

IMPACT OF SCHOOL FINANCE REFORM ON RESOURCE EQUALIZATION AND ACADEMIC PERFORMANCE: EVIDENCE FROM MICHIGAN

Joydeep Roy

IBO and
Columbia University
New York, NY 10027
jr3137@columbia.edu

Abstract

Michigan radically altered its school finance system in 1994. The new plan, called Proposal A, significantly increased state aid to the lowest-spending school districts and limited future increases in spending in the highest-spending ones, abolishing local discretion over school spending. I investigate the impact of Proposal A on the distribution of resources and educational outcomes. I analyze the differential effects on the lowest-spending and the highest-spending districts, highlighting the role of local discretion, which has so far been neglected in the literature. I also provide important evidence on the effect of spending on academic performance. Proposal A was quite successful in reducing interdistrict spending disparities. There was also a significant positive effect on student performance in the lowest-spending districts as measured in state tests. However, the constraints on future increases in spending may have had a negative effect on student performance in the highest-spending districts.

1. INTRODUCTION

In 1971 the California Supreme Court asked the state legislature to overhaul its public school finance system, arguing that the education of a child should not be a function of the wealth of its neighbors. Since then, school finance reforms have been debated and instituted in many states throughout the country. In 1994, Michigan radically altered its school financing rules, and the Michigan school finance reform, known as Proposal A, was enacted. In this article I investigate Proposal A's effects on the distribution of resources and academic performance in Michigan. Michigan provides a particularly interesting case because there were significant increases in per pupil spending following the reform, and local discretion over school spending was largely abolished—though there were no adverse court rulings providing the impetus—making Proposal A one of the more unique school finance reforms.

In most U.S. states, local property taxes are a major source of school revenues, and since people often sort themselves among neighborhoods based on their incomes, taxable property wealth differs substantially from one school district to another, even within the same state.¹ School finance reforms are aimed at weakening this nexus between school district wealth and per pupil expenditures. They typically achieve this by large increases in state aid to poorer districts, often coupled with restrictions on spending in the richer ones. The rationale behind such an overhaul is that children in poorer districts may be lagging behind others because of inadequate resources at their disposal. However, critics of such reforms argue that the way these are implemented—giving large windfalls of money to poorer districts—makes it highly unlikely that they will lead to any meaningful improvement for the students concerned (Hanushek 1991).

Michigan's comprehensive overhaul of its school finance program, Proposal A, significantly increased the state share of K–12 spending in all Michigan districts. It also entailed giving large sums of money to the lowest-spending districts, which were allowed to increase their spending at a much faster rate than others. Concurrently, Proposal A also ended local discretion over school spending. Based on a formula, the state now decides the amounts by which each school district can raise its expenditures.

I find that Proposal A was quite successful as far as equalization of school spending is concerned. By the end of the decade the lowest-spending districts had witnessed large increases in per pupil spending. The gap between the highest-spending and the lowest-spending districts in the state, large and widening before the reform, had significantly narrowed. I also look at the program's effect on academic performance. I employ various strategies,

1. See table 1 in Fernandez and Rogerson (1996) for an illuminating example of inequalities in per pupil spending across school districts in nearby communities.

including using the changes in state aid formula as instruments for actual spending, to estimate whether the lowest-spending districts, the chief beneficiaries of this reform, witnessed any additional improvements. The results based on tests administered by the state show significant gains in academic performance by these districts, and these gains are robust to several sensitivity checks. However, I also look at other measures of performance, including the college preparatory test ACT.² There is not much evidence for any improved performance by these lowest-spending districts either in rates of taking the ACT or in ACT performance, highlighting the fact that achievement benefits on one outcome measure do not always generalize to other measures.

As mentioned, one of the important effects of the reform was the loss of local discretion over per pupil spending at the school district level. This may have a significant effect in the highest-spending districts, which were constrained by the limits put on further increases in spending. I test whether the reform differentially affected student performance in these districts, which were also the highest-performing districts at the time of the reform. The results suggest that these constraints negatively affected student achievement in these high-spending districts.

Among the empirical studies of school finance reforms in the United States, Murray, Evans, and Schwab (1998) conclude that court-mandated finance reforms have had a large positive effect on equalization of school resources. Card and Payne (2002) find that such reforms also led to a convergence in SAT scores across family background groups. The articles that analyze individual school finance reforms mostly concern states where the financing system was changed following directives from the courts.³ In Michigan, on the other hand, the campaign for reform was led by the legislative and executive branches.⁴ There is an as yet unresolved issue in the school finance literature about whether legislature-mandated changes can have real effects. Evans, Murray, and Schwab (1997, p. 10), based on their analysis of sixteen thousand public school districts from 1972 to 1992, conclude that “reforms that

-
2. The ACT assessment, known as the American College Testing Program until 1996, is the more widely accepted college entrance examination in the United States. The ACT is particularly important in the midwestern and southern states, where it is often required for admission to colleges.
 3. Downes (1992) finds that in California there has been a significant convergence across school districts in per pupil expenditures, but it has not translated into academic performance. Guryan (2001), who studies the Massachusetts reform of 1993, finds that the increase in spending due to an increase in state aid improved fourth-grade test scores. Clark (2003) finds that the Kentucky Education Reform Act of 1990 spurred significant increases in spending in the lowest-spending districts but still failed to produce gains in ACT performance. Yinger (2004) includes case studies of several important school finance reforms, but most of these studies focus on the equalization of per pupil spending and abstract from the effect on academic performance.
 4. There were two court cases, *Milliken v. Green* in 1973 and *East Jackson Public Schools v. Michigan* in 1984. Both found the existing finance system constitutional.

were initiated by the states without judicial prodding were typically ineffective.” Card and Payne (2002), on the other hand, argue that even in these states, state aid is now becoming more targeted toward the lower-income districts. One of the important objectives of this study is to provide some evidence on the efficacy of legislature-induced school finance reforms.⁵

This article is most closely related to Cullen and Loeb (2004) and Papke (2005, 2008). Cullen and Loeb provide an excellent description of the Michigan program and discuss some of its effects on per pupil spending and academic outcomes. Papke (2005) uses data on standardized test scores in Michigan from 1992 to 1998 to determine the effects of spending on academic performance. She finds that the increases in spending had significant effects on mathematics pass rates and that the effects were largest for schools that were initially lagging behind. Papke (2008) further investigates this issue, using additional years of data, and she finds similar results as in the earlier study.

The present study differs from these studies in some fundamental ways. First, the questions posed here are somewhat different. While Papke is interested in the effect of money on test scores, my focus is on the broader school finance reform. In addition to investigating the effect of the reform on academic achievement, I rigorously analyze whether the reform actually led to convergence of revenues and expenditures and whether it weakened the link between income and spending. Further, I analyze in detail the heterogeneity of effects across the high-spending and low-spending districts and highlight the crucial role of local discretion over school spending, an issue that is of significant importance. Also, unlike the above studies, a motivation for the current article is to study the effectiveness of legislature-mandated school finance reforms, which is why I chose the Michigan school finance reform.

Methodologically, the present article also differs from the earlier studies in several important ways, including discussing issues like the spread of charter schools and public school choice, which might bias the estimates, and employing a cohort-based measure of tracking improvements in academic performance, in addition to looking at repeated cross sections (panel data on school districts). I also employ a much longer time series of data, from 1990 to 2001, which allows me to not only control for any differences in pre-reform trends but also to capture program effects that occur only with a significant lag.⁶ Further, to check whether the academic gains of students in

5. Both the Evans, Murray, and Schwab and the Card and Payne articles use data only up to 1992, so they exclude the Michigan reform, which was initiated in 1994.

6. I do not use data from 2002 and beyond, since any true effect of the reform is likely to be confounded with effects due to the introduction of the federal No Child Left Behind Act (NCLB). Papke (2008) uses data up to 2003–4, but using data from the post-2001 period is problematic for the above reason. This is particularly true for Michigan, which had a large number of failing schools under NCLB in the early years. For example, in 2001–2 Michigan, with rigorous standards, had 1,513 failing schools, the most in the nation (Schemo 2002).

the lowest-spending districts have also been translated into other achievement outcomes, this study looks at changes in college prep test performance (ACT participation and test scores) across districts. Finally, an important component of my study is exploring and analyzing possible heterogeneity in effects between districts at different points of the spending distribution, which were differentially affected by the school finance reform. Papke (2008) also looks for heterogeneity in effects on spending and test scores. However, her grouping of districts is very broad, dividing all districts into only two groups, and hence is less likely to reveal all nuances in effects across districts, particularly those in the top and bottom quintiles. Further, the classification that Papke (2008) employs is not uniform. For spending she classifies districts based on spending in 1991–92 (even though the last year before Proposal A was 1993–94), while for test scores she classifies districts based on average pass rates in the first three years in her sample (1991–92 to 1993–94). This is problematic, more so as the school finance reform targeted low-spending districts instead of low-performing districts, and the two groups did not always overlap completely. Moreover, the use of test scores averaged over three years precludes controlling for pre-reform trends, which may significantly bias the results.

The remainder of the article is organized as follows. In section 2 I outline the main features of the Michigan reform. Section 3 discusses the various data sources used. In section 4 I examine the effects of the program on equalization of school finances. Section 5 looks at the effects on academic performance, including performance in state tests conducted by the Michigan Department of Education and in college prep tests (ACT). Section 6 concludes.

2. THE MICHIGAN SCHOOL FINANCE REFORM

In 1994, just before the school finance reform, Michigan's property tax burden was the seventh highest in the country, and it was fourth among U.S. states in the share of locally financed school spending (61 percent).⁷ In March 1994, Michigan voters overwhelmingly ratified Proposal A, which reduced the reliance of school revenues on property taxes, replacing them primarily with an increase in the state sales tax from 4 percent to 6 percent.⁸ This resulted in a large increase in the state share of K–12 spending and was followed by efforts to make a significant dent in existing inequalities. An important point to note

7. The three states with a higher share of school expenditures financed locally were New Hampshire (86 percent), Illinois (62 percent), and Vermont (61 percent). Subsequently both Illinois and Vermont overhauled their school finance programs in 1997.

8. Taxes on homestead property came down from an average of thirty-four mills to a uniform statewide rate of six mills. The tax on non-homestead property was reduced to twenty-four mills. The share of the state in K–12 spending went up quickly, from 31.3 percent in 1993 to 77.5 percent in 1997. For more details on the Michigan program, see Courant and Loeb (1997) and Cullen and Loeb (2004).

Table 1. Increase in Foundation Allowances, Post-reform Michigan (selected districts at different percentiles of pre-reform spending distribution)

	1994	1995	1996	1997	1998	1999	2000	2001
Standish Sterling schools (1st percentile)	3,738	4,200	4,506	4,816	5,124	5,170	5,700	6,000
Harrison Community schools (5th percentile)	3,905	4,200	4,506	4,816	5,124	5,170	5,700	6,000
Adams Township schools (25th percentile)	4,321	4,566	4,832	5,099	5,362	5,362	5,700	6,000
Concord Community schools (50th percentile)	4,669	4,900	5,130	5,308	5,462	5,462	5,700	6,000
Oxford area schools (75th percentile)	5,249	5,458	5,611	5,766	5,920	5,920	6,158	6,458
Waverly Community schools (95th percentile)	6,998	7,159	7,312	7,467	7,621	7,621	7,743	7,914
Bloomfield Hills schools (99th percentile)	10,294	10,454	10,607	10,762	10,916	10,916	11,091	11,335

Notes: Foundation allowances refer to the base level of per pupil spending. 1994 refers to the academic year 1993–94, and so on. Note that due to court-related problems, foundation allowances did not increase between 1998 and 1999, except marginally for the lowest-spending districts.

Source: Michigan Department of Education 2010.

is that local discretion over future school spending was largely abolished; all increases in per pupil spending were henceforth governed by the state.⁹

The new school spending plan, effective from 1994 to 1995, works as follows. First, the 1993–94 level of spending in each district was taken as the starting point, and future increases were calculated based on this. Second, future increases in all districts' per pupil spending were governed entirely by the state legislature—the lowest-spending districts were allowed to increase spending at much faster rates than their higher-spending counterparts. In theory, over time this would lead to a substantial narrowing of the spending gap across districts. Further, all districts, regardless of how much they were spending, were held harmless—that is, no school district suffered any absolute decline in per pupil spending. The new state-mandated level of spending came to be called the district's foundation allowance.

Table 1 shows the changes in foundation allowances in Michigan school districts in the post-reform period. I show seven districts located at different percentiles in the pre-reform spending distribution. The large catch-up exhibited by the lowest-spending districts is immediately evident. For example, a

9. Prior to the reform and since the 1970s, Michigan had been using a district power equalizing (DPE) formula, where districts were allocated state funds based on their tax efforts. This was intended to make the system wealth neutral, leaving the choice of millage rates (property tax rates) to the local districts. However, since preference for school spending is in general a positive function of income (Feldstein 1975), the DPE system did not equalize per pupil spending across districts, which was one of its main objectives.

district like Standish-Sterling Community Schools, in the bottom percentile of the spending distribution in 1994, witnessed an increase of about 61 percent in foundation allowances over the next seven years. However, for Bloomfield Schools school district, already spending over \$10,000 per pupil in 1994, the increase was only about 10 percent.

As mentioned, unlike some other school finance reforms, in Michigan all school districts were held harmless. However, the fact that local discretion over spending was largely abolished following Proposal A (future increases in spending were dictated solely by the state) has interesting implications for the effect of the program on the high-spending districts. In these districts, per pupil spending increased barely at the rate of inflation after the reform and by much less than was the case just prior to the reform.¹⁰ It is important to explore whether the constraints on spending faced by these districts in the aftermath of Proposal A had any adverse consequences as far as student achievement is concerned. Below I study the effect of the reform on both spending and student performance in these districts.

3. DATA

In my analysis, I use data from multiple sources. Most of these come from the Michigan Department of Education (MDE). The revenue and expenditure figures, as well as those on K–12 enrollment and teacher salaries, are taken from the Bulletin 1014, published annually.¹¹ The data on ethnic and gender compositions and free lunch eligibility come from the Pupil Headcount Files and the Food and Nutrition Files of the MDE K–12 database.¹² The MDE K–12 database is also the source for the data on student achievement in state tests. I use student achievement data for reading and mathematics in both fourth and seventh grades.¹³

I use data from 1990 through 2001 for my analysis.¹⁴ This allows me to adequately capture differences in pre-reform trends across districts and also to capture program effects that may occur only with a lag.

-
10. The Consumer Price Index for Midwest Urban, which would cover Michigan, went up by 21 percent between 1994 and 2001.
 11. I mainly use data on general fund revenues and general fund expenditures because these are the most appropriate and widely used measures of spending. The exact definitions are available in the Bulletin 1014 (www.michigan.gov/mde/0,1607,7-140-5235-6539-21514-,00.html). I also use current operating expenditures as a validity check on my results.
 12. Some of the data on ethnicity and free lunch eligibility for the early years come from the Common Core of Data, a statistical database maintained by the National Center for Education Statistics.
 13. The state assessment in Michigan is known as the Michigan Educational Assessment Program (MEAP). Henceforth I will refer to these tests as MEAP tests. The exact proficiency measure is the percentage of students scoring at or above the *satisfactory* levels.
 14. Henceforth I refer to school years by the calendar year of the spring term—e.g., 1990 refers to academic year 1989–90, and so on.

Table 2. Inequality in Per Pupil Spending in Michigan before and after Proposal A

Measures of Inequality	Michigan	
	1994	2001
Ln(95th/5th)	0.562	0.395
Gini coefficient	0.104	0.084
Theil index	0.017	0.011
Coefficient of variation	0.190	0.151

Notes: Ln (95th/5th) is defined as the log of the ratio of school spending at the 95th and 5th percentiles. All statistics have been weighted by district enrollment.

The data on median household income in the school districts in 1989 come from the 1990 census, as published in the School District Data Book. Data on ACT participation and performance of Michigan school districts are obtained from Standard and Poor's School Evaluation Services.

4. EFFECT ON SCHOOL SPENDING

Table 2 compares inequality in per pupil spending in Michigan in the pre- and post-reform periods. I show the values of four common measures of spending inequality across districts—the Gini coefficient, the coefficient of variation, the Theil index, and the ratio of spending at the 95th and 5th percentiles. The values are shown for 1994, the last year before the reform, and 2001, the last year in my analysis. There has been a large decrease in each of the measures in Michigan since the reform, indicating a significant decline in differences in per pupil spending between districts.¹⁵

These numbers look even more impressive when examining figure 1, which compares the decline in variance of spending in Michigan with the corresponding estimates for an average court-mandated reform.¹⁶ Note that for Michigan I am only looking at a seven-year period, unlike a ten-year one as for the estimated effect of a court-ordered reform, and that because of the staggered nature of the reform in Michigan, inequality in spending has continued to further decline.

An important feature of school financing in the United States that troubles lawmakers, judges, and educators alike is the fact that per pupil spending depends largely on the affluence of the local school district. Many court cases,

15. I report only the weighted statistics (weighted by student enrollment of the district). For the unweighted statistics, the decline in inequality is even greater (except in the Gini), presumably due to the fact that most of the lowest-spending districts are located in rural areas and are relatively small in size.

16. These numbers show the average estimated decline in inequality in a state ten years after a typical court-mandated school finance reform and are taken from Murray, Evans, and Schwab (1998).

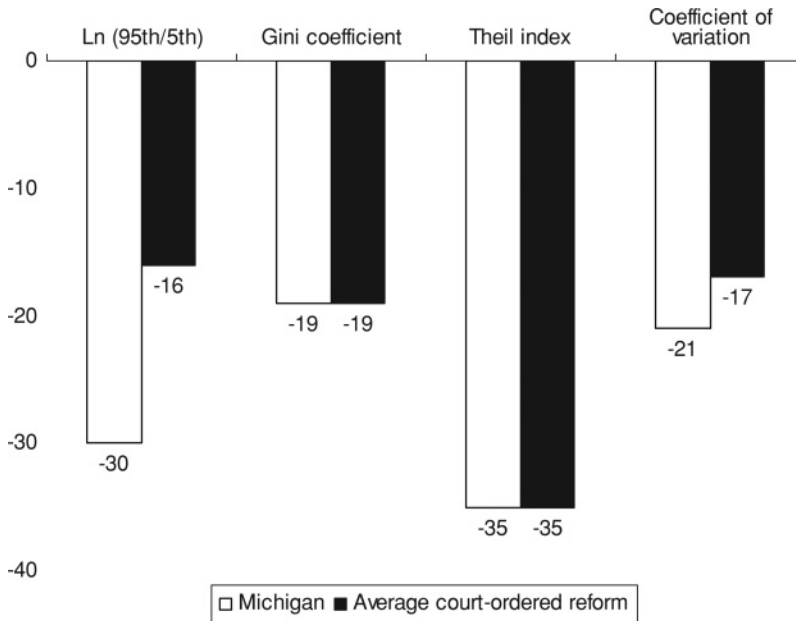


Figure 1. Decline in Inequality of Spending in Michigan, Compared with an Average Court-Ordered Reform. Note: The estimated reductions in the different inequality measures ten years after a typical court-mandated reform are taken from Murray, Evans, and Schwab (1998).

such as *Serrano I* in California (1971), have relied on this wealth-expenditure relationship as a yardstick of existing inequality. To investigate the strength of this relationship in pre-reform Michigan and the effect of the reform, if any, I proxy school district affluence with median household income in the district and estimate the following regression.¹⁷

$$\begin{aligned}
 Spending_{st} = & \alpha + \alpha_1 * D94 + \alpha_2 * D98 + \alpha_3 * D01 + \beta_0 * Income \\
 & + \beta_1 * (Income * D94) + \beta_2 * (Income * D98) \\
 & + \beta_3 * (Income * D01) + X_{st} + \varepsilon_{st}
 \end{aligned}
 \tag{1}$$

where $Spending_{st}$ denotes per pupil spending in district s in year t , $t \in \{1990, 1994, 1998, 2001\}$ and $Income$ denotes median household income of the district in 1989. $D94$, $D98$, and $D01$ are year dummies for 1994, 1998, and 2001, and X_{st} are the controls.¹⁸ The coefficient β_0 measures the strength of the income-expenditure relationship in 1990, the initial year in my analysis,

17. Strictly speaking, per pupil spending in a district is mostly a function of district property wealth, since most revenues come from the local property tax. However, district property wealth is very highly correlated with district incomes. Further, since incomes also reflect socioeconomic status in general, they are a strong predictor of desirability of the district as a potential place of residence.

18. The controls include racial composition, location of the district (proportion rural), and size; however, the results are very similar if I omit these variables.

Table 3. Relationship between Median Household Income in a School District and Its Per Pupil Spending, Michigan, 1990, 1994, 1998, and 2001

	(1)	(2)	(3)
Median income	0.31*** (0.04)	0.53*** (0.03)	0.47*** (0.03)
Median income * year 1994	-0.05 (0.05)	-0.10*** (0.04)	-0.04 (0.04)
Median income * year 1998	-0.25*** (0.05)	-0.26*** (0.04)	-0.24*** (0.04)
Median income * year 2001	-0.37*** (0.05)	-0.43*** (0.04)	-0.32*** (0.04)
Observations	2,079	2,079	2,075
R ²	0.74	0.82	0.80
Exclude Detroit	No	No	Yes

Notes: See equation 1 in the text. Column 1 shows the unweighted estimates, while regressions reported in columns 2 and 3 are weighted by respective district enrollments. All regressions control for racial composition, location of the district (proportion rural), and size. Column 3 excludes Detroit, which alone accounts for about 10% of Michigan's K–12 population. Robust standard errors are in parentheses.

***significant at 1%.

while β_1 gives the change in this gradient between 1990 and 1994, the last year before the program. The coefficients β_2 and β_3 show post-program changes; compared with β_1 they give an idea of the program's effectiveness in narrowing spending inequalities.

The results, in table 3, show a large and positive relationship between district income and school spending in 1990, which was only slightly attenuated between then and 1994. Post reform, however, there has been a very significant weakening, suggesting that Proposal A was instrumental in significantly equalizing school resources across districts in Michigan.

To further examine the effect of Proposal A on school spending in Michigan, I classify the 524 K–12 districts into five equal groups based on the 1993–94 level of per pupil spending (Group 1 consists of the lowest-spending 105 districts, Group 2 consists of the next 105 districts, and so on).¹⁹ Some summary statistics on these groups of districts are shown in table 4. As can be seen, the lowest-spending districts in Groups 1 and 2 are most similar to those in

19. Note that increases in spending in the post-reform period were related to spending in the base year (1994). I experimented with a variety of alternative classifications, such as grouping the districts such that there are an equal number of K–12 students in each group. The results are qualitatively similar. This is also true when I classify the districts based on their spending in 1990, which is not surprising, since the correlations in spending between 1990 and 1994 are very high. Note also that the groups of districts are defined based on pre-reform characteristics—spending in the base year (1994)—and these groups remain the same over time (as opposed to being redefined every year based on that year's relative spending levels).

Table 4. Summary Statistics for Different Groups of Michigan School Districts, 1994

	Group 1 Lowest-Spending Group	Group 2 Lower Middle Group	Group 3 Middle Group	Group 4 Upper Middle Group	Group 5 Highest-Spending Group
Ethnicity (percentage)					
Whites	93.59	91.06	93.29	55.13 (77.26)	82.37
Blacks	1.49	3.65	1.68	38.32 (15.49)	12.04
Hispanics	1.96	1.74	1.72	2.80 (3.16)	0.77
Free lunch eligibility	23.62	18.43	16.1	31.64 (22.97)	14.68
Pupil-teacher ratio	23.00	23.16	23.07	24.54 (23.88)	22.43
Grade 4 reading (MEAP)	40.59	43.66	42.72	40.36 (40.95)	49.44
Grade 4 math (MEAP)	60.01	63.51	64.04	56.08 (59.82)	67.45

Notes: For Group 4, the figures in parentheses correspond to the statistics when I leave out Detroit. Detroit is the largest school district in Michigan, accounting for about 10% of the total student population in the state. All figures have been weighted by enrollment of the districts in 1994. The results for MEAP scores correspond to 1995, as tests in Michigan during that time were administered in early fall.

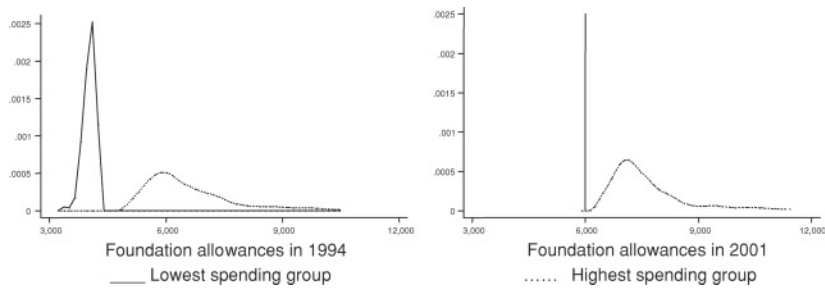


Figure 2. Distribution of Foundation Allowances in 1994 and 2001, Groups 1 and 5. Note: For 2001, the foundation allowances for all districts in Group 1 were at \$6,000.

Group 4. This is particularly so when Detroit, the largest school district in the state and that falls in group 4, is excluded. These districts were performing at very similar levels just before the program, and their percentage of students eligible for free or reduced price lunch and pupil-teacher ratios are also quite similar. There are some significant differences in racial composition—the percentage of black students is much higher in Group 4 districts, for example. Note, however, that I control for the racial composition of students, along with gender composition and free or reduced price lunch eligibility, in my regressions. Because of the similarity in performance levels in the immediate pre-reform period and the fact that most other indicators like free lunch eligibility and pupil-teacher ratios are also quite similar, I use Group 4 districts as the control group for districts in Groups 1 and 2.

Figure 2 shows the distributions of foundation allowances across school districts in Groups 1 and 5 in 1994 and 2001.²⁰ There has been a significant convergence between these groups in the post-reform period, in both absolute and relative terms. To formally compare the trends in spending in these different groups, before and after reform, I next run the fixed effects (FE) regression

$$\begin{aligned}
 Y_{sgt} = & \alpha + \alpha_s + \beta_o * t + \sum_{g \neq 4} \beta_g * (D_g * t) + \gamma_o * (reform) \\
 & + \sum_{g \neq 4} \gamma_g * (D_g * reform) + \theta_o * (reform * t) \\
 & + \sum_{g \neq 4} \theta_g * (D_g * reform * t) + \delta * X_{sgt} + \varepsilon_{sgt}
 \end{aligned} \tag{2}$$

20. These show the kernel smoothed plots of foundation allowances in the two groups of districts. All figures have been weighted by district enrollment. For 2001, the foundation allowances for all districts in Group 1 were at \$6,000.

where $g \in \{1, \dots, 5\}$, Y_{sgt} is the per pupil revenue or expenditure of district s in group g in year t , α_s is the district fixed effect, and X_{sgt} are the time-varying characteristics (controls).²¹ D_g is the dummy variable for the respective groups of districts. Group 4, comprising districts in the fourth quintile of the spending distribution in 1994 (the upper middle group), is the omitted category.²² *Reform* is a binary variable that takes the value of 0 in the pre-reform period (1990–94) and 1 afterward (1995–2001). The variable t represents time trend. The variables *reform* and (*reform* * t) control for post-reform common intercept and trend shifts, respectively. The coefficients on the interaction terms ($D_g * \text{reform}$) and ($D_g * \text{reform} * t$) estimate the program effects: γ_g captures the intercept shifts, while θ_g captures the trend shifts of different groups of districts. However, as mentioned earlier, the reform was staggered over several years and hence the immediate increase in spending was not large. So the estimated intercept effects are small, and below I focus on the trend shifts in the post-program period.²³

I use annual data from 1990 to 2001, which straddle 1994, the last year before the reform. The results are in table 5. I show the results for two samples—the first includes all 524 districts, the second excludes Detroit.²⁴ The first two columns show the results for revenues, the third and fourth columns show the results for expenditures. The gap between the highest-spending and the lowest-spending districts had been increasing in the years before the reform. In fact there was a clear hierarchy—the higher a group's level of spending (revenues or expenditures), the higher its growth rate of spending. Post reform, however, this has completely reversed itself. Once again there is a clear hierarchy, but in the opposite direction. The reform has been instrumental

-
21. Because free lunch data for 1990 and 1991 are either not available or not reliable because of small and inconsistent values, I have included only enrollment and racial composition in X_{sgt} . Running the regression on a subsample when data on all controls are available does not change the qualitative results.
 22. For ease of exposition, I use Group 4 districts as both the control group and the omitted category. This allows me to directly show the effect of the reform on the lowest-spending districts. As noted above, districts in Group 4 were performing at similar levels compared with districts in Groups 1 and 2, and they were also similar in other important indicators, though there were some differences in the respective racial mix-ups. Further, based on the percentile ranks of different districts in fourth-grade tests in 1994, the last year before the reform, districts in Groups 1 and 2 are most similar to districts in Group 4. Among other things, the similarity of performance in the pre-reform period between Groups 1 and 2 and Group 4 helps mitigate issues relating to mean-reversion, which is often important. However, my results are robust to the use of alternate comparison groups (Group 3 or Group 5).
 23. Since the districts are of different sizes, the regression errors are likely to be heteroskedastic. So I employ heteroskedasticity-robust standard errors in all regressions reported here. In addition, to control for serial correlation across observations for the same district, I cluster the standard errors at the school district level—the results are similar and hence are not reported separately.
 24. Detroit is the biggest school district in Michigan, alone accounting for about 10 percent of all Michigan K–12 students.

Table 5. Effect of Reform on Trends in Per Pupil Spending and Teacher Salaries, Michigan, 1990–2001

	General Fund Revenues		General Fund Expenditures		Teacher Salaries	
	(1)	(2)	(1)	(2)	(1)	(2)
Trend (t)	275*** (21.64)	265*** (16.06)	270*** (33.46)	223*** (15.69)	2,094*** (240.39)	2,325*** (102.64)
Group 1 * t	-63*** (23.01)	-53*** (17.79)	-43 (34.25)	4 (17.31)	-435* (263.77)	-666*** (147.67)
Group 2 * t	-27 (24.11)	-18 (19.41)	-23 (34.63)	23 (18.38)	-296 (258.11)	-518*** (141.47)
Group 3 * t	-15 (23.86)	-6 (18.91)	-27 (34.36)	19 (17.82)	-254 (257.66)	-484*** (138.80)
Group 5 * t	71*** (28.59)	81*** (24.52)	11 (37.53)	57** (23.26)	130 (276.26)	138 (163.88)
Reform * t	-20 (27.02)	-37* (18.77)	-10 (39.62)	17 (18.38)	-1,601*** (318.88)	-1,746*** (122.99)
Group 1 * reform * t	146*** (28.51)	162*** (20.78)	111*** (40.82)	83*** (20.59)	902*** (340.44)	1,057*** (168.61)
Group 2 * reform * t	46* (29.47)	64*** (22.47)	44 (41.03)	48** (21.69)	772** (337.33)	900*** (168.74)
Group 3 * reform * t	-16 (29.07)	1 (21.74)	-7 (40.70)	-34 (20.81)	502 (335.48)	637*** (162.22)
Group 5 * reform * t	-105*** (33.43)	-87*** (27.66)	-24 (43.75)	-53** (26.69)	10 (350.41)	103 (193.13)
Observations	6,269	6,257	6,269	6,257	6,266	6,254
R ²	0.96	0.96	0.96	0.96	0.87	0.90

Notes: See equation 2 in the text. Group 4 is the omitted category. Columns marked (1) include all 524 school districts, while columns marked (2) exclude Detroit, which is the largest district in the state (accounting for about 10% of the total number of students in the state). All regressions are weighted by district enrollment, include district fixed effects, and control for enrollment and ethnicity. For brevity, I do not report the other coefficients. Robust standard errors are in parentheses.

*significant at 10%; **significant at 5%; ***significant at 1%.

in significantly raising both revenues and expenditures in school districts in Groups 1 and 2.²⁵

In order to see whether the increases in spending have been translated into inputs, the last two columns of table 5 show the corresponding trends for teacher salaries. The trends for teacher salaries mirror those for revenues and expenditures: there has been a significant narrowing of the gap between Group 1 and Group 5 districts in the post-reform period. As with spending, this

25. The ordinary least squares (OLS) results, where the district fixed effects are replaced by group dummies for the four groups (Group 4 is the omitted category), are very similar and hence are not reported. In addition, I experimented with a variety of samples, often excluding the very small and the very large districts. The results are similar. The results are also similar if I consider current operating expenditures, another widely used measure of spending.

looks more impressive when compared with the divergent trends just prior to the program. This would be expected to improve academic performance if the new higher salaries attract better and more experienced teachers to the lowest-spending districts. However, any positive effect might be muted if the districts mostly use the money to pay existing teachers more.

To sum, the evidence points to a substantial program effect on equalization of school finances and teacher salaries across different districts. I next estimate whether this affected academic performance.

5. EFFECT ON ACADEMIC PERFORMANCE

If students in the lowest-spending districts are lagging behind others because of a lack of adequate resources, equalization of spending brought about by programs like Michigan's can spur significant gains in achievement. In this section I document the changes in student performance in post-reform Michigan, as measured by standardized test scores. I begin with the MEAP tests, which are administered annually in all Michigan school districts, and then analyze different districts' performance in ACT tests.

Performance in MEAP Tests

Trends in Academic Performance across Different Groups of Districts

Figure 3 shows the distributions of the changes in fourth-grade proficiency results (reading and mathematics) in Groups 1 and 5 between 1995 and 2001.²⁶ The top panel compares the change in reading proficiency, while the bottom panel is for mathematics proficiency. In either subject, the line for the lowest-spending group lies to the right of that for the highest-spending group, suggesting convergence in the post-reform period, with the lowest-spending districts narrowing the achievement gaps between the lowest- and the highest-spending districts.

Table 6 reports the results from running equation 2 on test scores. I show the results for two samples—the first includes all 524 districts, the second excludes Detroit. For reading, the estimates reflect the general trend seen earlier for spending—the lowest-spending districts were lagging behind in the pre-reform period but have made significant improvements since the reform. For mathematics, too, there is some evidence for improvements in these districts in the post-reform period; however, the coefficients are not generally significant.

Interestingly, though there are no clearly discernible trends in student performance across the years, the results are suggestive of a modest relative decline in the highest-spending districts in later years. For example, though

26. Since at that time MEAP tests were administered in early fall, I take 1995, instead of 1994, to be the base year when I look at improvements in academic performance. Note that I refer to school years by the calendar year of the spring term.

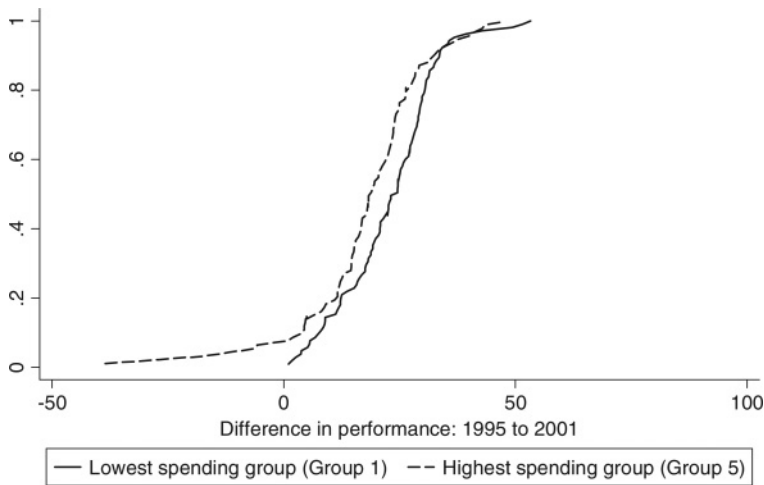


Figure 3a. Change in Performance between 1995 and 2001: Group 1 Districts versus Group 5 Districts—Reading.

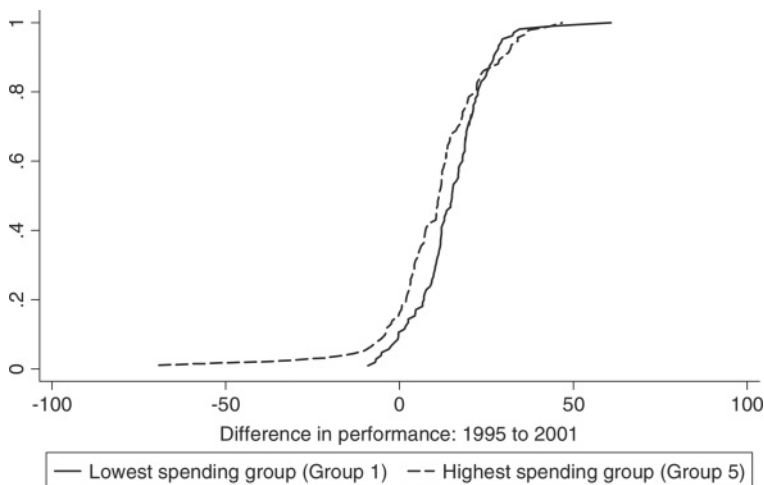


Figure 3b. Change in Performance between 1995 and 2001: Group 1 Districts versus Group 5 Districts—Mathematics. *Note:* Figures 3a and 3b show the cumulative density function differences between 1995 and 2001, separately for Group 1 and Group 5 districts, for reading and mathematics.

the coefficients in the early years (1996 and 1997) are positive, modestly large, and often significant, the coefficients in the later years are generally small, often negative, and always insignificant. This is true for both reading and mathematics, and it is particularly the case in regressions excluding Detroit (my preferred specification).

Overall, the evidence is suggestive of an improved performance by the lowest-spending districts later in the decade (and a relative deterioration by the

Table 6. Effect of Reform on Grade 4 Reading and Mathematics Tests, Michigan, 1993–2001

	Reading		Mathematics	
	(1)	(2)	(1)	(2)
Trend (t)	7.59*** (0.72)	6.60*** (0.55)	6.78*** (0.81)	6.30*** (0.72)
Group 1 * t	-2.23*** (0.80)	-1.14** (0.56)	-1.00 (0.85)	-0.63 (0.70)
Group 2 * t	-1.30 (0.83)	-0.2 (0.61)	0.10 (0.84)	0.79 (0.69)
Group 3 * t	-1.85** (0.79)	-0.76 (0.55)	-1.08 (0.84)	-0.71 (0.68)
Group 5 * t	0.94 (0.85)	2.07*** (0.63)	1.13 (0.89)	1.53** (0.75)
Reform * t	-4.86*** (1.01)	-3.14*** (0.59)	-4.27*** (1.03)	-3.50*** (0.77)
Group 1 * reform * t	2.97*** (1.14)	1.15* (0.68)	1.77* (1.14)	1.10 (0.84)
Group 2 * reform * t	2.22** (1.16)	0.40 (0.72)	0.02 (1.13)	0.64 (0.81)
Group 3 * reform * t	0.71 (1.12)	0.91 (0.66)	1.49 (1.12)	0.85 (0.81)
Group 5 * reform * t	-0.7 (1.20)	-2.48*** (0.78)	-1.32 (1.16)	-1.94** (0.87)
Observations	4,678	4,671	4,678	4,671
R ²	0.83	0.84	0.85	0.86

Notes: See equation 2 in the text. Group 4 is the omitted category. Columns marked (1) include all 524 school districts, while columns marked (2) exclude Detroit, which is the largest district in the state (accounting for about 10% of the total number of students in the state). All regressions are weighted by district enrollment, include district fixed effects, and control for enrollment and ethnicity. For brevity, I do not report the other coefficients. Robust standard errors are in parentheses.

*significant at 10%; **significant at 5%; ***significant at 1%.

highest-spending districts). However, looking at trends in student achievement for a particular grade has its limitations. I next present a cohort-based measure, which controls for idiosyncratic cohort-specific shocks across districts and is arguably a better way of judging true gains in student achievement.

Cohort-Based Analysis of Gains in Student Achievement

The idea here is that if the reform has had a direct effect on student performance, cohorts in the lowest-spending districts will exhibit higher growth rates, and these gains will be progressively higher as one moves onto more recent cohorts, which have spent more years under the program. This cohort-based measure should give an accurate picture of actual academic gains in the lowest-spending districts.

In Michigan, students are tested in reading and mathematics in both grades 4 and 7. I compare the improvement in student proficiency between the fourth and seventh grades across different groups of districts. I begin with the cohort that was in the fourth grade in 1993. These children would be in grade 7 in 1996, the second year of the reform, and hence would be affected for little more than a year.²⁷ The cohort in grade 4 in 1994 would similarly experience a little more than two years under the program. The next cohort, in grade 4 in 1995, would be affected for the entire period. However, if there are some lag effects, as seems likely, it may be the next cohorts (fourth graders in 1996, 1997, etc.) that show the most improvement.

For each cohort, I run both OLS and FE versions of the following model. For example, for the cohort in fourth grade in 1993 (and in seventh grade in 1996), the FE regression is

$$T_{sgt} = \alpha_o + \alpha_s + \beta_o * t + \sum_{g \neq 4} \beta_g * (D_g * t) + \delta * X_{sgt} + \varepsilon_{sgt}, \quad (3)$$

where T_{sgt} is the fourth-grade test score for $t = 1993$ and the seventh-grade test score for $t = 1996$. I run equation 3 for each cohort and record the coefficients β_1 , β_2 , and β_5 . An increasing trend in β_1 and β_2 , the coefficients for the two lowest-spending groups, over successive cohorts will be evidence in favor of the program. Similarly, a declining trend in β_5 , the coefficient for the highest-spending group, will suggest that the constraints imposed by Proposal A on future increases in spending are adversely affecting student performance in these districts. As earlier, Group 4 is the omitted category.²⁸

The results are given in table 7, which reports β_1 , β_2 , and β_5 for the five cohorts (fourth graders in 1993, 1994, 1995, 1996, and 1997).²⁹ For reading, the coefficients (which show the gains between the fourth and seventh grades for districts in Groups 1 and 2 compared with districts in Group 4) were significantly negative in the pre-reform years. But the differences narrowed down, and for the later post-reform years the coefficients are either mostly positive (Group 1) or only modestly negative (Group 2). The results for mathematics are similar but stronger—here the improvements for both Groups 1 and 2 follow a trend.

27. Ideally I should begin with a cohort unaffected by the reform. However, data on the control variables are not available for the previous year with fourth-grade test score data (1991, mathematics). Further, inclusion of this cohort does not change any of the results qualitatively.

28. The results are similar when, instead of Group 4, I use either Group 3 or Group 5 as the comparison group.

29. The sample includes all districts in Michigan except Detroit, since proficiency data for 1993 and 1994 are not available for Detroit. Including Detroit for the later years yields very similar results. These are not reported separately but are available on request.

Table 7. Gains in Student Achievement between Grades 4 and 7, Michigan: Reading and Mathematics, Cohorts in Groups 1, 2, and 5

Cohorts	Cohorts in Group 1 Districts	Cohorts in Group 2 Districts	Cohorts in Group 5 Districts
Reading			
4th graders in 1993 (7th graders in 1996)	-4.10*** (1.57)	-4.59*** (1.60)	-0.05 (1.55)
4th graders in 1994 (7th graders in 1997)	-5.02*** (1.36)	-3.54*** (1.38)	-2.36* (1.30)
4th graders in 1995 (7th graders in 1998)	-1.04 (1.35)	-0.39 (1.61)	-1.97 (1.46)
4th graders in 1996 (7th graders in 1999)	1.61 (1.62)	0.10 (1.66)	-7.65*** (1.79)
4th graders in 1997 (7th graders in 2000)	-0.68 (1.76)	-1.78 (1.76)	-4.65** (2.17)
4th graders in 1998 (7th graders in 2001)	1.98 (1.38)	0.53 (1.45)	-4.52*** (1.34)
Mathematics			
4th graders in 1993 (7th graders in 1996)	-0.79 (1.82)	-2.55 (1.81)	1.05 (1.80)
4th graders in 1994 (7th graders in 1997)	-0.69 (2.01)	-1.92 (1.96)	-1.87 (1.70)
4th graders in 1995 (7th graders in 1998)	1.35 (1.70)	0.04 (2.15)	0.84 (1.52)
4th graders in 1996 (7th graders in 1999)	4.47*** (1.82)	2.33 (2.00)	-3.43** (1.66)
4th graders in 1997 (7th graders in 2000)	4.53** (1.80)	1.94 (2.09)	-2.88* (1.69)

Notes: See equation 3 in the text. A separate regression is run for each cohort. Group 4 is the omitted category. The sample includes all districts in Michigan except Detroit. (Since proficiency data for 1993 and 1994 are not available for Detroit, I do not include Detroit in regressions reported in this table. Including Detroit for the later years yields very similar results; these are not reported separately but are available on request.) All regressions are weighted by the number of test takers and control for ethnicity, gender, enrollment, and free lunch eligibility. Robust standard errors are in parentheses.

*significant at 10%; **significant at 5%; ***significant at 1%.

Figure 4 shows the estimated coefficients for the different cohorts in the lowest-spending districts, together with the 95 percent confidence interval bands. An improvement in performance over the years is readily apparent.

It is also instructive to look at the results for the highest-spending districts. The trends here strongly suggest a relative deterioration in performance in Group 5 districts following Proposal A. For example, though in the early years the difference in test score gains between the fourth and the seventh grades for districts in Group 5 were generally not significantly different from those in

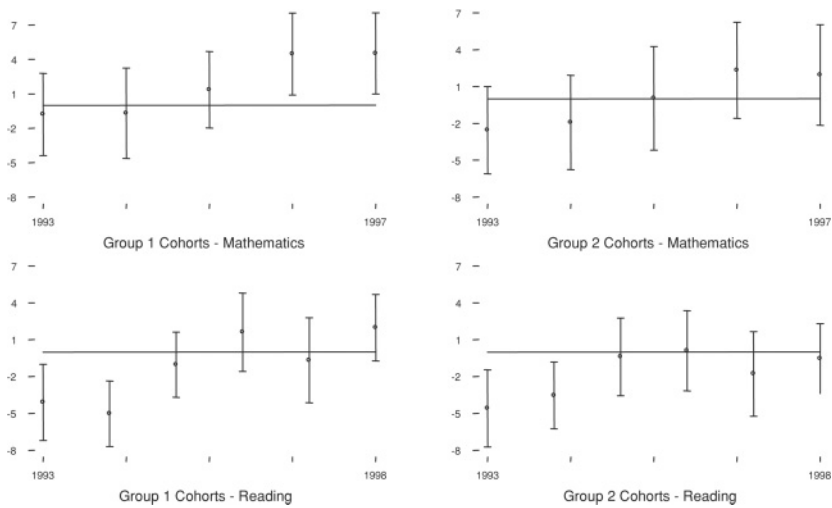


Figure 4. Gain in Reading and Mathematics between Grades 4 and 7 for Group 1 and 2 Cohorts. *Notes:* The figure shows the estimated coefficients on Groups 1 and 2 from regressions of equation 3, as reported in table 7. The small circles represent the estimated values, while the lines show the respective 95 percent confidence interval bands. Cohorts are identified by the year they were in grade 4; for example, the 1993 cohort is the one that was in grade 4 in 1993 and in grade 7 in 1996.

Group 4 districts (the omitted category), this was not the case in the later years. The coefficients in the later years are often large and negatively significant, particularly for reading but also for mathematics. The restraint on increases in spending in these districts may have led to adverse consequences for student performance, despite the fact that these districts were already spending at relatively high levels at the time of the program.

These results on cohort-based gains are particularly interesting, since they suggest that the extra resources in the lowest-spending districts had positive effects on their performance. I next present some direct evidence on the effects of increased expenditures on student performance.

Instrumental Variables Estimates of the Effect of Spending on Academic Performance

An alternative way to test for the reform's effect on academic achievement is to directly assess the relationship between the increases in spending (resulting from higher state aid) and student performance. However, to get consistent estimates we need district spending to be orthogonal to all the determinants of academic performance excluded from this equation, an assumption whose veracity may be in doubt. Note that this is true even for regressions that control for district fixed effects: even though this controls for all unobserved characteristics of a district that do not vary over time, there might be other, time-varying characteristics that affect both performance and spending.

I address this concern by using an instrumental variables strategy. I instrument district spending by the state-mandated foundation allowance, which meets the two valid criteria for an instrument. First, the foundation allowances are very closely correlated with district spending. Second, since it is determined at the state level, foundation allowances should not be correlated with the unobserved determinants of test performance at the district level.

Note, however, that the state-mandated foundation allowances, although following clear-cut rules, are functions of school spending in the base year 1994. Thus our maintained hypothesis, which is that foundation allowances should not affect student performance except through their effect on current spending, boils down to the following: 1994 spending should not affect performance from 1996 onward except through its effect on current spending. As is obvious, this may not always hold if there are some lag effects.

But note that I am primarily looking at fourth-grade test performance. So it is plausible to assume that any (noncapital) expenditure incurred by the district before the current fourth graders began attending school will have only negligible effects on them. That is, I can restrict attention to fourth-grade scores from 1998 onward—these fourth graders would be in grade 1 in 1995 or later and should presumably not be affected much by 1994 spending (except inasmuch as it affects current spending). I report results for both periods, 1996–2001 and 1998–2001. Moreover, since districts at different points of the spending distribution may be different from each other in a way that is not entirely captured by a district fixed effect, I curtail my sample by excluding the highest-spending districts (districts in Group 5), since they look somewhat different from the other groups.³⁰

As a further check, I use two different (but close) sets of instruments. The IV(1) estimates present results from regressions where I instrument spending with the foundation allowances. For IV(2) estimates I use the group indicators (dummies) and interactions of the group indicators with the year dummies as instruments. Intuitively, the latter is a way of rescaling the group effects I found earlier in dollar terms.³¹

I run both OLS and FE regressions. The FE regressions use the changes in foundation allowances as instruments for changes in per pupil spending and control for all time-invariant characteristics of the school districts, such as their pre-reform levels of spending. The FE regressions are my preferred specification; henceforth I refer to the instrumented FE estimates as FE-IV.³² In addition, because for much of this period the MEAP tests were conducted

30. Results using the entire sample are similar and are available on request.

31. The results from the IV(2) regressions are similar to those from the IV(1) regressions. These are not shown separately but are available on request.

32. The OLS results are not reported but are available on request.

Table 8. Effect of Per Pupil Spending on Grade 4 Test Performance, Michigan: Grade 4 Reading and Mathematics, 1996–2001 and 1998–2001

	1996–2001		1998–2001	
	(1)	(2)	(1)	(2)
Panel A				
First-stage regressions				
Foundation allowances	0.73*** (0.04)	0.69*** (0.04)	0.80*** (0.06)	0.71*** (0.06)
Observations	2,507	2,501	1,673	1,669
R ²	0.52	0.53	0.42	0.45
F-statistic	207.27	216.77	111.16	123.04
Panel B				
Instrumental variables regressions				
Reading				
Lagged spending	0.022 (0.022)	0.033* (0.020)	0.049* (0.025)	0.057** (0.028)
Observations	2,507	2,501	1,673	1,669
R ²	0.50	0.57	0.14	0.11
Mathematics				
Lagged spending	0.089*** (0.024)	0.083*** (0.023)	0.082*** (0.028)	0.061** (0.028)
Observations	2,507	2,501	1,673	1,669
R ²	0.49	0.56	0.20	0.06

Notes: Lagged spending is per pupil spending lagged one year. The regressions reported in columns marked (1) include all school districts in Groups 1–4, while those reported in columns marked (2) exclude Detroit and are weighted by respective district enrollments. All regressions include separate year dummies and control for racial and gender compositions and free lunch eligibility. Robust standard errors are in parentheses.

*significant at 10%; **significant at 5%; ***significant at 1%.

in early fall of the academic year, I use one-year lagged spending instead of contemporaneous spending as my independent variable.³³

Table 8 shows the results from the IV estimation. The first two columns show the results for 1996–2001, while the last two columns report results for 1998–2001. Panel A shows the results of the first-stage regressions. As expected, the coefficients on foundation allowances are highly significant, showing that they are the major determinant of per pupil spending in the post-reform period.³⁴ Panel B shows the results for the second-stage regressions. The estimated coefficients are modest to large and sometimes highly significant. These imply a modest effect of spending on academic performance.

33. The results from using contemporaneous spending are qualitatively similar.

34. In regressions reported here, I use general fund expenditures as the measure of per pupil spending. The results are very similar if I use general fund revenues as the measure of spending. Also, all the nominal variables used below—general fund expenditures and foundation allowances—have been deflated using the Consumer Price Index for Midwest Urban, obtained from the Bureau of Labor Statistics.

For example, the model in columns marked 2, which excludes Detroit and weights the observations by student enrollment, is my preferred specification. For reading the estimates imply an increase of between 3 and 6 percentage points in pass rates (percent of students scoring at or above the satisfactory level) for every \$1,000. With an in-sample standard deviation of about 14, this translates to about 0.20–0.40 standard deviations increase per \$1,000 of extra spending. For mathematics, the estimates in column 2 imply an increase of between 6 and 8 percentage points in pass rates for every \$1,000, an increase of 0.40–0.55 in standard deviations.³⁵

These estimates—0.20–0.40 standard deviations for reading and 0.40–0.55 standard deviations for mathematics—are for an increase in real per pupil expenditure of \$1,000. Typically even the lowest-spending districts in Michigan would have seen such an increase over a period of about four years.³⁶ Thus this boils down to roughly one-quarter to well over one-half a standard deviation increase over four years. This is a modest to large effect.

Finally, it might be useful to compare the results obtained above with the estimates in Papke (2005). The IV estimates in her study suggest that a 10 percent increase in spending increases the pass rates in fourth-grade mathematics by about 2.2 percentage points. Since the average per pupil spending in Michigan during her sample period was slightly over \$5,000 (see tables 2 and 3 in her study), a 10 percent increase in spending translates to about \$500. This is comparable to but slightly smaller than my estimates reported above. I find an increase of between 3 and 4 percentage points in fourth-grade mathematics pass rates for every \$500 of additional spending. The differences in actual magnitudes can be attributed to differences in sample and methodology—for example, she uses a smaller number of years (ending at 1998, while this study uses data through 2001). The regressions here also control for racial composition, etc.

Robustness of the Results

In this section I perform some robustness checks. First, it is possible that the rapid spread of charter schools in Michigan could have contaminated some of my results. Michigan passed a charter school law in the mid-1990s, along with the sweeping changes in school financing.³⁷ Many commentators believe that

35. The standard deviations in the reading tests for 1996–2001 and 1998–2001 are 14.32 and 13.63, respectively. The standard deviations in the mathematics tests are 14.88 and 13.31, respectively.

36. The foundation allowances for the lowest-spending districts were increasing by about \$300 per year. Cumulated over four years and appropriately deflated, this should be roughly around \$1,000.

37. There is also an interdistrict choice program in Michigan. However, it is very small. Only about 1 percent and 1.5 percent, respectively, of Michigan public school students enrolled in public schools outside their home district in 2000 and 2001 (see Arsen, Plank, and Sykes 2001). As is somewhat true for charter schools too (see below), public school choice is mainly concentrated in and around Detroit. As Cullen and Loeb (2004, p. 242) note, “Student participation in schools of choice has

Table 9. Concentration of Charter Schools across Different Groups, Michigan, 1996–2001

	Percentage of Students in Charter Schools					
	1996	1997	1998	1999	2000	2001
Lowest-spending group	0.10	0.29	0.58	0.84	0.96	1.07
Lower middle group	0.10	0.41	0.61	0.95	1.24	1.46
Middle group	0.25	0.63	1.14	1.60	2.13	2.45
Upper middle group	0.27	0.78	1.33	1.91	2.63	3.03
Highest-spending group	0.31	0.71	1.41	2.03	2.77	3.26
Michigan	0.25	0.68	1.25	1.95	2.71	3.40

Notes: In Michigan, charter schools are attached to an intermediate school district (ISD), comprising several school districts, rather than to one individual school district. The table shows the average ISD-wide percentages of charter students across various groups of districts; for example, a student in the lowest-spending group in 2000 would on average belong to an ISD where 0.96% of the students are enrolled in charter schools. The data in the last row (Michigan) come from Arsen, Plank, and Sykes (2001).

the beneficial effect of charter schools will spill over to students who remain in traditional public schools by increasing the latter's productivity in the face of intense competition for students. Hoxby (2003) shows that in Michigan this is exactly what happened—districts that had a larger percentage of students in charter schools increased their productivity at a faster rate than others not similarly threatened.

However, even though charter schools have spread very rapidly in Michigan,³⁸ they still serve only a fraction of overall K–12 students (see table 9, last row). Second, the presence of charter schools would bias some of my results only if it were true that these schools are relatively more concentrated in the lowest-spending districts. However, the opposite is true in Michigan. The lowest-spending groups (Groups 1 and 2) are predominantly rural, while charter schools in Michigan mostly serve urban children and are located in the higher-spending districts. Table 9 shows the growth of charter schools across the different groups of districts in Michigan. For each of the years, the percentage of students enrolled in charter schools is the smallest in the lowest-spending districts and vice versa. In fact, if competition from charter schools indeed encouraged increased effort in public schools (Hoxby 2003), my results would be an underestimate, since the competition effect would be strongest in districts in Groups 4 and 5. Geographically, many charter schools are located in southeast Michigan, particularly in Wayne County, where they serve mostly students living in the poorer suburbs or inner-city Detroit. The

largely been a Detroit phenomenon, with more than one-third of all transfers taking place within the Detroit metropolitan area."

38. The proportion of Michigan public school students enrolled in charter schools is one of the highest in the country.

results reported above are robust to excluding the ten school districts, including Detroit, located in Wayne County that fall in Group 4, the comparison group.³⁹

I compared my estimates for the lowest-spending districts with those for districts in Group 5. These results, where Group 5 is the comparison group, are stronger. There were significant improvements by the lowest-spending districts post reform, particularly in later years, narrowing the performance gap. When I similarly use Group 3 as the comparison group, the estimates are slightly attenuated, but Group 1 districts still exhibit considerable improvements. Figures 3a and 3b compare the improvements in test scores in Group 1 districts between 1995 and 2001 with those in Group 5 districts. As is evident, districts in Group 1 improved at a much faster rate in both reading and mathematics.

Third, note that mean reversion, the statistical tendency whereby high- and low-scoring schools tend to score closer to the mean subsequently, is unlikely to explain much of the improvement. This is because the initial performance gap between the lowest-spending districts and those in the comparison group (districts in Group 4) was quite small, so the degree of mean reversion, if any, and ceiling effects on test scores, as identified by Cullen and Loeb (2004), should be similar for either group. In addition, the improvements in the lowest-spending districts roughly parallel the increase in per pupil spending (see tables 6 and 7 and figure 4), suggesting that the increases in spending and improvements in academic performance are causally related.

Fourth, the presence of some time-varying factors, which affect student performance and happen to be changing during this period, can bias the results. Note that the regressions control for the racial and gender compositions of the districts and the percentage of students eligible for free and reduced price lunch, so perhaps the two most important correlates of student performance are accounted for. Further, I control for pre-reform trends in the regressions, so some of the time-varying characteristics that affected student performance in the pre-reform period, and that happen to also be changing in the post-reform period, will be accounted for.

The evidence thus points to a significant positive effect of the Michigan reform on student achievement. While not ruling out substantial inefficiencies in the utilization of additional funds in Michigan, it seems that lack of resources may have been partially responsible in holding down achievement in some school districts.

39. None of the school districts in Wayne County fall in the lowest-spending quintiles (Groups 1 and 2). In fact, most of the school districts in Wayne County fall in Group 5, the highest-spending quintile, as these are located in the affluent suburbs of Detroit.

However, one point should be noted. So far I have relied exclusively on performance in the MEAP tests as my measure of educational progress. Focusing solely on results from a state assessment may, however, be misleading. In Michigan, the MEAP results are highly publicized, and there may be incentives for teachers and administrators to boost their test scores artificially.⁴⁰ Several such phenomena, such as “teaching to the test,” transferring regular students to special categories, and even outright cheating, have been noted by researchers for other states (see, e.g., Cullen and Reback 2006; Figlio and Getzler 2002; Jacob and Levitt 2003). Note, however, that this will bias my results only if teachers and administrators in the lowest-spending districts indulge in such activities at a higher rate than those in other districts. It is important to note that Group 4 districts (my comparison group for most analyses) were performing at a similar level compared with the lowest-spending districts before the reform, so it is not obvious that compared with them the lowest-spending districts would have a greater incentive for such behavior. Nevertheless, I supplement my analysis by looking at a national college preparatory test, the ACT, that largely avoids the above pitfalls.

Participation and Performance in ACT Tests

Across the nation, most college-bound students take at least one of the two college entrance tests, SAT and ACT. In Michigan, as in most midwestern states, ACT is more common. I check whether the reform has had any effect on ACT taking and performance.

The underlying idea is that since students mostly take the ACT in twelfth grade, those taking the exam in later years would have been exposed to the finance reform for a larger number of years. For example, someone taking the exam in 1997 would have spent only about two years under the reform, but someone taking the same exam in 2001 would have been exposed for six years. Beneficial effects of the finance reform, if any, would show up as gains for the more recent cohorts in the lowest-spending districts.

Data on college prep tests have been available only since 1997, so I have data for five cohorts—twelfth graders in 1997, 1998, 1999, 2000, and 2001. I run OLS and FE regressions, allowing for unrestricted year effects (each year effect corresponds to the respective cohort). The FE equation is

$$\begin{aligned}
 ACT_{sgt} = & \alpha + \alpha_s + \beta_{98,0} * D_{98} + \dots + \beta_{01,0} * D_{01} \\
 & + \sum_{g \neq 4} \beta_{98,g} * (D_g * D_{98}) + \dots + \sum_{g \neq 4} \beta_{01,g} * (D_g * D_{01}) \\
 & + \delta * X_{sgt} + \varepsilon_{sgt}.
 \end{aligned} \tag{4}$$

40. Though they are not literally “high stakes” tests as in some other states, there are awards (Governor’s Cup) for districts that turn in the best performances.

Here ACT_{sgt} is the percentage of high school seniors in district s in group g in year t who take the ACT, or the average score of ACT takers. Again, Group 4 serves as the comparison group.

Table 10 presents the estimates from regressions of equation 4. The first two columns show changes in ACT participation rates between 1997 and 2001, while the third and fourth columns report results for the ACT scores. There is no evidence of any positive impact of the reform on the lowest-spending districts in Michigan. Though most of the coefficients on Groups 1 and 2 are positive, they are small in magnitude, never significant, and do not show any trends toward improvement. This is true for ACT scores as well, although for Group 1 the coefficients in the later years are less negative. The differences are minimal and never significant. In addition, there is no evidence of any effect of Proposal A on the highest-spending districts, either on ACT participation or on ACT performance (average scale scores). All the coefficients are small in size and statistically insignificant.

The ACT results are somewhat intriguing in that the improvement seen earlier for the lowest-spending districts in Michigan did not get translated in ACT tests. This underlines the importance of looking at multiple outcome measures, as gains in performance in one achievement test do not always translate into similar gains in other achievement tests. However, it may be too early to expect sizable improvements in ACT measures. Even the most recent cohort included here would have spent just about half of its school life under the reform. This may also be the reason behind the lack of any effect for Group 5 districts as well—the adverse impacts may be evident only after a while.

6. CONCLUSION AND EXTENSIONS

In this article I study the Michigan school finance reform (Proposal A). In 1994, quite unexpectedly and without the prodding of any courts, Michigan initiated a drastic overhaul of its school financing system. Among other things, Proposal A significantly increased state aid to the lowest-spending districts and largely ended local control over school spending. Using data for both the pre- and the post-reform periods, which allows me to control for differences in pre-existing trends, I find that the program has been quite successful in reducing inequalities of school spending.

I next examine whether improvements in student performance followed the large increase in resources witnessed by the lowest-spending districts and whether there were any adverse effects on the highest-spending districts stemming from the constraints on spending. I find that there was significant positive improvement in performance in the lowest-spending districts, which had

Table 10. Trends in Performance in College Prep Examination, Michigan: ACT Participation and Scores, 1997–2001

	Participation in ACT		ACT Average Score	
	(1)	(2)	(1)	(2)
Group 1 * 1998	0.23 (1.35)	0.84 (1.25)	-0.14 (0.09)	-0.17* (0.09)
Group 2 * 1998	0.37 (1.21)	1.02 (1.12)	-0.05 (0.09)	-0.08 (0.10)
Group 3 * 1998	-1.2 (1.39)	-0.06 (1.22)	0.07 (0.09)	0.02 (0.10)
Group 5 * 1998	0.25 (1.21)	0.7 (1.10)	-0.01 (0.08)	-0.03 (0.08)
Group 1 * 1999	-0.59 (1.41)	0.76 (1.19)	-0.11 (0.10)	-0.12 (0.10)
Group 2 * 1999	-0.44 (1.39)	1.09 (1.15)	0.16 (0.09)	0.15 (0.09)
Group 3 * 1999	0.160 (1.39)	1.610 (1.14)	0.01 (0.10)	0 (0.10)
Group 5 * 1999	0.58 (1.35)	1.42 (1.06)	0.02 (0.08)	-0.01 (0.09)
Group 1 * 2000	0.45 (1.35)	0.13 (1.23)	-0.03 (0.10)	-0.01 (0.10)
Group 2 * 2000	1.44 (1.30)	1.21 (1.18)	0.09 (0.09)	0.11 (0.10)
Group 3 * 2000	-0.19 (1.36)	-0.39 (1.22)	0.16 (0.10)	0.18* (0.10)
Group 5 * 2000	2.60** (1.23)	2.84** (1.11)	0.07 (0.08)	0.08 (0.08)
Group 1 * 2001	0.4 (1.38)	0.25 (1.30)	-0.01 (0.10)	-0.01 (0.10)
Group 2 * 2001	-0.62 (1.33)	-0.56 (1.25)	0.05 (0.09)	0.04 (0.10)
Group 3 * 2001	-1.32 (1.33)	-1.08 (1.25)	0.04 (0.09)	0.02 (0.10)
Group 5 * 2001	0.96 (1.28)	1.27 (1.23)	0.02 (0.08)	0 (0.09)
Observations	2,613	2,608	2,600	2,595
School districts	523	522	522	521
R ²	0.86	0.86	0.95	0.91

Notes: See equation 4 in the text. For brevity, I report only the coefficients on the different groups interacted with the year dummies. All the groups are based on base year spending (1994), and Group 4 is the omitted category. Columns marked (1) include all 524 districts, while columns marked (2) exclude Detroit. All regressions are weighted by grade 12 enrollment and control for ethnicity, gender, and free lunch eligibility. Robust standard errors are in parentheses.

*significant at 10%; **significant at 5%.

erstwhile been lagging behind. However, the improvements do not seem to have spilled over to other outcome measures such as participation and performance in college prep tests (ACT). Moreover, there is suggestive evidence that the constraints imposed by Proposal A on discretionary increases in spending had a negative effect on student performance in the highest-spending districts. This is an intriguing result, with important policy implications. Note that this relative decline in performance occurred despite the facts that Michigan did not level down expenditures in the highest-spending districts, which were held harmless, and the state did not witness a decline in average school spending compared with the rest of the country in the post-Proposal A period (Cullen and Loeb 2004).

These findings have significant policy implications. First, they show that state legislatures can initiate and implement a comprehensive school finance reform, even one that is largely redistributive in nature. This is important considering the recent debate on the effectiveness of legislature-led reforms vis-à-vis court-ordered ones in equalizing per pupil spending. Second, although the increase in resources in the lowest-spending districts was instrumental in raising student performance, there still remain significant gaps in achievement across different districts. Third, the fact that the highest-spending districts may have been adversely affected implies that one has to be careful in designing appropriate school finance policies in order to avoid unintended consequences.

I am grateful to Cecilia Rouse, Roland Benabou, and Jeffrey Kling for their comments and suggestions. I also thank the editors, David Figlio and David Monk; two anonymous referees; Melissa Clark; Rajashri Chakrabarti; and seminar and conference participants at Princeton University, University of Colorado at Boulder, American Economic Association, Econometric Society, American Education Finance Association, and Southern Economic Association. Thanks are also due to Glenda Rader of the Michigan Department of Education and Jim Brown of the Ohio Department of Education for providing part of the data used in this analysis. All errors are my own.

REFERENCES

- Arsen, David, David N. Plank, and Gary Sykes. 2001. A work in progress. *Education Next* 1(4): 14–19.
- Card, David, and Abigail A. Payne. 2002. School finance reform, the distribution of school spending, and the distribution of student test scores. *Journal of Public Economics* 83: 49–82.
- Clark, Melissa A. 2003. Education reform, redistribution, and student achievement: Evidence from the Kentucky Education Reform Act. Unpublished paper, Princeton University.

- Courant, Paul N., and Susanna Loeb. 1997. Centralization of school finance in Michigan. *Journal of Policy Analysis and Management* 16(1): 114–36.
- Cullen, Julie Berry, and Susanna Loeb. 2004. School finance reform in Michigan: Evaluating Proposal A. In *Helping children left behind: State aid and the pursuit of educational equity*, edited by John Yinger, pp. 215–50. Cambridge, MA: MIT Press.
- Cullen, Julie Berry, and Randall Reback. 2006. Tinkering toward accolades: School gaming under a performance accountability system. In *Improving school accountability: Check-ups or choice, advances in applied microeconomics*, edited by Timothy J. Gronberg and Dennis W. Jansen, pp. 1–34. Amsterdam: Elsevier Science.
- Downes, Thomas A. 1992. Evaluating the impact of school finance reform on the provision of public education: The California case. *National Tax Journal* 45(4): 405–19.
- Evans, William N., Sheila E. Murray, and Robert M. Schwab. 1997. Schoolhouses, courthouses and statehouses after *Serrano*. *Journal of Policy Analysis and Management* 16(1): 10–31.
- Feldstein, Martin S. 1975. Wealth neutrality and local choice in public education. *American Economic Review* 65(1): 75–89.
- Fernandez, Raquel, and Richard Rogerson. 1996. Income distribution, communities and the quality of public education. *Quarterly Journal of Economics* 111: 135–64.
- Figlio, David N., and Lawrence S. Getzler. 2002. Accountability, ability and disability: Gaming the system. NBER Working Paper No. 9307.
- Guryan, Jonathan. 2001. Does money matter? Regression discontinuity estimates from education finance reform in Massachusetts. NBER Working Paper No. 8269.
- Hanushek, Eric. 1991. When school finance “reform” may not be a good policy. *Harvard Journal on Legislation* 28: 423–56.
- Hoxby, Caroline. 2003. School choice and school productivity: Could school choice be the tide that lifts all boats? In *The economics of school choice*, edited by Caroline Hoxby, pp. 287–342. Chicago: University of Chicago Press.
- Jacob, Brian, and Steven D. Levitt. 2003. Rotten apples: An investigation of the prevalence and predictors of teacher cheating. *Quarterly Journal of Economics* 118(3): 843–77.
- Michigan Department of Education. 2010. *Foundation allowances 1995–2010*. Available www.michigan.gov/documents/fdtnall_49926_7.pdf. Accessed 30 June 2010.
- Murray, Sheila E., William N. Evans, and Robert M. Schwab. 1998. Education finance reform and the distribution of education resources. *American Economic Review* 88(4): 789–812.
- Papke, Leslie. 2005. The effects of spending on test pass rates: Evidence from Michigan. *Journal of Public Economics* 89(5–6): 821–39.

Papke, Leslie. 2008. The effects of changes in Michigan's school finance system. *Public Finance Review* 36(4): 456–74.

Schemo, Diana Jean. 2002. Few exercise new right to leave failing schools. *New York Times*, 28 August.

Serrano v. Priest (Serrano I). 1971. 5 Cal. 3d 584.

Yinger, John. 2004. *Helping children left behind: State aid and the pursuit of educational equity*. Cambridge, MA: MIT Press.