail to allow for the possibility that the effects of such reforms depend on each school district's initial financial conditions. The results in Table 5 indicate that, in states

---

[20]Since, for reading test performance, the estimated policy effects are generally not statistically significant, in Table 5 we report only results for mathematics test performance.

TABLE 4A
Differences-in-Differences-in-Differences Results: Heterogeneity in the Estimated Effect of School Finance Policies on Student Math and Science Performance (excluding California observations)

Policy Variable	Reading Test	Math Test	Reading Test	Math Test
Tax limit	-0.0948 (0.0741)	-0.1580** (0.0550)	-0.0128 (0.0588)	-0.0557 (0.0486)
Court-mandated reform	0.1033 (0.0802)	0.0573 (0.0578)	0.0625 (0.0733)	-0.0143 (0.0574)
Legislative reform	0.0453 (0.0633)	0.1851** (0.0460)	0.1214** (0.0510)	0.2026** (0.0411)
Tax limit * per pupil expenditures in 1972 (/1000)	0.0333 (0.0241)	0.0342** (0.0168)		
Court reform * per pupil expenditures in 1972 (/1000)	-0.0301 (0.0261)	-0.0077 (0.0181)		
Legislative reform * per pupil expenditures in 1972 (/1000)	-0.0111 (0.0202)	-0.0591** (0.0140)		
Tax limit * percent local in 1972			0.0288 (0.0999)	0.0021 (0.0766)
Court reform * percent local in 1972			-0.0948 (0.1329)	0.0952 (0.0978)
Legislative reform * percent local in 1972			-0.2162** (0.0853)	-0.3718** (0.0656)

Note: The regressions also include all variables reported in Table 1, as well as a time dummy and *state-specific* variables reflecting seven urban status possibilities (large central city, mid-sized central city, suburb of large central city, suburb of mid-sized central city, large town, small town, and rural, as defined by the Census Bureau).
Robust standard errors are in parentheses beneath parameter estimates.
Parameter estimates marked ** are significant at the 5 percent level; those marked * are significant at the 10 percent level.

TABLE 4B
Differences-in-Differences-in-Differences Results: Heterogeneity in the Estimated Effect of School Finance Policies on Student Math and Science Performance (excluding California observations)

Policy Variable	Reading Test	Math Test	Reading Test	Math Test
Tax limit	-0.0411 (0.0560)	-0.0822* (0.0443)	0.0443 (0.0490)	0.0046 (0.0409)
Anti-spending reform	-0.0020 (0.1188)	0.1260 (0.1000)	0.0795 (0.1081)	0.1515* (0.0781)
Pro-spending reform	-0.0314 (0.0945)	-0.0463 (0.0748)	-0.1064 (0.0694)	-0.1320** (0.0574)
Tax limit * per pupil expenditures in 1972 (/1000)	0.0121 (0.0169)	0.0079 (0.0130)		
Anti-spending reform * per pupil expenditures in 1972 (/1000)	0.0107 (0.0325)	-0.0274 (0.0254)		
Pro-spending reform * per pupil expenditures in 1972 (/1000)	-0.0026 (0.0223)	-0.0018 (0.0170)		
Tax limit * percent local in 1972			-0.1019 (0.0757)	-0.1319** (0.0581)
Anti-spending reform * percent local in 1972			-0.1989 (0.2697)	-0.4118** (0.1925)
Pro-spending reform *percent local in 1972			0.1402 (0.1194)	0.1648* (0.0935)

Note: The regressions also include all variables reported in Table 1, as well as a time dummy and *state-specific* variables reflecting seven urban status possibilities (large central city, mid-sized central city, suburb of large central city, suburb of mid-sized central city, large town, small town, and rural, as defined by the Census Bureau). Pro-spending and anti-spending school finance reforms are categorized by Hoxby (1996). Robust standard errors are in parentheses beneath parameter estimates.
Parameter estimates marked ** are significant at the 5 percent level; those marked * are significant at the 10 percent level.

TABLE 4C
Differences-in-Differences-in-Differences Results: Heterogeneity in the Estimated Effect of School Finance Policies on Student Math and Science Performance (including California observations)

Policy Variable	Reading Test	Math Test	Reading Test	Math Test
Tax limit	-0.0805 (0.0734)	-0.1351** (0.0546)	-0.0022 (0.0587)	-0.0414 (0.0485)
Court-mandated reform	0.0866 (0.0749)	0.0762 (0.0545)	0.0183 (0.0690)	-0.0568 (0.0540)
Legislative reform	0.0386 (0.0615)	0.2056** (0.0450)	0.1063** (0.0499)	0.1962** (0.0403)
Tax limit * per pupil expenditures in 1972 (/1000)	0.0280 (0.0239)	0.0254 (0.0167)		
Court reform * per pupil expenditures in 1972 (/1000)	-0.0190 (0.0239)	0.0076 (0.0168)		
Legislative reform * per pupil expenditures in 1972 (/1000)	-0.0065 (0.0195)	-0.0551** (0.0136)		
Tax limit * percent local in 1972			0.0039 (0.0996)	-0.0339 (0.0764)
Court reform *percent local in 1972			0.0363 (0.1200)	0.3027** (0.0889)
Legislative reform * percent local in 1972			-0.1703** (0.0840)	-0.3021** (0.0889)

Note: The regressions also include all variables reported in Table 1, as well as a time dummy and *state-specific* variables reflecting seven urban status possibilities (large central city, mid-sized central city, suburb of large central city, suburb of mid-sized central city, large town, small town, and rural, as defined by the Census Bureau).
Robust standard errors are in parentheses beneath parameter estimates.
Parameter estimates marked ** are significant at the 5 percent level; those marked * are significant at the 10 percent level.

TABLE 4D
Differences-in-Differences-in-Differences Results: Heterogeneity in the Estimated Effect of School Finance Policies on Student Math and Science Performance (including California observations)

Policy Variable	Reading Test	Math Test	Reading Test	Math Test
Tax limit	-0.0300 (0.0510)	-0.0461 (0.0438)	0.0443 (0.0488)	0.0081 (0.0408)
Anti-spending reform	0.0089 (0.0864)	0.2306** (0.0661)	-0.0131 (0.0825)	0.0058 (0.0632)
Pro-spending reform	-0.0178 (0.0941)	-0.0165 (0.0744)	-0.0987 (0.0690)	-0.1263** (0.0572)
Tax limit * per pupil expenditures in 1972 (/1000)	0.0081 (0.0167)	-0.0029 (0.0127)		
Anti-spending reform * per pupil expenditures in 1972 (/1000)	0.0119 (0.0248)	-0.0237 (0.0182)		
Pro-spending reform * per pupil expenditures in 1972 (/1000)	-0.0062 (0.0221)	-0.0091 (0.0170)		
Tax limit * percent local in 1972			-0.1009 (0.0752)	-0.1308** (0.0576)
Anti-spending reform * percent local in 1972			0.1289 (0.1420)	0.2872** (0.1046)
Pro-spending reform *percent local in 1972			0.1222 (0.1184)	0.1503 (0.0931)

Note: The regressions also include all variables reported in Table 1, as well as a time dummy and *state-specific* variables reflecting seven urban status possibilities (large central city, mid-sized central city, suburb of large central city, suburb of mid-sized central city, large town, small town, and rural, as defined by the Census Bureau). Pro-spending and anti-spending school finance reforms are categorized by Hoxby (1996).
Robust standard errors are in parentheses beneath parameter estimates.
Parameter estimates marked ** are significant at the 5 percent level; those marked * are significant at the 10 percent level.

TABLE 5
Predicted Effects of Tax Limits and School Finance Reforms for Districts with Different Initial Levels of Financial Variables Results Using Mathematics Test (samples including and excluding California)

	Mean Effect		+1 Std. Dev.		-1 Std. Dev.	
Policy Variable	With CA	Without CA	With CA	Without CA	With CA	Without CA
I. Expenditure-Based Specification						
Tax limit	-0.0600	-0.0569	-0.0378	-0.0270	-0.0823	-0.0869
Court reform	0.0991	0.0341	0.1066	0.0265	0.0916	0.0417
Legislative reform	0.0454	0.0132	-0.0016	-0.0372	0.0924	0.0637
Anti-spending reform	0.1518	0.0348	0.1292	0.0088	0.1743	0.0609
Pro-spending reform	-0.0536	-0.0536	-0.0634	-0.0556	-0.0437	-0.0517
II. Percent Local-Based Specification						
Tax limit	-0.0601	-0.0546	-0.0555	-0.0542	-0.0536	-0.0549
Court reform	0.1097	0.0381	0.1642	0.0552	0.0552	0.0209
Legislative reform	0.0331	0.0018	-0.0243	-0.0689	0.0905	0.0725
Anti-spending reform	0.1724	-0.0873	0.2212	-0.1574	0.1236	-0.0173
Pro-spending reform	-0.0361	-0.0331	-0.0046	0.0015	-0.0677	-0.0677

Note: In the case of tax limits, we report the estimated effects using the specifications with court-ordered and legislative reforms.

in which a legislative reform has been implemented, students in schools in districts with the mean initial financial characteristics have mathematics test scores 3.5 to 4.5 percent higher, relative to students in states with no finance reform. Further, the results in Table 5 support the conclusion that legislative reforms appear to reduce the dispersion in student performance, a result that contrasts with the findings of Evans, Murray, and Schwab (1997) on the impact of such reforms of the distribution of spending. As expected, the effect of such reforms on test scores is larger the lower a district's initial spending level and the lower the initial share of a district's revenues raised locally.

Tax limits have smaller effects in districts with larger initial spending levels, as expected. However, as is apparent from Table 5, the extent of this heterogeneity is relatively small. The bottom line is that mathematics test scores of students in all districts subject to a tax limit decline relatively.

From Table 4B, we can see that the effects of anti-spending and pro-spending finance reforms do not vary with the initial spending levels of districts in the states that implement these reforms. We can also see that the effects of reforms of these types do vary with the initial share of revenues raised locally. Taken as a whole, the results in Table 5 appear to indicate that anti-spending reforms actually reduce the dispersion in student performance. In this sense, anti-spending reforms more closely resemble legislative reforms than they do court-ordered reforms. On the other hand, in states that implement pro-spending reforms, we find students in districts that raise larger-than-average fractions of their revenue locally can actually benefit, even though there is a relative decline in mean performance in the state. This result is consistent with the fact that, in general, districts that already raise much of their revenue locally tend to be favored by finance reforms that increase the incentives to raise revenue locally.

The final set of policy effects, given in Table 6, result when we relax the assumption that the implementation of these policies was exogenous. To do so, we adopt the two-stage dummy endogenous-variable technique suggested by Heckman (1979). To predict whether states adopted particular policies during the period between the NLS-72 and NELS samples, we build on the logic above by using a set of

TABLE 6
Differences-in-Differences Results: Estimated Effect of School Finance Policies on Student Math and Science Performance When Policy Variables Are Modeled as Endogenous

Policy Variable	Reading Test	Math Test	Reading Test	Math Test
I. Including California Observations				
Tax limit	-0.0165	-0.0531*	0.0056	-0.0476**
	(0.0439)	(0.0343)	(0.0366)	(0.0204)
Court-mandated reform	0.0574	0.1296**		
	(0.0353)	(0.0272)		
Legislative reform	-0.0050	0.1110**		
	(0.0369)	(0.0309)		
Anti-spending reform			0.0646*	0.1748**
(Hoxby 1996)			(0.0385)	(0.0315)
Pro-spending reform			-0.0460	-0.0696**
(Hoxby 1996)			(0.0395)	(0.0328)
II. Excluding California Observations				
Tax limit	-0.0225	-0.0720**	-0.0004	-0.0835**
	(0.0444)	(0.0348)	(0.0388)	(0.0318)
Court-mandated reform	0.0350	0.0446		
	(0.0393)	(0.0308)		
Legislative reform	-0.0066	0.1065**		
	(0.0369)	(0.0308)		
Anti-spending reform			0.0323	0.0298
(Hoxby 1996)			(0.0602)	(0.0498)
Pro-spending reform			-0.0331	-0.0359
(Hoxby 1996)			(0.0408)	(0.0340)

Note: The regressions also include all variables reported in Table 1, as well as a time dummy and *state-specific* variables reflecting seven urban status possibilities (large central city, mid-sized central city, suburb of large central city, suburb of mid-sized central city, large town, small town, and rural, as defined by the Census Bureau).
Robust standard errors are in parentheses beneath parameter estimates.
Parameter estimates marked ** are significant at the 5 percent level; those marked * are significant at the 10 percent level.

instruments that combines information on constitutional language with variables reflecting each state's demographic and fiscal conditions in 1970. Specifically, we constructed a dummy variable indicating whether the state's constitution permitted direct voter initiatives (Schmidt 1989) and three dummy variables reflecting whether the education clause in the state's constitution requires "equality," "uniformity," or "efficiency" (McUsic 1991). The full set of instruments included these dummy variables, regional dummies, two dichotomous variables reflecting whether in 1960 or 1970 the state's governor was a Democrat, the state's population in 1970, the fraction of the 1970 population that was school age, the fraction of the 1970 population that was over 65, the state's tax effort in 1967, the change in the state's tax effort from 1967 to 1972, and the interaction of this last variable with the initiative dummy. Our results are qualitatively insensitive to the choice of instruments used; we select this set of variables in an attempt to get the best possible prediction of the adoption of these policies.

These variables have substantial explanatory power in discerning among states that did or did not impose policies during the relevant time period. For instance, while the mean estimated probability that a state would impose a tax limitation is 77 percent for states that eventually imposed tax limits, this mean estimated probability is 16 percent for states that did not pass a tax limit during the relevant period. These differences are similar for the other policy variables: while the mean estimated probability that a state would impose a court-ordered school finance reform is 74 percent for states that eventually imposed one, this mean estimated probability is only 6 percent for states that did not impose a court-ordered reform during the relevant period. For legislative reforms, the relevant numbers are 70 and 29 percent, respectively. All differences are statistically significant at the 1 percent level.

Comparing the results in Tables 3 and 6 reveals that estimates of the mean effects of these policy changes are altered little when the policies are treated as endogenous. In fact, all of the conclusions concerning the impact of the policies are robust to treating the policies as endogenous.[21]

6. CONCLUSION

During the late 1970s and early 1980s, a majority of states substantially changed the ways in which schools were funded, either directly through court- or legislatively mandated school finance reform, or indirectly through tax and expenditure limits. To date, very few academics have attempted to gauge the effects of these policy changes on actual outcomes of education, such as student academic achievement. Given the recent resurgence of both types of policy changes, be they equalization court cases in states such as New Jersey, Ohio, or Vermont, or new tax and expenditure limits in states such as California, Oregon, and Illinois, the need to evaluate the long-term effects of these types of policy changes has become more immediate.

This paper is an attempt to fill this gap in the literature. We use detailed, individual-level data from two cohorts of high school seniors, one that attended school in an era before any recent substantial changes in school finance policy occurred, and another that spent most (or all) of their academic careers in schools subject to tax limits or school finance reforms. Because of the large number of schools in each cohort, and because we have the opportunity to observe arguably observationally equivalent students before and after these school finance policy changes occurred, we have a unique opportunity to conduct a natural experiment concerning the effects of these policy changes.

---

[21] Versions of Tables 4A through 4D in which the instrumental variable estimates are provided are available from the authors.

We find economically and statistically significant evidence that tax limits have led to decreased mathematics performance, but not reading performance. Moreover, there is some evidence that the effects of tax limits have been borne principally by initially low-spending school districts. Within any state and urban status, initially low-spending districts tend to have low family incomes and high minority percentages. Thus, if anything, tax limits may have led not only to a "leveling down" of student performance in mathematics, but also to increased inequality in educational opportunity and outcomes. These tax limits results are robust to inclusion or exclusion of California observations from the estimation, strongly suggesting that we are not merely finding a "California effect."

We find, however, that school finance reforms may have led to the opposite outcomes. Our evidence suggests that court-mandated and legislatively mandated school finance reforms have led, on average, to *increased* student performance. Legislative reforms apparently increase average student achievement in mathematics (and perhaps, even, in reading), and, unlike tax limits, appear to be redistributive. Students in initially low-spending districts and those that, in the early 1970s, relied the least on local revenue sources apparently benefitted the most from legislative reforms. For students in initially high-spending districts, legislative reforms may have led to reduced achievement levels. Unlike our results for tax limits, however, these results are sensitive to the inclusion or exclusion of California. Generally, the estimated effects of school finance reforms are greater when California is part of our sample.

This considerable heterogeneity in the estimated effects of school finance reforms, even within a state, is masked by researchers who act as if school finance reforms affect all school districts in a state equally. Of the types of school finance reforms considered in this paper, only court-mandated finance reforms appear to have relatively uniform effects across districts. It is clear from intuition, and bolstered by the theoretical literature, that school finance reforms should differentially affect school districts in a state, and our empirical evidence bears this prediction out.

Hoxby (1996) also makes the point that the effects of reforms on school districts in a state may be heterogeneous, but argues that this heterogeneity comes through differences in the spending incentives provided to local school districts. While this is a compelling argument, we are not comfortable accepting that the only interesting aspect of school finance reforms is the effect of changing marginal tax prices. Our results provide suggestive evidence that while marginal tax prices surely matter, other aspects of school finance reforms—even in our crude categorizations—matter too. Indeed, if we use Hoxby's (1996) rough classification of school finance reforms based on their effects on the marginal tax prices faced by the districts, we find (especially in California) evidence that anti-spending reforms may be associated with *increased* achievement, and pro-spending reforms may be associated with *decreased* achievement. Although in our sample, local spending went down in anti-spending reform states and up in pro-spending reform states, *state* spending moved in the opposite direction. Nominal growth in state spending per pupil in anti-spending states was more than double that in other states, while nominal growth in state spending per pupil in pro-spending states was only 44 percent of that in other states. Thus, while we agree that the effects of local spending incentives are important, we also argue that the role of direct state spending in school finance reforms should not be ignored.

In all of our sensitivity tests to date, the basic results presented above appear indelible. Tax limits apparently reduce student achievement, and school finance reforms, if anything, level up student performance. If these results continue to hold up to further testing, they present substantial implications for states currently faced with the multiple dilemmas of desiring tax relief while protecting or further equalizing educational opportunities.

References

Advisory Commission on Intergovernmental Relations and Center for Urban and Environmental Policy at Indiana University. 1995. *Tax and Expenditure Limits on Local Governments.* Washington, D.C. March.

Bénabou, Roland. 1996. “Equity and Efficiency in Human Capital Investment: The Local Connection.” *Review of Economic Studies* 63 (April): 237–264.

Betts, Julian R. 1996. “Is There a Link between School Inputs and Earnings? Fresh Evidence of an Old Literature.” In *Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success,* ed. Gary Burtless. Washington, D.C.: The Brookings Institution.

Brunner, Eric, and Jon Sonstelie. 1996. “Coping with *Serrano*: Voluntary Contributions to California’s Public Schools.” Santa Barbara: University of California at Santa Barbara. October. Mimeo.

Courant, Paul N., and Susanna Loeb. 1997. “Centralization of School Finance in Michigan.” *Journal of Policy Analysis and Management* 16 (Winter): 114–136.

Dewey, James, Thomas A. Husted, and Lawrence W. Kenny. 1997. “Are Educational Inputs Irrelevant?: A Reexamination of the Evidence.” Gainesville, FL: University of Florida. Mimeo.

Downes, Thomas A. 1992. “Evaluating the Impact of School Finance Reform on the Provision of Public Education: The California Case.” *National Tax Journal* 45 (December): 405–419.

Downes, Thomas A. 1996. “An Examination of the Structure of Governance in California School Districts Before and After Proposition 13.” *Public Choice* 86 (March): 279–307.

Downes, Thomas A., Richard F. Dye, and Therese J. McGuire. Forthcoming 1997. “Do Limits Matter? Evidence on the Effects of Tax Limitations on Student Performance.” *Journal of Urban Economics*.

Downes, Thomas A., and Shane M. Greenstein. 1996. “Understanding the Supply of Non-Profits: Modeling the Location of Private Schools.” *RAND Journal of Economics* 27 (Summer): 365–390.

Downes, Thomas A., and David Schoeman. Forthcoming 1997. “School Financing Reform and Private School Enrollment: Evidence from California.” *Journal of Urban Economics*.

Downes, Thomas A., and Mona Shah. 1996. “The Effect of School Finance Reform on the Level and Growth of Per Pupil Expenditures.” Medford, MA: Tufts University Working Paper No. 95-4, June.

Dye, Richard, and Therese McGuire. Forthcoming 1997. “The Effect of Property Tax Limitation Measures on Local Government Fiscal Behavior.” *Journal of Public Economics*.

Ehrenberg, Ronald, and Dominic Brewer. 1994. "Do School and Teacher Characteristics Matter? Evidence from *High School and Beyond*." *Economics of Education Review* 13: 1–17.

Evans, William N., Sheila Murray, and Robert M. Schwab. 1995. "Education Finance Reform and the Distribution of Resources." College Park, MD: University of Maryland, College Park. July. Mimeo.

Evans, William N., Sheila Murray, and Robert M. Schwab. 1997. "Schoolhouses, Courthouses, and Statehouses after *Serrano*." *Journal of Policy Analysis and Management* 16 (Winter): 10–31.

Ferguson, Ronald F. 1991. "Paying for Public Education: New Evidence on How and Why Money Matters." *Harvard Journal on Legislation* 28 (Summer): 465–498.

Ferguson, Ronald F., and Helen F. Ladd. 1996. "How and Why Money Matters: An Analysis of Alabama Schools." In *Holding Schools Accountable: Performance-Based Reform in Education*, ed. Helen F. Ladd. Washington, D.C.: The Brookings Institution.

Fernandez, Raquel, and Richard Rogerson. 1997. "Education Finance Reform: A Dynamic Perspective." *Journal of Policy Analysis and Management* 16 (Winter): 67–84.

Figlio, David N. 1996. "Does School Quality Matter? More Than We Thought But Less Than We Hoped." Eugene, OR: University of Oregon. November. Mimeo.

Figlio, David N. Forthcoming 1997a. "Did the 'Tax Revolt' Reduce School Performance?" *Journal of Public Economics.*

Figlio, David N. Forthcoming 1997b. "Short-Term Effects of a 1990s-Era Tax Limit: Panel Evidence on Oregon's Measure 5." *National Tax Journal*.

Figlio, David N., and Arthur O'Sullivan. 1997. "Do Local Governments Respond Strategically to Tax Limits." Eugene, OR: University of Oregon, February. Mimeo.

Figlio, David N., and Kim Rueben. 1997. "How Tax Limits Affect Teacher Quality." Eugene, OR: University of Oregon, February. Mimeo.

Fischel, William A. 1996. "*Serrano* after Twenty-Five Years: Are America's Schools Better and Property Taxes Fairer?" Hanover, NH: Dartmouth College. December. Mimeo.

Fisher, Ronald C. 1996. *State and Local Public Finance,* 2nd ed. Chicago: Richard D. Irwin.

Hanushek, Eric A. 1986. "The Economics of Schooling : Production and Efficiency in the Public Schools." *Journal of Economic Literature* 24 (September): 1141–1177.

Hanushek, Eric A. 1996. "School Resources and Student Performance." In *Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success,* ed. Gary Burtless. Washington, D.C.: The Brookings Institution.

Heckman, James J. 1979. “Sample Selection Bias as a Specification Error.” *Econometrica* 47 (January): 153–161.

Hedges, Larry V., and Rob Greenwald. 1996. “Have Times Changed? The Relationship between School Resources and Student Performance.” In *Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success*, ed. Gary Burtless. Washington, D.C.: The Brookings Institution.

Hoxby, Caroline M. 1996. “All School Finance Equalizations Are Not Created Equal: Marginal Tax Rates Matter.” Cambridge, MA: Harvard University. March. Mimeo.

Husted, Thomas A., and Lawrence W. Kenny. 1996. “Evidence from the States on the Political and Market Determinants of Efficiency in Education.” Washington, D.C.: American University. October. Mimeo.

Kenny, Lawrence W., and Thomas A. Husted. 1996. “The Legacy of *Serrano*: The Impact of Mandated Equal Spending on Private School Enrollment.” Gainesville, FL: University of Florida. August. Mimeo.

Krueger, Alan B. 1997. “Experimental and Nonexperimental Estimates of Education Production Functions.” Princeton, NJ: Princeton University. Mimeo.

Leyden, Dennis P. 1992. “Court-Mandated Changes in Educational Grant Structure.” *Public Finance* 47: 229–247.

Levine, Phillip B., and David J. Zimmerman. 1995. “The Benefit of Additional High-School Math and Science Classes for Young Men and Women.” *Journal of Business and Economic Statistics* 13 (April): 137–149.

Manwaring, Robert L., and Steven M. Sheffrin. 1995. “The Effects of Education Equalization Litigation on the Levels of Funding: An Empirical Analysis.” Davis, CA: Department of Economics, University of California-Davis. June. Mimeo.

McUsic, Molly. 1991. “The Use of Education Clauses in School Finance Legislation.” *Harvard Journal on Legislation* 28 (Summer): 307–339.

Nechyba, Thomas J. 1996. “Public School Finance in a General Equilibrium Tiebout World: Equalization Programs, Peer Effects, and Private School Vouchers.” Cambridge: NBER Working Paper No. 5642. June.

New York Times. 1996. “California Spends Tax Windfall on Smaller Class Sizes.” September 11, p. D18.

Rueben, Kim. 1995. “Tax Limitations and Government Growth: The Effect of State Tax and Expenditure Limitations on State and Local Government.” Cambridge: MIT. October. Mimeo.

Schmidt, David D. 1989. *Citizen Lawmakers: The Ballot Initiative Revolution.* Philadelphia: Temple University Press.

Silva, Fabio, and Jon Sonstelie. 1995. "Did *Serrano* Cause a Decline in School Spending?" *National Tax Journal* 48 (June): 199–215.