

Impact of School Finance Reform on Resource Equalization and Academic Performance: Evidence from Michigan*

Joydeep Roy[†]

Princeton University

October 2003

Abstract

The state of Michigan radically altered its school finance system in 1994. This was a legislature-led reform that took place somewhat unexpectedly and without the intervention of any courts. The new plan, called Proposal A, significantly increased state aid to the lowest spending districts and limited future increases in spending in the richest ones. I investigate the impact of Proposal A on distribution of resources and educational outcomes in Michigan. In the process this paper offers a first detailed look at the effectiveness of a legislature-led school finance reform. Using panel data on Michigan K-12 districts from 1990 to 2001, I find that Proposal A was quite successful in reducing inter-district spending disparities. The effect on academic performance is more modest. Though there is evidence of significant gains by the lowest spending districts in state tests, as yet these do not show up in nationwide tests like NAEP and ACT. The results are largely unchanged when I use Indiana and Ohio as control states.

Keywords: School Finance, Inequality and Redistribution, Academic Performance

JEL Classification: H4, H7, I2

*I am grateful to Jeffrey Kling and Cecilia Rouse for their invaluable advice and constant encouragement. I would like to thank Roland Benabou and Melissa Clark and participants at the Princeton University Labor Lunch and Princeton University Labor Economics Seminar for helpful comments and discussions. Thanks are also due to Glenda Rader and Jim Brown of the Michigan and Ohio Departments of Education. All errors are my own.

[†]Department of Economics, Princeton University, Princeton, NJ 08544, email: jroy@princeton.edu

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 become commonplace across the U.S. Michigan is among the latest entrants to this group, having radically altered its school financing rules in 1994. In this paper I evaluate the Michigan school finance reform.

In most U.S. states, local property taxes are the major source of school revenues, and since people often sort themselves among neighborhoods based on income, taxable property wealth differs substantially from district to district, even within the same state.¹ School finance reforms are aimed at weakening this nexus between district wealth and higher per pupil expenditures. They achieve this typically by large increases in state aid to poorer districts, often coupled with restrictions on spending in the richer ones.

The rationale behind such overhaul is that children in poorer districts may be lagging behind others because of inadequate resources at their disposal. With education becoming more and more important in a knowledge-driven economy, this is of immediate concern to parents and educators alike. It is interesting to note here U.S. Supreme Court Judge Sandra O'Connor's recent interview where she suggested that some affirmative action in U.S. universities was needed because in their current position, U.S. schools were unable to bridge the socioeconomic divide.² An important aspect of school finance reforms is the fact that often a substantial share of the relevant population is affected, so potentially this can have large effects on educational outcomes. On the other hand, critics of such reforms argue that the way these are implemented - large windfalls to poorer districts - make it highly unlikely that they will lead to any meaningful improvement for the students concerned (Hanushek,

¹ See Table 1 in Fernandez and Rogerson (1996) for an illuminating example of inequalities within nearby communities.

² Chicago Tribune, June 24, 2003.

1996).

In 1994, Michigan embarked on a comprehensive overhaul of its school finance program when it enacted a new plan called Proposal A. This significantly increased the state share of K-12 spending and entailed 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. Given spending in 1993-94, the last year before the program, it is the state that now decides by how much each district can raise its subsequent expenditures.

The theoretical literature about school spending concerns itself primarily with the difference between local and state finance, and their implications for growth and inequality. In a Tiebout model, where people sort themselves into neighborhoods based on their income, a local system is believed to be preferable to a state system, if one is only concerned about available resources to education and not their distribution.³ However, in an important contribution, Benabou (1996a) notes the major (and hitherto neglected) role of complementarity between the educational production function and human capital in the neighborhood (local 'social capital'). He shows that, depending on the strength of this complementarity, a move from local to state financing may or may not be welfare-improving. Benabou (1996b) explores the effects of different types of school finance systems on growth and inequality in a dynamic framework. In the presence of credit market imperfections, though there might be an inter-temporal tradeoff under certain circumstances, the economy's growth rate is higher under a state system, assuming rational forward-looking agents.

The empirical studies undertaken so far fall under two broad groups - those that deal cross-sectionally with many school finance reforms and those that study individual states. Among the former, Murray et al (1998) conclude that court-mandated finance reforms have had a large positive effect on equalization of school resources.⁴ Card and Payne (2002) study not only the relative

³ See e.g. the calibration results in Fernandez and Rogerson (1999). This result assumes that the income elasticity of demand for school expenditures is greater than the price elasticity.

⁴ However, they caution that the scope of these reforms is somewhat limited, since they cannot make a dent into

equalization in spending across districts, but also its consequences for academic performance. They find that such reforms, which were successful in reducing inter-district disparities in spending, also led to a convergence in SAT ⁵ scores across family background groups.

Among the case studies of individual states, Downes (1992) finds that in California there has been a significant convergence across school districts in per pupil expenditures, though it has not been reflected in academic performance. For Vermont too, Downes finds that spending has become much more equally distributed (2002).⁶ Guryan (2001), who studies the Massachusetts reforms of 1993, finds that the increase in spending due to an increase in state aid improved 4th grade test scores, primarily through a better performance by low-scoring students. 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.⁷

However, all these studies concern states where the financing system was changed following directives from the courts. In Michigan, on the other hand, the campaign for reform was led by the legislature and the executive.⁸ There is an as-yet unresolved issue in the school finance literature about whether legislature-mandated changes can have real effects. Evans et al (1997), based on their analysis of 16,000 public school districts from 1972 to 1992, conclude that “reforms that 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 towards the lower income districts.⁹ This is the first paper to examine a legislature-led school finance reform between-state inequalities which dominate within-state inequalities for the country as a whole.

⁵ SAT is the acronym for Scholastic Aptitude Test which, along with ACT, is the most popular college prep test in the U.S.

⁶ Since the Vermont legislation was enacted only in 1997, it was considered too early to look at test scores.

⁷ The ACT assessment, known as the American College Testing Program till 1996, is America’s most widely accepted college entrance examination. It is particularly important in the Midwestern and Southern states where it is often required for admission to colleges.

⁸ The two court cases of the previous two decades, Milliken vs. Green (1973) and East Jackson Public Schools vs. Michigan (1984), had both found the existing finance system constitutional.

⁹ See Table 4 in Card and Payne. Both studies use similar data, though Card and Payne only consider the change between 1977 and 1992. The approach to the problem, and the empirical methodology, is however quite different.

closely, and should provide important evidence as to its effectiveness.¹⁰

I begin by looking at the effects on equalization of school spending. I find that the program was indeed quite successful on this count - by the end of the decade the lowest spending districts had witnessed large increases in spending. Next I look at the trends in academic performance. I employ various strategies, 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 test score gains by these districts. These gains are robust to alternative control groups, and hold good when I look at the experience of two neighboring states, Indiana and Ohio. However, there is not much evidence for any improved performance by these lowest spending districts in college prep test (ACT). Neither is any relative improvement in nationally-conducted NAEP¹¹ tests visible, at least in mathematics.

The remainder of the paper is arranged as follows. In Section 2 I outline the main features of the Michigan program. Section 3 discusses the sources of the data 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. I begin with state tests conducted by the Michigan Department of Education, but I supplement this with results from nationwide tests and compare the Michigan experience to those of two neighboring states, Indiana and Ohio. Section 6 concludes.

¹⁰ Both of the cross-section studies mentioned above use data only up to 1992, so they exclude the Michigan reforms, which were initiated only in 1994.

¹¹ NAEP is the acronym for National Assessment of Educational Progress. It consists of tests in various grades and subjects, conducted periodically across the U.S. states.

2 The Michigan Program

Since the Michigan program has been discussed at length elsewhere,¹² I confine myself to the main features. As already mentioned, the program was not a response to any adverse court ruling or to a sudden rise in public concern over inequalities. Rather, it was a by-product of the prevailing debate over pervasively-high property taxes, whose main purpose was supporting local schools. In 1994, just before the program, Michigan's property tax burden was the seventh highest in the country and it was 4th among U.S. states in the share of school spending financed locally (61%).¹³ In March 1994 Michigan voters overwhelmingly ratified Proposal A, which reduced the reliance of school revenues on property taxes, replacing them primarily by an increase in the sales tax from 4 to 6%. This resulted in a large rise in the state share of K-12 spending, and was followed by efforts to make a significant dent in existing inequalities.¹⁴

Prior to the reform, Michigan had been using a district power equalizing (DPE) formula, where districts are allocated state funds based on their tax efforts. This was intended to make the system wealth-neutral,¹⁵ leaving the choice of millage rates (property tax rates) to the local districts, but its equalizing power had considerably eroded over the years. In 1994 about one-third of all districts were too rich to be affected.

The new school spending plan, effective from 1994-95, works as follows. First, the 1993-94 level of spending in each district was taken as its base, and came to be called the district's Foundation Allowance (FA). Second, future increases in all districts' FA's were governed entirely by the state

¹² See e.g., Addonizio et al (1995), Courant et al (1995) and Courant and Loeb (1997).

¹³ after New Hampshire (86%), Illinois (62%) and Vermont (61%) - subsequently, both Illinois and Vermont overhauled their school finance programs in 1997.

¹⁴ Taxes on homestead property came down from an average of 34 mills to a uniform statewide rate of 6 mills. The tax on non-homestead property was reduced too, but kept at 24 mills. The share of the state in K-12 spending went up quickly, from 31.3% in 1993 to 77.5% in 1997.

¹⁵ The idea behind wealth-neutrality is that high tax wealth in a district should not lead to high revenues except through a higher tax effort. In general, as preference for school spending is a positive function of income, this does not equalize per pupil expenditures across districts, see Feldstein (1975).

legislature - the lowest spending districts were allowed to increase spending at much faster rates than their richer counterparts. In theory, over time this would lead to a substantial narrowing of the revenue gap across districts. Further, all districts, however rich, were held harmless - none suffered any absolute decline in per pupil spending. Table 1 shows the changes in expenditures in six Michigan school districts in the post-reform period. The large catch-up exhibited by the lowest spending districts is immediately evident.

Another feature of the Michigan program is that it abolishes, for all practical purposes, local control over future increases in spending. This is a controversial issue - Hoxby (1996, 2001) argues that finance programs that tamper with local control over spending are harmful for school productivity and decrease support for overall school spending. In her opinion, finance formulas should take the form of flat grants and not distort incentives.¹⁶

3 Data

All the data that I use, except for the NAEP, are at the level of individual K-12 districts. Most of these come from the Michigan Department of Education (henceforth, MDE). The revenue and expenditure figures, as well as those on K-12 enrollment, teacher salaries and pupil-teacher ratios, are taken from the Bulletin 1014's, published annually.¹⁷ The data on ethnic and gender compositions and free lunch eligibility (4th and 7th grades) come from the Pupil Headcount Files and the Food and Nutrition Files of the MDE K-12 database.¹⁸ The MDE K-12 Database is also the source for

¹⁶ That is, if a district wishes to tax itself \$1 to increase its school spending by \$1, it should be allowed to do so. The California experience provides much of the fodder for this debate. In the years after the Serrano II case (1976), where the court ruled a difference of more than 200 dollars in spending between school districts to be unconstitutional, California witnessed a drastic fall in per pupil expenditures. Some critics believe this to be a harmful fallout of the move to a state system. There are other possible explanations though, such as a significant rise in enrollment, particularly of Hispanics, and Proposition 13, the 1978 tax limit bill which restricts increases in state expenditures.

¹⁷ I mostly use data on general fund revenues and general fund expenditures, these are broad measures that include most categories of spending. For the exact definitions, see any of the Bulletin 1014's.

¹⁸ Some of the data on ethnicity and free lunch eligibility for the early years come from the Common Core of Data (CCD) of the National Center for Education Statistics (NCES).

the state test (Michigan Educational Assessment Program, henceforth MEAP) scores. I only use the results for reading and mathematics (4th and 7th grades, percent scoring at or above the satisfactory levels).

The data on median income of the school districts in 1989 come from the 1990 census, as published in the School District Data Book. I also use the 1990 and 2000 censuses to look at changes in socioeconomic and demographic characteristics of Michigan school districts. Data on ACT and SAT participation and performance of Michigan school districts come from the Michigan website of Standard and Poor's School Evaluation Services. The NAEP test scores for Michigan and other states are available on the NCES website.

The data on Indiana and Ohio come from the respective departments of education. For Ohio, these have been extracted from the annual Local Report Cards, available for each district since 1996. For Indiana, the K-12 School Data is the repository of much of the information. The measures of spending used are General Expenditures for Indiana and Total Expenditures for Ohio, which are broadly similar to General Fund Expenditures in Michigan.¹⁹ For test scores, I use results from 3rd and 6th grades reading and mathematics in Indiana and 4th and 7th grades reading and mathematics in Ohio.²⁰

4 Effect on School Spending

Table 2 shows 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.²¹ The values are shown for 1994, the last year before the reform, and 2001, the last year

¹⁹ For Ohio, due to a change in definition, comparable figures are not available pre-1995. The expenditure data for the early years have been generously provided by Jim Brown of the Ohio Department of Education.

²⁰ For Indiana, the exact measure is total mean NCE (Normal Curve Equivalent) score, for Ohio, this is the percent of students scoring at or above the proficient level.

²¹ For the exact formulas and the advantages and disadvantages of each of these measures, see Murray et al (1998), or Litchfield (1999).

in my analysis. For comparison, I also put estimates for Kentucky (1990-2000) and court-mandated reforms in the last two columns.

There has been a large fall in each of the measures in Michigan after the reform. They look even more impressive when compared to the corresponding estimates in the last two columns, since for Michigan I am only looking at a seven year period.²² Because of the staggered nature of the program in Michigan the measures should continue falling for several more years.²³

An important feature of school financing in the U.S. that troubles lawmakers, judges and educators alike is the fact that local school expenditures depend largely on district property wealth. Many court cases, beginning with Serrano I in California (1971), have relied on this wealth-expenditure relationship as a yardstick of existing inequality. To estimate this gradient for Michigan and the effect of the program, if any, I next regress school spending in 1990, 1994, 1998 and 2001 on district median income.

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

D94, D98, etc. are the respective year dummies. MI is the median income of the district in 1989. β_0 measures the strength of the income-expenditure relationship in 1990, the first year in my analysis. β_1 gives the change in this gradient between 1990 and 1994, the last year before the program. β_2 and β_3 show post-program changes, if any - together with β_1 they give an idea of the effectiveness of the program 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 diluted between then and 1994. Post-reform, however, there

²² Part of the reduction may have occurred even without the reforms. However, estimates from Table 3 below suggest that pre-reform trends, though moving towards equalization, were not particularly large.

²³ I only report the weighted statistics. For the unweighted statistics, the fall is even greater (except in the Gini), presumably due to the fact that most of the lowest spending districts are relatively small (see Table 4).

has been a very significant weakening - in fact, for the unweighted estimates, the income-expenditure relationship is slightly negative in 2001.²⁴

To examine whether Proposal A actually equalized expenditures in Michigan I classify the 524 K-12 districts into 5 equal groups based on the 1993-94 level of spending (Group 1 consists of the lowest spending 105, and so on).²⁵ Some summary statistics on these groups of districts are shown in Table 4. The lowest spending districts are overwhelmingly white, mostly rural and small, and have a high proportion of free lunch students. They also perform significantly worse than the highest spending districts in 4th grade tests, though the differences from the other groups is not that large.

Figure 1 shows the spending distributions for Groups 1 and 5 in 1994 and 2001. The top panel shows changes in revenues, the lower one is for expenditures.²⁶ There seems to have been a significant convergence between these groups post-reform. To formally compare the trends in spending in these different groups, pre and post-reform, I next run the fixed effects (FE) regression

$$Y_{sgt} = \alpha + \alpha_s + \beta_0 * t + \sum_{g=1}^{g=4} \beta_g * (D_g * t) + \gamma_0 * (reform) + \sum_{g=1}^{g=4} \gamma_g * (D_g * reform) + \theta_0 * (reform * t) + \sum_{g=1}^{g=4} \theta_g * (D_g * reform * t) + \delta * X_{sgt} + \varepsilon_{sgt} \quad (2)$$

Y_{sgt} is the general fund revenue (GFR) or general fund expenditure (GFE) of district s in group g in year t . α_s is the district fixed effect while X_{sgt} are the time-varying characteristics.²⁷ D_g 's are the dummy variables for the respective groups. Group 5, comprising of the highest spending districts, is

²⁴ Similar large declines have been observed for some of the states with court-mandated reforms. Downes (1992) reports that in California the estimated coefficient on district income fell from 6.70 in 1977 to 1.00 in 1986. Clark (2003) finds that in Kentucky this fell from -4.38 to -38.27 between 1990 and 2000.

²⁵ Note that future increases in foundation allowances were monotonically related to this base level of spending. 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 - not surprising, since the correlations in spending between the two years is very high.

²⁶ These show the kernel smoothed plots of general fund expenditures in the two groups of districts. All figures have been weighted by district enrollment.

²⁷ Since free lunch data for 1990 and 1991 are either not available or not reliable, I have only included ethnic composition in X_{sgt} . Running the regression on a sub-sample when data on both controls are available does not change the qualitative results.

the omitted category. ‘Reform’ is a binary variable, taking the value 0 pre-reform and 1 afterwards. t is a time-trend, equal to 0 in 1994, 1 in 1995 and so on. In this specification, θ_g ’s give the differences in post-reform trends over and above the differences in β_g ’s, the pre-reform trends - they can thus be interpreted as a sort of a difference-in-differences estimate for trends.

I use annual data from 1990 to 2001, which straddle 1994, the last year before the reforms. The results are in Table 5.²⁸ The gap between the highest spending and the lowest spending districts had been increasing in the years before the program. In fact there was a clear hierarchy - the richer a group, the higher its growth rate. Post-reform however, this has completely reversed itself. The reform seems to have been instrumental in raising spending in Groups 1 and 2 in both absolute and relative terms. This result looks robust, and holds for both revenues and expenditures.²⁹³⁰

In order to see whether the increases in spending have been translated into real inputs, the last two columns of Table 5 show the corresponding trends for teacher salaries and student-teacher ratios. The trends for teacher salaries mirror those for revenues and expenditures - there seem to have been a significant narrowing of the gap between Group 1 and Group 5 districts post-reform. As for spending, this looks more impressive when compared to the divergent trends just prior to the program. For student-teacher ratios however, the lowest spending districts were improving in the pre-reform period too, and there is not much of a change following the program.

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

²⁸ The OLS results, not reported, are very similar. In addition, I experimented with a variety of samples, often excluding the very small and the very large districts. The results were similar.

²⁹ One thing to note is that in general, when revenues are increasing fast, expenditures often lag behind - conversely, limits on the revenue side may not constrain expenditures by as much. This is true in both the pre and post-reform period.

³⁰ The results are also similar if I consider another widely used measure of spending viz., current operating expenditures. The only thing to note is that for the lowest spending districts, current operating expenditures in the post-reform period lag behind general fund expenditures.

5 Effect on Test Scores

Though from a moral or a judicial point of view, differences in spending across schools are undesirable in themselves, the practical importance of school finance reforms stems from their likely effect on academic performance. If students in lowest spending districts are lagging behind others because of a lack of adequate resources, then equalization of spending brought about by programs like Michigan's can spur significant gains in achievement.³¹ 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 in all Michigan school districts.

5.1 MEAP

5.1.1 Trends in Test Scores across Different Groups of Districts

Figure 2 shows the distributions of 4th grade test scores (reading and mathematics) in Groups 1 and 5 in 1995 and 2001. There seems to have been some convergence post-reform, though it is less pronounced than that for spending. To formally estimate changes in student performance in post-reform Michigan³² I first employ two variants of equation (2) - one where I do not separately control for differences in pre-reform trends across groups, the other controlling for such differences. However, in most of these regressions, the restrictions involving linearity of the time trends are rejected, so I instead run regressions allowing for unrestricted year effects. I run three variants of these - first, with only a common pre-reform trend across groups, second, controlling for differences in pre-reform trends in a linear way, and third, running a completely non-linear specification. Because the results

³¹ Apart from the direct effect of increased resources, there may be an indirect effect operating through increases in school effort - see Appendix I for a theoretical digression of whether public school effort should go up following a finance reform.

³² As a first pass at the data, I regressed 4th grade test scores for various years on district median income from the 1990 census to estimate the income-test score gradient. The results are suggestive of a positive movement towards equality in the second half of the decade, however, it does not appear that the widespread success of the Michigan reforms in narrowing spending gaps was replicated in test scores.

are quite similar, I only report the results from the first specification, which is

$$T_{sgt} = \alpha_0 + \alpha_s + \beta_0 * t + \beta_{96,0} * D96 + \dots + \beta_{01,0} * D01 + \sum_{g=1}^{g=4} \beta_{96,g} * (D_g * D96) + \dots + \sum_{g=1}^{g=4} \beta_{01,g} * (D_g * D01) + \delta * X_{sgt} + \varepsilon_{sgt} \quad (3)$$

Here β_0 is the overall trend, while $\beta_{96,0}$, $\beta_{97,0}$, etc. give the year effects for the post-reform years. I am interested in the group-year interaction terms $\beta_{96,g}$'s, $\beta_{97,g}$'s, etc. - positive values of these for the lowest spending groups would be evidence that the program has helped to equalize test scores. Because from Table 4, it seems that the Group 5 districts were somewhat different from Groups 1 and 2, it is possible that they would have followed different paths even in the absence of any program. So instead of omitting Group 5 as earlier I now use Group 4 as my omitted category.³³

Table 6 reports the results. For ease of comparison, I show the estimated coefficients on spending in the first two columns.³⁴ For mathematics, there is some evidence for improvements in the later years. Though there is not a strict hierarchy the coefficients are suggestive of an improving trend. The same seems to be true for reading if one leaves out the sudden deterioration in 2000. In the immediate pre-reform period Group 1 had a negative trend for reading, controlling for this would make the estimates much higher.

The evidence is suggestive of an improved performance by the lowest spending districts later in the decade, though there are considerable differences from year to year. However, looking at trends in scores for a particular grade is not the ideal way to gauge gains in achievement - among other things, it cannot control for any idiosyncratic cohort-specific shocks across districts. I next present a cohort-based measure, which controls for this and is a proxy for arguably the best way of judging gains in achievement, viz., comparing test scores for the same student over time.

³³ When I assign the different districts percentile ranks based on their performance in 4th grade tests in 1994, the distribution of ranks within Groups 1 and 2 seem quite similar to that within Group 4.

³⁴ I show the results for two samples - the first includes all 524 districts, the second excludes Detroit. Exclusion of Detroit changes the result somewhat since Detroit belongs to Group 4, which is now the omitted category or control group.

5.1.2 Cohort-based Analysis of Test Score Gains

In Michigan, students are tested in reading and mathematics both in grades 4 and 7. I now test for improvements in academic performance by looking at the experience of particular cohorts. For each group of districts and for each cohort, I first measure the improvement between grades 4 and 7. Then I compare the improvement in the lowest spending districts with that in the higher spending ones.

The idea is that if the reforms have had a direct effect on test scores, then cohorts in the lowest spending districts would exhibit higher growth rates, and these gains would be progressively higher as one moves onto more recent cohorts, which would have spent more years under the program.

I begin with the cohort that was in the 4th grade in 1993. These kids would be in grade 7 in 1996, the second year of the reforms, and hence would be affected for little more than a year.³⁵ The cohort in grade 4 in 1994 would similarly experience a little more than 2 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 (4th graders in 1996, 1997, etc.) that show the most improvement.

For each cohort, I run both OLS and FE regressions. For example, for the cohort in 4th grade in 1993 (and in 7th grade in 1996), the FE equation is

$$T_{sgt} = \alpha_0 + \alpha_s + \beta_0 * t + \sum_{g=1}^{g=4} \beta_g * (D_g * t) + \delta * X_{sgt} + \varepsilon_{sgt} \quad (4)$$

where T_{sgt} is the 4th grade test score for $t=1993$, and the 7th grade test score for $t=1996$. I run (4) for each cohort and record the coefficients β_1 and β_2 . As earlier, Group 4 is the omitted category.

The results are given in Table 7, which reports β_1 and β_2 for the 5 cohorts (4th graders in 1993,

³⁵ Ideally I should begin with a cohort unaffected by the program. However, data on the control variables are not available for the only previous year with 4th grade test score data (1991, mathematics) - further, inclusion of this cohort does not change any of the results qualitatively.

1994, 1995, 1996 and 1997).³⁶ An increasing trend in these coefficients can be construed as evidence in favor of the program.

For reading, the gains (between the 4th and 7th grades) for Groups 1 and 2 in the early years were significantly less than those of Group 4. But the differences narrowed down, and for the later years the coefficients are either mostly positive (Group 1) or only modestly negative (Group 2). The results for mathematics are similar - here the improvements seem to follow a rough trend, both for Groups 1 and 2. Figure 3 shows the estimated coefficients for the different cohorts, together with the 95% confidence interval bands.

These results on cohort-based gains are particularly interesting, since they seem to reflect what one would expect if one were to believe that ‘money matters.’ I now present some direct evidence on the effects of increased expenditures on test scores.

5.1.3 Instrumental Variables Estimates of Effect of Spending on Test Scores

In 5.1.1 and 5.1.2 I divided the Michigan K-12 districts into 5 groups, and compared the respective post-program evolutions. An alternative way to proceed for test scores would be as follows. Note that the “treatment” here takes the form of state-mandated foundation allowance (FA), which in turn is closely related to district revenues and expenditures.

The basic equation relating test scores (T_{st}) to district spending (GFE_{st} , general fund expenditures) can be written as ³⁷

$$T_{st} = \alpha_0 + \alpha_s + \beta_t + \phi * GFE_{st} + \delta * X_{st} + \varepsilon_{st} \quad (5)$$

If the coefficient ϕ for the post-reform years is positive and large one may assume the improvement in

³⁶ As earlier, I present estimates from two different samples. Since test scores for 1993 and 1994 are not available for Detroit, I instead show the results for a sample that omits the largest districts (districts with more than 800 4th grade students).

³⁷ α_s is the district fixed effect while β_t is the year effect

test scores to be causally related to spending. This in turn can be taken to be a measure of success for the program, since GFE is closely related to the actual treatment intensity (FA). However, OLS or FE on (5) may give inconsistent results. Consistency here requires district spending to be orthogonal to all the determinants of test scores excluded from this equation - an assumption whose veracity may be in doubt.³⁸

Use of instrumental variables provides one way out. In this case I can instrument district GFE by the state-mandated foundation allowance, which meets the two valid criteria for an instrument. First, as is seen later, the foundation allowances are very closely correlated with district expenditures. Second, since it is determined at the state level, FA should not be correlated with the unobserved determinants of test performance at the district level.

Note however that the state-mandated FAs, though following clear-cut rules set out in advance, are monotonic functions of school spending in the base year, 1994. Thus our maintained hypothesis, which is that FA should not affect test scores except through their effect on current expenditures, boils down to the following - 1994 spending should not affect test scores from 1996 onwards 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 4th grade test scores. So it is plausible to assume that any expenditure incurred by the district before the current 4th graders began attending school will only have negligible effects. That is, I can restrict attention to 4th grade scores from 1998 onwards - these 4th graders would be in grade 1 in 1995 or later, and should presumably not be affected much by 1994 expenditures (except in as much as it affects current spending). I report results for both periods, 1996-2001 and 1998-2001.

Another problem with using the changes in foundation allowances as instruments, where these

³⁸ This is true even for FE, which controls for all unobserved characteristics that do not vary over time.

depend on district spending in 1994, is the following. 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. Ideally, if there were enough non-linearities in the state aid formula one could control for any smooth (linear) effects of base spending while at the same time exploiting the sharp changes to estimate effects of spending on test scores. However, such a strategy would not have adequate power in Michigan where the discontinuities are not large enough. I instead curtail my sample, excluding the highest spending districts since they look somewhat different from the other groups.³⁹

As a further check, I use two different (but close) sets of instruments. The IV(1) estimates present results from regressions where I instrument the changes in spending with the changes in 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.

A final concern is that of discretionary spending by the state. School districts may lobby hard for extra funds and depending on their bargaining power some of them may indeed be successful. However, it seems reasonable to conclude that such spending would not be reflected in the foundation allowances, which are based on a fixed set of rules. Second, the magnitude of such spending can be expected to be more important from 2001-02 onwards when the passage of No Child Left Behind increased the stakes for states with a large number of failing schools.⁴⁰

Table 9 shows the results of the first-stage regressions for the IV estimates.⁴¹ As expected, foundation allowances are highly significant and seem to be the major determinant of per pupil expenditures in the post-reform period.

³⁹ As is shown later (Table 10), between the 1990 and 2000 censuses Group 5 evolved somewhat differently from the other groups.

⁴⁰ Since the passing standards in Michigan are higher than in most states, Michigan has a large number of ‘failing’ schools.

⁴¹ For brevity, I only show the results for two samples, and for the IV-FE(1) estimates, the other results are qualitatively similar.

Table 8 shows the results of the OLS and FE regressions of equation (5), and of the second stage regressions of the IV estimation.⁴² The upper half shows the results for reading, the lower half is for mathematics. Considering the OLS estimates, the effect of lagged spending is seen to be positively significant only in Panel A, which shows the unweighted estimates. The same is true for OLS-IV(1) estimates, which are qualitatively similar to the OLS results.⁴³

The FE estimates are quite close to the simple OLS estimates, particularly for 1996-2001. However, as often happens, the IV-FE estimates are larger and sometimes highly significant. These imply a modest effect of spending on performance. For example, for reading, the estimates in Panel C,⁴⁴ imply an increase of between 2 and 5 percentage points for every \$1000. With an in-sample standard deviation of about 14, this translates to about 0.15-0.40 standard deviations increase per \$1000 of extra spending.⁴⁵ For mathematics, the estimates in Panel C imply an increase of between 6 and 8 percentage points for every \$1000 - an increase of 0.40-0.55 in standard deviations.⁴⁶

To put these estimates into perspective, I note the results from some recent studies of state interventions in schooling. Krueger (1999) reports that smaller class sizes (about a third reduction) in the Tennessee STAR project resulted in about 0.28-0.22 standard deviations increase in test scores per year. Rouse (1998) reports that low-income, minority students who participated in the Milwaukee voucher program had annual effect sizes of the order of 0.08-0.12 standard deviations.⁴⁷ Guryan finds that in Massachusetts 4th grade test scores increased by about 0.50 standard deviation in average district scores due to an increase of about \$1000 in spending (Guryan, 2001). Hoxby (2002), who in her study of charter schools in Michigan uses the same MEAP data, finds that a significant threat of

⁴² All the nominal variables used below - general fund expenditures and foundation allowances - have been deflated using the consumer price index for Midwest Urban, available on the website of the Bureau of Labor Standards. The coefficients and their standard errors are in terms of 10^{-1} .

⁴³ For brevity, I do not report the OLS-IV(2) estimates.

⁴⁴ my preferred specification

⁴⁵ The standard deviations for reading scores for 1996-2001 and 1998-2001 are 14.32 and 13.63 respectively.

⁴⁶ The standard deviations for 1996-2001 and 1998-2001 are 14.88 and 13.31 respectively.

⁴⁷ This is only for mathematics, she does not find any statistically significant effects in reading.

charter competition increases 4th grade reading and mathematics scores in public schools annually by 0.23 and 0.07 standard deviations respectively.

My estimates imply an effect of between 0.15-0.40 standard deviations for reading and 0.40-0.55 standard deviations for mathematics. This is for an increase in real per pupil expenditure of \$1000 - typically even the lowest spending districts in Michigan would have seen such increase over a period of about four years.⁴⁸ Thus this boils down to roughly a quarter to well over half a standard deviation increase over four years. This is a modest to large effect.

To sum, though one can be a bit skeptical of the IV estimates because of the reasons mentioned above, the results generally suggest that the increases in spending led to some improvements in academic performance. Before I proceed further I perform some robustness checks.⁴⁹

5.1.4 Robustness of the Results

First I looked at the changes in the Michigan school districts between the two censuses. If over time the lowest spending districts are evolving differently from the rest, then it is likely that some of the improvement in test scores is due to this factor, rather than increases in school resources or school effort.⁵⁰ Table 10 shows the results from FE regressions of different socioeconomic characteristics on a time-trend t and t interacted with the 4 group dummies. Group 1 is now the omitted category. Overall, there is not much difference in the time-paths of the different groups, except Group 5. This is true both for demographic characteristics as well as economic ones. All districts witnessed modest to large increases in population, per capita income and in the proportion of adults with at least a bachelor's degree. The poverty and employment variables also changed favorably. But generally, the

⁴⁸ 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 \$1000.

⁴⁹ Most of these results are not reported, but they are available from the author on request.

⁵⁰ Note that I control for ethnicity and free lunch eligibility in my regressions - changes in these will mostly be accounted for. Further, some of the changes in the 2000 census may be a result of the reform itself, since choosing neighborhoods on the basis of desirable schools has long been commonplace in the U.S.

lowest spending districts do not seem to have done any better than Groups 3 and 4.

Second, I repeated my earlier exercises (in 5.1.1 and 5.1.2) using Group 5 as the control. The results are stronger - there were significant improvements by the lowest spending districts post-reform, particularly in later years (see Figure 2). When I similarly use Group 3 as the control, the results are somewhat diluted, but Group 1 districts still exhibit considerable improvements, particularly in the cohort-based analysis.

The rapid spread of charter schools in Michigan may have contaminated some of my results. Michigan instituted these reforms regarding school choice in the mid '90s, along with the sweeping changes in school financing.⁵¹ Many commentators believe that 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. In fact, Hoxby (2002) argues that in Michigan this is exactly what happened - districts that witnessed sprouting up of 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,⁵² they still serve only a fraction of overall K-12 students (see Table 11, 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, and given that charter schools mostly serve urban city children it is not surprising that they are located mostly in the richer districts.⁵³ Table 11 shows the growth of charter schools across the different groups in Michigan. For each of the years,

⁵¹ There is also a small inter-district choice program in Michigan - in 2000 and 2001 e.g., 17,440 and 26,024 students enrolled in public schools outside their home district (about 1% and 1.5% of Michigan's total public enrollment, respectively). See Arsen et al, 2001. A proposal to introduce vouchers was however defeated in the 2000 elections.

⁵² The proportion of Michigan public school students enrolled in charter schools is third highest in the nation, after Arizona and Washington D.C.

⁵³ The demand for such schools is presumably higher in urban areas where public schools are often seen to be 'failing.' Another reason may be that in Michigan most of the charters are granted by state universities, so some of the schools are located near these universities (Bettinger, 1999).

the percentage of students enrolled in charter schools is the smallest in the lowest spending districts and vice versa. In fact, if the story about increased competition from charter schools encouraging increased effort in public schools is indeed true, my results would be an underestimate - since I am netting out the improvement in the Group 4 districts as a common shock to all districts, I am discarding more test score gains in the lowest spending districts than I actually should.

To sum, controlling for factors like charter schools and selective migration across school districts does not seem to affect the results substantially. The same is true when I consider alternate comparison groups. The exact results are somewhat magnified or diluted but the overall picture is unchanged - there were significant improvements by the lowest spending districts post-reform, particularly in later years.

If valid, the above will be an important argument in favor of Michigan-like programs. 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 districts. However, two caveats should be mentioned. First, the '90s witnessed an across-the-board rise in public concern for K-12 education, and some of the effects seen so far may not be unique to Michigan. To get some idea of the strength of these forces I compare the Michigan experience to those in two neighboring states, Indiana and Ohio. Second, so far I have relied exclusively on the MEAP test scores 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 publicized⁵⁴ and there may be incentives for teachers and administrators to boost their test scores artificially - several such phenomena like "teaching to the test", transferring regular students to special categories, even outright cheating, have been noted by researchers for other states (see e.g. Figlio and Getzler (2002), Jacob and Levitt (2003)). So I supplement my analysis by looking at two tests that largely avoid

⁵⁴ though they are not literally 'high stakes' as in some other states, there are awards (Governor's Cup) for districts that turn in the best performances

the above pitfalls, NAEP and ACT, in section 5.3.

5.2 Experience of Neighboring States

I compare the Michigan experience to those of two neighboring states, viz. Indiana and Ohio, to check for any potential bias in the results.⁵⁵

5.2.1 Equalization of School Spending

First I briefly look at changes in school spending. Table 12 shows the trends in spending across different groups of districts in each state in the 1995-2000 period.⁵⁶ Michigan's superior performance is immediately evident. In Indiana and Ohio the lowest spending groups move mostly in line with their richer counterparts, or exhibit catch-up which is modest by Michigan's standards. It seems that Proposal A was responsible for at least a major part of the relative equalization of school finances seen in post-1995 Michigan.⁵⁷

5.2.2 Effect on Academic Performance

The next step is to look for changes in school performance across the three states. The first problem that arises in this context is ensuring comparability of the test results. The particular measures used - percent scoring at or above the satisfactory level (Michigan), percent scoring at or above the proficient level (Ohio) and mean NCE ⁵⁸ score (Indiana) - are different for different states.

⁵⁵ See Appendix II for why these two states were chosen.

⁵⁶ As in Michigan, I divide Indiana and Ohio districts into five groups each, based on their 1995 spending.

⁵⁷ I also ran OLS regressions of district spending on district median income (1990 census) to estimate the income-expenditure gradient in each state. For Indiana, which already had a negative gradient in 1990, this has fallen further between 1995 and 2000. However, part of this decline seems due to trends existing prior to 1995. For Ohio, the picture is similar to Michigan's - a large drop in the gradient for the unweighted estimates, though for the weighted estimates, the decline is relatively modest. However, part of the improvement may also be due to the changes enacted after a 1997 court ruling or to pre-existing trends. The change in Michigan looks particularly impressive in view of the fact that prior to 1994 there do not seem to have been any definite trends towards convergence.

⁵⁸ Normal Curve Equivalent

Unfortunately, either these are the only measures for which continuous data are available, or the other available measures are very similar.

The second problem is to ensure comparability of the respective groups. For example, unlike in Michigan and Ohio, the lowest spending group in Indiana is among the best performers in the state, while the higher spending groups lag behind.

For these reasons, even though I take efforts to create appropriate comparison groups in each state⁵⁹ the results are only suggestive. For brevity these are not reported, but are available from the author on request. To summarize, lowest spending school districts in Indiana and Ohio do not seem to exhibit test score gains that match those in Michigan. The results are similar whether I look at the trends in test scores for a particular grade, or perform a cohort-based analysis, focusing on the improvements between the 4th and 7th grades (Michigan), 3rd and 6th grades (Indiana) and 4th and 6th grades (Ohio). There is evidence of a modest mean-reversion in either state. But this is unlikely to explain much of the improvement in Michigan - first because the initial performance gap between the groups was quite small, and second, because the improvements there seem to follow a rough trend.

5.3 Performance in Nationwide Tests

Finally, I examine whether the gains seen so far show up in alternate measures of achievement. I show results for two important nationwide tests, viz. NAEP and ACT.

5.3.1 NAEP Performance

NAEP consists of tests conducted periodically on a nationwide basis. Michigan participated in all of the mathematics assessments conducted by NAEP since 1990 (1990, 1992, 1996 and 2000). It

⁵⁹ based on their average percentile ranks as well as base year spending

also participated in all the three reading assessments (1992, 1994 and 1998) - however, the minimum participating guidelines were not met in 1994.

I analyze the NAEP trends for Michigan, and compare them to the corresponding trends for Indiana, Ohio and the Nation.⁶⁰ Since the results for reading are available only for 1992 and 1998, which straddle 1995 - I confine myself to mathematics (4th and 8th grades).⁶¹ Moreover, since I am mostly interested in the bottom half of the distribution, I show the scores for selected categories - the average scores, those at the 10th, 25th and 50th percentiles, and the percentages scoring at or above the Basic and Proficient levels. I also include the scores for Black and Hispanic students.

Table 13 shows the change in scale scores in Michigan over and above that in the control state or region.⁶² One can get a rough idea of the effectiveness of the reform by comparing the change between 1996 and 2000 with the corresponding change between 1992 and 1996. Most of the 1996-2000 changes with respect to Indiana and the Nation are negative, as are the 1992-2000 changes relative to Ohio. For grade 8 the results weigh even more overwhelmingly against a relative improvement in the later half of the decade.

Thus, there does not seem to be much of a positive effect of the finance reform on NAEP scores.⁶³ However, many of the trends seem to reflect a move towards convergence than anything else.⁶⁴ This is also true for the individual ethnic groups - Michigan witnessed modest to large relative gains between 1992 and 1996 for Blacks but most of these were lost between 1996-00. The converse is true for the Hispanics.

⁶⁰ Ohio did not participate in the mathematics tests in 1996.

⁶¹ Even if there are gains for Michigan in reading between 1992 and 1998, it would be difficult to ascribe it to pre or post-reform with any degree of certainty.

⁶² For example, between 1992 and 1996, scale scores for grade 4 mathematics increased by 6 points in Michigan but only 3 points in the Nation. So Michigan's net improvement is +3 with respect to the Nation.

⁶³ If we consider Central as a comparison group (not shown) then Michigan does have a positive 'program effect'.

⁶⁴ For example, large net improvements for Michigan seem to be followed by net declines; whereas large relative declines, like for Hispanic between 1992 and 1996 seem to be followed by large improvements or small declines. The evidence for converging trends is all the more strong if I look at the 1990-92 changes for grade 8 (not shown) and compare those to the 1992-96 changes.

5.3.2 ACT Performance

Across the nation, most college-bound students appear in 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 if the reforms have had any effect on ACT taking and performance.

The underlying idea is that since students mostly take the ACT in their 12th grade, kids taking the exam in later years would have been exposed to the program for a larger number of years. For example, someone taking the exam in 1997 would have spent only about 2 years under the reforms, but someone taking the same exam in 2001 would have been exposed for 6 years. So beneficial effects of the finance reform, if any, would show up as gains for the more recent cohorts in the lowest spending districts.

I run OLS and FE regressions, allowing for unrestricted year effects. The FE equation is,

$$\begin{aligned}
 ACT_{sgt} = & \alpha + \alpha_s + \beta_{98,0} * D98 + \dots + \beta_{01,0} * D01 + \sum_{g=1}^{g=4} \beta_{98,g} * (D_g * D98) + \dots \\
 & + \sum_{g=1}^{g=4} \beta_{01,g} * (D_g * D01) + \delta * X_{sgt} + \varepsilon_{sgt} \quad (6)
 \end{aligned}$$

To control for any secular changes that may differentially affect the various groups, I use Indiana as the control state.⁶⁵ In Indiana, however, the more important college entrance test is SAT, not ACT. As earlier, I take Group 4 as the control in Michigan since in terms of participation and performance in the base year (1997), Group 4 is much closer to Groups 1 and 2 than Group 5. In Indiana too I use Group 4 as the control, since Group 5, the richest group, has a SAT participation rate that is significantly lower than those of Groups 1 and 2 and also scores well below other groups.⁶⁶

Table 14 presents the estimates from regressions of equation (6). The first three columns show the changes in ACT and SAT participation rates between 1997 and 2001. There is no evidence of

⁶⁵ Data on college prep tests are not available for Ohio, so here I focus only on Indiana and Michigan.

⁶⁶ The ACT participation rate of Group 5 districts, on the other hand, is much higher than those of Groups 1 and 2.

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 towards improvement. In Indiana, there is no evidence of any significant changes in SAT participation rates. For ACT participation rates, there seems to be an improving trend for the poorest districts. However, Group 1 districts had significantly lower rates of ACT participation in 1997, so part of the improvement is presumably a reversion to the mean.

The last three columns of Table 14 report the results for ACT and SAT scores.⁶⁷ The Michigan ACT results show no significant differences from Group 4. Though for Group 1 the coefficients in the later years are less negative, the differences are minimal and never significant. The same is true in Indiana, where neither for ACT nor for SAT is there any evidence of improved performance by the lowest spending districts.

To sum, it is somewhat puzzling that the improvement seen earlier for the lowest spending districts in Michigan have not yet been reflected in these nationwide tests. However, two caveats should be mentioned. First, the NAEP scores are compiled at an aggregate level, and this may mask important changes within districts or regions in a state. Though I have tried to use those indicators where improvement would have been most expected, this is only suggestive.⁶⁸ Second, it may be too early to expect sizeable improvements in ACT measures. Even the most recent cohort included here would have spent just about half of its school life under the reforms.⁶⁹

⁶⁷ For brevity I only show the results for SAT mathematics, the results for SAT verbal scores are similar.

⁶⁸ This is all the more so since the lowest spending districts are on average much smaller - in a sample of 100 Michigan students, for example, only about 24 would belong to Groups 1 and 2. Detroit, on the other hand, would alone account for 10 students.

⁶⁹ A recent study by Arizona State University (ASU) researchers has found that, in states with 'high-stakes' state tests, "while students show considerable improvement on .. state exams, the opposite is typically true of their performance on other, independent measures of academic achievement" (The New York Times, December 28, 2002). The 'other, independent measures' referred to here include NAEP, SAT and ACT tests. It is interesting that this may be true of a state like Michigan, which does not have a high-stakes testing regime. (The results of the ASU study have since been disputed, see e.g. Raymond and Hanushek, 2003).

6 Conclusion and Extensions

In this paper I have studied the Michigan school finance reforms. In 1994, somewhat unexpectedly and without the prodding of any courts, Michigan initiated a drastic overhaul of its school financing system. Among other things, Proposal A entailed large sums of money over the next few years to the lowest spending districts, and largely ended local control over total school spending. Using data for both the pre and post-reform periods, I find that the program has been quite successful in reducing inequalities of school spending. Though some inequalities exist, because of the staggered nature of the reforms, these are no longer considered a matter of significant concern.

The bulk of the paper is devoted to an analysis of academic performance in the years after the program. I tried to see whether improvements in test scores followed the large increase in resources witnessed by the lowest spending districts. Based on the MEAP tests conducted by the state department of education, it seems that there was some significant improvement in these districts, which had erstwhile been lagging behind. However, the improvements do not seem to have spilled over to other areas, like participation and performance in ACT or in NAEP (Mathematics). This conclusion remains largely unchanged when I look at the experience of two neighboring states, Indiana and Ohio.

These findings have significant policy implications. First, these show that state legislatures can initiate and implement a comprehensive school finance reform, even one which is largely redistributive in nature. Second, it is interesting to note the significant academic progress registered by the lowest spending districts in the post-reform period. However, and third, the gains in student achievement look relatively modest, particularly when compared to the large increases in spending. It seems that even complete equalization of school resources across districts will not be enough to ensure complete equality in school outcome measures. One may have to look beyond school financing to issues of school effort and favorable peer group quality.

A school finance reform has many effects, only some of which have been addressed in this paper. Among the other effects mentioned in the literature are the following. First, there are the possible effects on stratification across school districts ⁷⁰ - community compositions may change as a result of the reform, and may get reflected in housing prices. Second, the move to state-level financing (implicit consolidation of school districts) may reduce the productivity of public schools ala Hoxby (1996). Third, there are the effects on spending in general - the substitution of local control by state-mandated spending may lead to a decline in average school expenditures in the state as a whole. It may also lead to increases in private school enrollment, as in California.⁷¹

A related question is the effect of the reform on academic performance in the highest spending districts. These districts were constrained by Proposal A and as seen from Table 5, they witnessed a significant decline in their rates of growth of spending. It would be interesting to see whether the program had any adverse effect on this group. However, since Proposal A held these districts harmless, they only suffered a decline in their *rates of growth of spending*, not in the absolute spending levels. So any adverse effects would show up in the rates of change in performance. It seems that this would ideally need a longer time-series to say anything definite.

One important question in Michigan is the sustainability of this reform. The most debated area is that of local control - the richer districts can be expected to exert considerable pressure over the years if they are unable to spend as much as desired. This would be particularly pressing if, like California, there is a drop in overall support for education. As of now, however, the program is quite popular.

⁷⁰ For example, some people with high preferences for schooling who erstwhile had been living in richer districts may now move back - one benefit of living in high property-value districts is no longer operative. On the other hand, because of the staggered nature of the reforms people may in fact move to the highest spending districts, which will continue to have higher spending for a while.

⁷¹ One interesting point to note in this respect is the recent debate on whether the presence of elderly may negatively impinge on expenditures on schooling. Poterba (1997) argues, using state level data, that this has indeed been the case in the U.S. Harris et al (2001) concur, but they find that the negative effects are more significant at the state level, and only modest at the district level. With an imminent rise in the share of the elderly, this should be of some concern to states with centralized school financing, where the elderly and others cannot sort themselves into districts based on their preferred levels of school spending.

Appendix I: Whether Public School Effort should go up after the Reforms

I briefly discuss whether one should expect public school effort to go up after the finance reforms. Unfortunately, there is no unanimity in the literature as how to model public school behavior, so I consider different possible specifications.

First, following the school choice literature, I model public schools as being ‘rent maximizers’. Formally, I assume that they choose effort level e to maximize their rent R , which I define as net revenues, or revenues minus costs

$$R = p * N(e) - C(N, e)$$

Here N is the enrollment of the public school district, assumed to depend positively on public school effort e (but N is concave in e). p is the per pupil spending and C is the cost function of the public school, assumed convex in both N and e .

The first-order condition for rent maximization is $(p - C_N) * N(e) = C_e$, which determines the optimal public school effort.⁷² (We assume here that $p > C_N$, i.e. enrolling the marginal student is always profitable for the public school.)

The effects of a school finance reform can be modeled as an increase in p , at least for the lowest spending districts. Differentiating the first-order condition, I get

$$\frac{de}{dp} = \frac{N_e}{\{-(p - C_N^1) * N_{ee} + C_{NN}^1 * N_e^2 + C_{ee}^2\}}$$

which is positive if $p > C_N^1$. The explanation is intuitive - an increase in p increases the payoff from the marginal effort exerted by the district, and it is profitable for the school to increase e to attract more students. Note however that the result depends crucially on $N_e \neq 0$. In Table 4 I saw that

⁷² The second order condition is satisfied if C_{Ne} is either positive or not too negative. This will hold if e.g. $C(N, e)$ is separable in N and e - for simplicity I assume $C(N, e) = C^1(N) + C^2(e)$ in what follows.

in Michigan most of the lowest spending districts are located in rural areas, where transportation costs may make N quite inelastic with respect to e . In this case one would not expect e to change by much following the reform.

Conversely, if the highest spending districts are constrained in their spending, this will be equivalent to a fall in p , and may lead to diminished effort. It is conceivable that there will be a leveling down of school effort - not much change in the lowest spending districts, but a significant decline in the richer ones, for whom $N(e)$ may be quite elastic.⁷³ Another channel by which the same thing can happen is if higher public school effort is capitalized in rising house prices through an increase in test scores. This will increase the budget of the school, which is financed primarily through property taxes under local finance.⁷⁴ In other words, under local finance p may be an increasing function of effort level e . Then the optimal effort would solve $(p - C_N^1) * N_e = C_e^2 - p_e * N$ - as is intuitively obvious, if $p_e = 0$ under state finance, then the public school exerts a lower effort, *ceteris paribus*. The extra benefit of higher subsequent expenditures is no longer relevant when the state equalizes all school budgets.

Another possible scenario is where schools are constrained in their expenditures. Assume that effort is costly, but the public school has to offer a minimum quality to their customers, even with budgetary restrictions. Formally, let total school expenditures E consist of local support (L) and state aid (S), both assumed given, and let quality Q be a positive function of effort e and expenditures E , $Q = Q(e, E)$. If the threshold quality level that can be offered is Q_{min} , then an increase in S will lead to a decline in public school effort. Intuitively, with an increase in state aid the school can afford to cut back on its effort, which it will do so since effort is costly.

Finally, a variant of this where public schools will increase effort is where they are maximizing quality net of costs, i.e. where they are choosing e to maximize $Q(e, E) - C(e)$, given E . It

⁷³ A program of school choice, such as charter schools in Michigan, will make $N(e)$ even more elastic.

⁷⁴ This assumes that if house prices rise, residents will just bear the higher school expenditures rather than giving themselves a tax break. This is not obvious - one needs a fuller model to work out the implications.

follows immediately that $\frac{de}{dE} = \frac{-Q_{eE}}{Q_{ee} - C_{ee}} > 0$ if e and E are complements in production, as is likely. Intuitively, even otherwise motivated teachers may be hampered by larger class sizes, antiquated equipments, etc. - it is common to hear teachers and administrators complaining of lack of resources affecting productivity.

Appendix II: Choosing Control States for Michigan

I confine myself to the Midwestern states neighboring Michigan, viz., Indiana, Illinois, Ohio, Minnesota and Wisconsin. The obvious criteria for selecting control states among these five are first, the absence of any important educational initiatives during the '90s, particularly the second half, and second, the availability of test score data similar to Michigan's. Based on these, Indiana and Ohio fit the bill best. There have not been any major school-related programs in either state during this period.⁷⁵ Further, both states have got test score data in grades and subjects similar to Michigan's, Indiana also has district level data on both SAT and ACT tests from 1997.

As for the other states, Wisconsin changed in its school finance system in 1997, moving to a three-tier equalization aid formula from a two-tier one. Further, the state switched to a new set of tests in 1997-98, so that the pre and post-1997 scores are not strictly comparable. For Minnesota too, test score data are available only for the later years - 3rd and 5th grade tests in Reading and Math were administered only from 1998 onwards. Finally, Illinois instituted major changes in its school finance programs in spring and winter of 1997. Illinois also looks different from Michigan in the organization of school districts and in ethnic composition of its students, see Table A-1(a).

Table A-1 summarizes the district and student characteristics in these five states, together with their performance in SAT, ACT and NAEP. Ohio and Indiana do not look much different from Michigan, except that in Indiana the more popular college prep test is SAT.

⁷⁵ Though following an adverse court ruling in 1997, Ohio had tried to explicitly tie its base funding levels to spending in districts in which students scored well on state tests. The only other exception is the school voucher program in Cleveland which began in 1997. Since there are more than 600 school districts in Ohio, omitting Cleveland is not a big concern. In any case, including or excluding Cleveland does not change the qualitative results.

References

- Addonizio, M. F., C. P. Kearney and H. J. Prince (1995), "Michigan's High Wire Act", *Journal of Education Finance*, Volume 20, Winter, 235-269.
- Arsen, David, David N. Plank and Gary Sykes (2001), "A Work in Progress", *Education Next*, Winter, available on the web at www.educationnext.org
- Benabou, Roland (1996a), "Equity and Efficiency in Human Capital Investment: The Local Connection", *Review of Economic Studies*, Vol. 63, April, 237-264.
- Benabou, Roland (1996b), "Heterogeneity, Stratification and Growth: Macroeconomic Implications of Community Structure and School Finance", *American Economic Review*, Volume 86 (3), June 1996, 584-609.
- Bettinger, Eric (1999), "The Effect of Charter Schools on Charter Students and Public Schools, mimeo, November, available on the web at www.ncspe.org.
- Card, David and A. Abigail Payne (2002), "School Finance Reform, the Distribution of School Spending, and the Distribution of Student Test Scores", *Journal of Public Economics*, Volume 83, 49-82.
- Clark, Melissa A. (2003), "Education Reform, Redistribution, and Student Achievement: Evidence from the Kentucky Education Reform Act", mimeo, Princeton University, January.
- Courant, Paul N., Edward M. Gramlich and Susanna Loeb (1995), "Michigan's Recent School Finance Reform: A Preliminary Report", *American Economic Review*, Vol. 85 (2), May, 372-377.
- Courant, Paul N. and Susanna Loeb, "Centralization of School Finance in Michigan" (1997), *Journal of Policy Analysis and Management*, Volume 16 (1), Winter, 114-136.
- Downes, Thomas A. (1992), "Evaluating the Impact of School Finance Reform on the Provision of

Public Education: The California Case”, *National Tax Journal*, Volume XLV, No. 4, December, 405-419.

Downes, Thomas A. (2002), “School Finance Reform and School Quality: Lessons from Vermont”, mimeo, Tufts University, March.

Evans, William N., Sheila E. Murray and Robert M. Schwab (1997), “School Houses, Court Houses and State Houses after Serrano”, *Journal of Policy Analysis and Management*, Volume 16 (1), Winter, 10-31.

Feldstein, Martin (1975), “Neutrality and Local Choice in Public Education,” *American Economic Review*, Vol. 115, No. 1, March.

Fernandez, Raquel and Richard Rogerson (1996), “Income Distribution, Communities and the Quality of Public Education”, *The Quarterly Journal of Economics*, Vol. 111, 135-164.

Fernandez, Raquel and Richard Rogerson (1999), “Equity and Resources: An Analysis of Education Finance Systems”, NBER Working Paper 7111, May.

Figlio, David N. and Lawrence S. Getzler (2002), “Accountability, Ability and Disability: Gaming the System”, NBER Working Paper 9307, November.

Guryan, Jonathan (2001), “Does Money Matter? Regression Discontinuity Estimates from Education Finance Reform in Massachusetts”, NBER working Paper 8269, May.

Hanushek, Eric (1991) “When School Finance ‘Reform’ may not be a Good Policy”, *Harvard Journal on Legislation*, Volume 28, 423-456.

Harris, Amy Rehder, William N. Evans and Robert M. Schwab (2001), “Education Spending in an Aging America”, *Journal of Public Economics*, Volume 81(3), September, 449-472.

Hoxby, Caroline M. (1996), “Are Efficiency and Equity in School Finance Substitutes or Complements?”, *Journal of Economic Perspectives*, Vol. 10(4), Fall, 51-72.

- Hoxby, Caroline M. (2001), "All School Finance Equalizations Are Not Created Equal", *The Quarterly Journal of Economics*, vol. 116(4), 1189-1231.
- Hoxby, Caroline (2002) "School Choice and School Productivity (or Could School Choice be a Tide that Lifts All Boats?)", NBER WP 8873, April.
- Jacob, Brian and Steven D. Levitt (2003), "Rotten Apples: An Investigation of the Prevalence and Predictors of Teacher Cheating", *Quarterly Journal of Economics*, Vol. 118(3), August, 843-877.
- Krueger, Alan B. (1999), "Experimental Estimates of Education Production Functions", *Quarterly Journal of Economics*, May 1999, 497-532.
- Litchfield, Julie A. (1999), "Inequality: Methods and Tools", March, available on the World Bank website at <http://www.worldbank.org/poverty/inequal/methods>.
- Murray, Sheila E., William N. Evans and Robert M. Schwab (1998), "Education Finance Reform and the Distribution of Education Resources", *American Economic Review*, Volume 88 (4), September, 789-812.
- Poterba, James M. (1997), "Demographic structure and the Political Economy of Public education", *Journal of Policy Analysis and Management*, Vol. 16(1), 48-66.
- Raymond Margaret E. and Eric A. Hanushek (2003), "Shopping for Evidence Against School Accountability", mimeo, March, available on the web at www.educationnext.org.
- Rouse, Cecilia (1998), "Private School Vouchers and Student Achievement: An Evaluation of the Milwaukee Parental Choice Program", *Quarterly Journal of Economics*, Vol. 113(2), May, 553-602.

Table 1: Increases in Foundation Allowances, post-reform Michigan
(Selected districts at different points of pre-reform spending distribution)

District	1994	1995	1996	1997	1998	1999	2000	2001
Standish Sterling	3738	4200	4506	4816	5124	5170	5700	6000
Delton-Kellogg	4501	4740	4988	5235	5462	5462	5700	6000
Kearsley Community	5008	5227	5380	5535	5689	5689	5927	6227
Carman-Ainsworth	6002	6181	6334	6489	6643	6643	6881	7182
Warren Woods	7069	7239	7392	7547	7701	7701	7824	7997
Grant Township	10,681	10,841	10,994	11,149	11,303	11,303	11,484	11,737
FA in Grant Town/ FA in Standish-Sterling	286	258	244	231	221	219	201	196

Source: Michigan Department of Education.

Table 2: Summary Measures of Inequality in Michigan before and after Proposal A

	General Fund Expenditures			Estimates	Estimates
	Michigan			from Kentucky	from court-ordered reforms
	1994	2001	% change 1994-2001	% change 1990-2000	% change 10 years after
$\text{Ln}(95^{th}/5^{th})$	0.562	0.395	-29.72	-32.80	-16
Gini Coefficient	0.104	0.084	-19.23	-28.68	-19
Theil Index	0.017	0.011	-35.29		-35
Coefficient of Variation	0.190	0.151	-20.53	-33.33	-17

$\text{Ln}(95^{th}/5^{th})$ 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 results for Kentucky come from Clark (2003). The last column shows the estimated reductions in the different measures 10 years after a court-mandated reform (Murray et al, 1998).

Table 3: Income-Expenditure Gradient, Michigan, 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	2079	2079	2075
R-squared	0.74	0.82	0.80

See equation (1) in the text. Column (1) shows the un-weighted estimates, while columns (2) and (3) are weighted by respective district enrollment. 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. +, *, ** denotes significance at the 10%, 5%, and 1% levels respectively.

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	93.90	91.38	93.40	57.28 (83.75)	81.64 (84.42)
Blacks	1.97	3.68	2.06	36.33 (10.59)	13.43 (10.60)
Hispanics	2.47	2.21	2.50	3.90 (3.25)	1.27 (1.23)
Asians	0.63	0.72	0.94	1.46 (1.44)	2.57 (2.79)
Free Lunch Eligibility	23.74	18.67	16.30	32.38 (17.87)	15.50 (13.82)
Average District Size	1788	1912	2323	5641	4224
Average School Size	457	458	493	558 (514)	520 (500)
Proportion Rural	75.62	68.64	56.03	16.15 (26.67)	8.80 (9.98)
4th grade Reading (MEAP)	43.50	44.64	44.47	41.11 (42.38)	48.60 (50.41)
4th grade Math (MEAP)	47.59	50.92	51.34	47.04 (49.01)	54.92 (57.10)

For Groups 4 and 5, the figures in parentheses correspond to the statistics when I leave out the 3 and 2 most populous districts in those groups respectively. All figures have been weighted by enrollment of the districts in 1994.

Table 5: Pre and Post-reform Trends in Spending, Michigan - FE Regressions, 1990-2001

	General Fund		General Fund		Teacher	Student-Teacher
	Revenues		Expenditures		Salaries	Ratios
	(1)	(2)	(1)	(2)	(1)	(1)
Trend (t)	346** (18.63)	360** (18.34)	281** (17.23)	316** (17.94)	2224** (134.50)	0.64** (0.11)
Group 1 * t	-134** (20.24)	-144** (19.85)	-54** (18.78)	-86** (19.16)	-564** (171.84)	-0.29* (0.13)
Group 2 * t	-98** (21.45)	-115** (22.53)	-34+ (19.64)	-73** (21.60)	-426** (165.91)	-0.31* (0.12)
Group 3 * t	-86** (21.15)	-100** (20.12)	-38* (19.23)	-73** (19.69)	-384* (163.16)	-0.15 (0.12)
Group 4 * t	-71* (28.59)	-64** (21.95)	-11 (37.53)	-56** (21.51)	-130 (276.26)	0.29 (0.18)
Reform * t	-124** (20.79)	-139** (20.31)	-33+ (19.60)	-71** (20.49)	-1590** (157.97)	-0.90** (0.11)
Group 1 * reform * t	251** (22.77)	261** (22.43)	135** (21.66)	166** (22.36)	892** (196.29)	0.09 (0.14)
Group 2 * reform * t	151** (24.02)	172** (25.15)	68** (22.57)	112** (24.96)	761** (194.14)	0.25+ (0.13)
Group 3 * reform * t	88** (23.50)	105** (22.44)	17 (21.83)	59** (22.58)	491** (188.78)	0.13 (0.13)
Group 4 * reform * t	105** (33.43)	67** (24.59)	24 (43.75)	49* (24.72)	-10 (350.41)	-0.14 (0.23)
Observations	6269	5981	6269	5981	6266	6266
R-squared	0.96	0.96	0.96	0.96	0.87	0.59

See equation (2) in the text. Group 5, the richest group of districts, is the omitted category. Columns marked (1) include all 524 districts, (2) excludes the 24 districts with 1994 enrollment more than 10,000. All regressions are weighted by district enrollment and control for ethnic composition. For brevity I do not report the other coefficients. Robust standard errors are in parentheses. +, *, ** denotes significance at the 10, 5, and 1 percent levels respectively.

Table 6: Effect of Reform on Spending and Grade 4 Reading and Mathematics Scores, Michigan - 1993-2001
(FE Regressions with unrestricted year effects)

	GF Expenditures		Reading		Mathematics	
	(1)	(2)	(1)	(2)	(1)	(2)
Group 1 * 1996	-224 (153.12)	10 (57.63)	-2.13 ⁺ (1.17)	-1.87 (1.23)	-1.63 (1.32)	-2.35 ⁺ (1.33)
Group 2 * 1996	-235 (152.10)	10 (55.43)	-0.91 (1.16)	-0.54 (1.22)	0.47 (1.24)	-0.13 (1.25)
Group 1 * 1997	61 (58.83)	73 (52.67)	-3.97 ^{**} (1.42)	-2.46 [*] (1.05)	-3.16 ^{**} (1.07)	-2.89 [*] (1.16)
Group 2 * 1997	46 (54.05)	65 (47.35)	-2.98 [*] (1.48)	-1.26 (1.12)	-0.86 (1.11)	-0.43 (1.19)
Group 1 * 1998	203 ^{**} (59.41)	200 ^{**} (52.39)	-1.94 ⁺ (1.14)	-1.43 (1.19)	-0.30 (1.40)	1.00 (1.23)
Group 2 * 1998	180 ^{**} (60.31)	191 ^{**} (53.37)	0.70 (1.13)	1.35 (1.18)	1.05 (1.38)	2.50 [*] (1.20)
Group 1 * 1999	133 ^{**} (55.49)	207 ^{**} (47.62)	2.25 (1.73)	1.18 (0.99)	1.55 (1.06)	1.84 ⁺ (1.12)
Group 2 * 1999	90 ^{**} (58.93)	179 ^{**} (51.75)	4.65 [*] (1.79)	2.73 ^{**} (1.09)	2.58 [*] (1.07)	3.02 ^{**} (1.14)
Group 1 * 2000	289 ^{**} (59.34)	309 ^{**} (51.90)	-5.91 ^{**} (1.33)	-4.41 ^{**} (1.07)	0.56 (1.15)	0.84 (1.13)
Group 2 * 2000	147 ^{**} (62.25)	186 ^{**} (55.43)	-2.19 (1.40)	-0.52 (1.14)	0.86 (1.16)	1.34 (1.12)
Group 1 * 2001	424 ^{**} (72.28)	374 ^{**} (61.96)	2.43 [*] (1.24)	1.21 (1.11)	3.53 ^{**} (1.35)	2.19 ⁺ (1.32)
Group 2 * 2001	228 ^{**} (79.55)	201 ^{**} (69.57)	2.45 [*] (1.46)	1.37 (1.35)	2.69 [*] (1.35)	1.52 (1.32)
Observations	6269	6257	4606	4599	4606	4599
R-squared	0.96	0.96	0.85	0.85	0.88	0.88

See equation (3) in the text. Group 4 is the omitted category. Columns marked (1) include all 524 districts, while (2) excludes Detroit. For brevity, I do not report the other coefficients. All regressions are weighted and control for ethnicity, gender, enrollment and free lunch eligibility. Robust standard errors are in parentheses. ⁺, ^{*}, ^{**} denote significance at the 10, 5, and 1 percent levels respectively.

Table 7: Cohort-based Analysis for Groups 1 and 2, Michigan - FE Regressions
(Reading and Mathematics, 4th and 7th grades)

	Group 1 Cohorts		Group 2 Cohorts	
	(1)	(2)	(1)	(2)
Reading				
4 th grade in 1993 (7 th grade in 1996)	-4.10** (1.57)	-3.72* (1.59)	-4.59** (1.60)	-4.15** (1.66)
4 th grade in 1994 (7 th grade in 1997)	-5.02** (1.36)	-5.37** (1.47)	-3.54** (1.38)	-3.55* (1.50)
4 th grade in 1995 (7 th grade in 1998)	-1.04 (1.35)	-0.65 (1.41)	-0.39 (1.61)	0.05 (1.66)
4 th grade in 1996 (7 th grade in 1999)	1.61 (1.62)	1.37 (1.72)	0.10 (1.66)	0.07 (1.71)
4 th grade in 1997 (7 th grade in 2000)	-0.68 (1.76)	-0.34 (1.91)	-1.78 (1.76)	-1.60 (1.87)
4 th grade in 1998 (7 th grade in 2001)	1.98 (1.38)	1.15 (1.35)	-0.53 (1.45)	-1.22 (1.42)
Mathematics				
4 th grade in 1993 (7 th grade in 1996)	-0.79 (1.82)	0.05 (1.90)	-2.55 (1.81)	-1.64 (1.91)
4 th grade in 1994 (7 th grade in 1997)	-0.69 (2.01)	-0.76 (2.00)	-1.92 (1.96)	-1.85 (1.97)
4 th grade in 1995 (7 th grade in 1998)	1.35 (1.70)	2.54 (1.62)	0.04 (2.15)	1.20 (2.08)
4 th grade in 1996 (7 th grade in 1999)	4.47** (1.82)	4.67** (1.87)	2.33 (2.00)	2.57 (2.00)
4 th grade in 1997 (7 th grade in 2000)	4.53* (1.80)	4.62* (1.71)	1.94 (2.09)	1.91 (1.93)

See equation (4). The table reports the values of β_1 and β_2 for the different cohorts (a separate regression is run for each cohort). Group 4 is the omitted category. Columns marked (1) include all 524 districts, while (2) includes only those with less than 800 4th grade test takers for the particular cohort. 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. +, *, ** denote significance at the 10, 5, and 1 percent levels.

Table 8: Effect of General Fund Expenditure (GFE) on grade 4 MEAP Scores, Michigan
(Reading and Mathematics, 1996-2001 and 1998-2001)

	1996-2001					1998-2001				
	OLS	OLS-IV(1)	FE	FE-IV(1)	FE-IV(2)	OLS	OLS-IV(1)	FE	FE-IV(1)	FE-IV(2)
Reading										
Panel A										
lag GFE	0.021** (0.006)	0.013+ (0.008)	0.018* (0.010)	0.022 (0.021)	0.024 (0.021)	0.027** (0.008)	0.027** (0.011)	-0.003 (0.011)	0.040 (0.027)	0.046+ (0.026)
R ²	0.40	0.40	0.74	0.50	0.50	0.23	0.23	0.73	0.14	0.14
Panel B										
lag GFE	0.000 (0.008)	-0.010 (0.009)	0.000 (0.016)	0.056** (0.021)	0.034 (0.035)	-0.003 (0.010)	-0.010 (0.013)	-0.028 (0.021)	0.037 (0.037)	0.029 (0.049)
R ²	0.51	0.51	0.82	0.49	0.50	0.46	0.46	0.84	0.10	0.12
Panel C										
lag GFE	0.005 (0.005)	-0.009 (0.008)	0.004 (0.008)	0.022 (0.018)	0.034+ (0.020)	0.009 (0.007)	-0.002 (0.012)	-0.001 (0.011)	0.053+ (0.026)	0.054* (0.027)
R ²	0.56	0.55	0.84	0.57	0.57	0.46	0.46	0.83	0.13	0.11
Mathematics										
Panel A										
lag GFE	0.033** (0.006)	0.034** (0.009)	0.033** (0.009)	0.070** (0.023)	0.083** (0.023)	0.030** (0.008)	0.028* (0.011)	0.005 (0.011)	0.049+ (0.029)	0.077** (0.028)
R ²	0.38	0.38	0.74	0.50	0.49	0.17	0.17	0.74	0.05	0.04
Panel B										
lag GFE	0.002 (0.007)	-0.005 (0.009)	-0.003 (0.014)	0.053* (0.021)	0.040 (0.032)	-0.006 (0.010)	-0.017 (0.012)	-0.038+ (0.022)	-0.012 (0.031)	-0.021 (0.040)
R ²	0.57	0.57	0.85	0.57	0.58	0.49	0.49	0.87	0.13	0.13
Panel C										
lag GFE	0.012* (0.006)	0.005 (0.008)	0.011 (0.008)	0.059** (0.019)	0.082** (0.021)	0.009 (0.007)	-0.002 (0.011)	0.001 (0.010)	0.060 (0.025)	0.062* (0.026)
R ²	0.53	0.53	0.84	0.58	0.56	0.40	0.40	0.85	0.09	0.06

Lag GFE is GFE lagged one year. All coefficients are in terms of 10^{-1} . Panels A and B include all 524 school districts, Panel C excludes Detroit. Panels B and C are weighted by respective district enrollments. All regressions include separate year dummies and control for racial and gender compositions and free lunch eligibility. There are 2507 (1996-2001) and 1673 (1998-2001) observations each for Panels A and B, and 2501 and 1669 observations respectively for Panel C. Robust standard errors are in parentheses. +, *, ** denote significance at the 10, 5, and 1 percent levels respectively.

Table 9: Effect of Foundation Allowances on General Fund Expenditures, Michigan (IV First Stage Regressions)

	Panel A				Panel C			
	1996-2001		1998-2001		1996-2001		1998-2001	
Foundation Allowance (FA)	0.76**	(0.07)	0.91**	(0.12)	0.72**	(0.07)	0.73**	(0.12)
FA * 1997	0.10 ⁺	(0.05)			0.07	(0.06)		
FA * 1998	0.16**	(0.06)			0.10	(0.06)		
FA * 1999	0.13	(0.08)	-0.03*	(0.07)	0.01	(0.09)	-0.10	(0.07)
FA * 2000	0.01	(0.02)	-0.07	(0.07)	0.02	(0.02)	-0.07	(0.07)
FA * 2001	0.23 ⁺	(0.12)	0.04	(0.10)	0.19	(0.13)	0.05	(0.12)
Observations	2511		1676		2501		1669	
F-statistic of joint significance of excluded instruments	29.62		17.04		31.71		11.98	

For brevity, I only report the results for the FE-IV(1) regressions. Panel A includes all 524 districts, Panel C excludes Detroit. Panel C results are weighted by number of test-takers in the 4th grade in the district. All regressions include separate year dummies and control for racial and gender compositions and free lunch eligibility. Robust standard errors are in parentheses.

Table 10: Changes in Selected Demographic and Economic Characteristics, Michigan School Districts (1990 and 2000 Censuses, FE Regressions)

	Number of Persons	Per Capita Income	Persons in Poverty	Children in Poverty	Unemployment Rate	Adults who are HS grads	Adults with BA degree
Time trend (t)	699* (327.43)	7367** (255.29)	-3.63** (0.33)	-4.50** (0.56)	-2.98** (0.33)	-0.40 (0.44)	3.69** (0.41)
Group 2 * t	906 ⁺ (501.02)	-30 (325.27)	1.14* (0.48)	1.43 ⁺ (0.72)	0.67 (0.46)	-0.70 (0.59)	0.79 (0.51)
Group 3 * t	1571 ⁺ (675.64)	262 (303.63)	1.17** (0.45)	0.96 (0.74)	0.58 (0.39)	-0.59 (0.60)	0.45 (0.52)
Group 4 * t	2437 ⁺ (1388.09)	286 (432.00)	0.89 (1.07)	-0.59 (1.72)	-0.05 (1.06)	0.21 (0.87)	0.65 (0.90)
Group 5 * t	2875 (2069.14)	1656** (514.48)	3.37** (0.48)	3.98** (0.71)	1.35** (0.41)	-1.23** (0.54)	3.11** (0.83)

Group 1 is the omitted category. All regressions are weighted by the number of persons in the district in 1990. There are 523 districts in each year, 1990 and 2000. Robust standard errors are in parentheses. ⁺, *, ** denote significance at the 10, 5, and 1 percent levels.

Table 11: 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

Author's calculations. In Michigan, charter schools are attached to an intermediate school district (ISD) rather than to individual school districts. The table shows the average ISD-wide percentages of charter students across various groups, e.g., 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 et al (2001).

Table 12: Trends in School Spending, Michigan, Indiana and Ohio, 1995-2000 - FE Regressions

	Michigan		Indiana		Ohio	
	(1)	(2)	(1)	(2)	(1)	(2)
Trend (t)	245** (10.22)	234** (9.19)	194** (11.31)	176** (10.90)	326** (23.12)	334** (23.70)
Group 1 * t	79** (11.48)	90** (10.43)	21 (39.62)	31+ (17.51)	12 (15.79)	3 (14.29)
Group 2 * t	42** (11.92)	52** (10.95)	9 (17.65)	30+ (17.24)	-2 (15.38)	-10 (14.93)
Observations	3642	3628	1746	1740	3626	3614
R-squared	0.95	0.96	0.93	0.92	0.94	0.94

For brevity I do not report the other coefficients. Columns marked (1) include all districts (except Detroit in Michigan). Columns marked (2) exclude Flint and Grand Rapids in Michigan (in addition to Detroit), Indianapolis in Indiana, and Cleveland and Columbus in Ohio. All regressions are weighted by enrollment and control for ethnicity. Robust standard errors are in parentheses. +, *, ** denote significance at the 10, 5, and 1 percent levels respectively.

Table 13: Changes in Michigan NAEP Scores net of the changes in Nation, Indiana and Ohio, 1992-2000

	4th grade Mathematics					8th grade Mathematics				
	Nation		Indiana		Ohio	Nation		Indiana		Ohio
	1992-96	1996-00	1992-96	1996-00	1992-00	1992-96	1996-00	1992-96	1996-00	1992-00
Average Scale Score	+3 (3.17)	+1 (2.87)	-2 (2.56)	0 (2.42)	-1 (2.82)	+6 (4.27)	-2 (4.17)	+4 (2.89)	-6 (3.12)	-4 (3.01)
Average Scale Score - Black	+5 (5.48)	-3 (4.82)	+3 (5.82)	-8 (5.21)	+2 (5.64)	+8 (4.80)	-8 (5.20)	+10 (5.25)	-17 (5.72)	-11 (5.38)
Average Scale Score - Hispanic	-5 (4.54)	-1 (5.42)	-6 (4.89)	0 (6.51)	-6 (6.42)	-5 (6.38)	+8 (6.45)	-4 (8.82)	0 (8.72)	-14 (8.38)
Percent at or above Basic	+2 (3.38)	-1 (3.20)	-5 (3.72)	-2 (3.47)	-5 (3.92)	+4 (4.13)	-1 (3.24)	+1 (3.68)	-5 (3.86)	-4 (3.81)
Percent at or above Proficient	+2 (2.71)	+1 (2.81)	-3 (2.99)	-1 (3.25)	+1 (3.47)	+6 (2.81)	-3 (3.05)	+5 (3.13)	-7 (3.66)	-4 (3.23)
Score at 10th percentile	+5 (4.91)	-2 (3.96)	+1 (5.65)	-2 (6.47)	+6 (6.61)	+1 (9.75)	-3 (9.91)	+1 (4.11)	-5 (5.07)	-6 (4.17)
Score at 25th percentile	+3 (5.25)	-1 (4.65)	-2 (3.48)	-2 (2.93)	-4 (3.88)	+4 (6.08)	0 (5.86)	+2 (4.17)	-5 (4.43)	-6 (4.48)
Score at 50th percentile	+2 (3.72)	+2 (2.88)	-3 (2.97)	0 (3.41)	-1 (3.73)	+6 (4.16)	-3 (4.19)	+5 (3.33)	-7 (3.50)	-3 (2.96)

The cells report the changes in Michigan scores net of the changes in Indiana, Ohio and the Nation. See NAEP, The Nation's Report Card (State Profiles) for details about the categories. Standard errors are in parentheses.

Table 14: Trends in Participation and Scores, ACT and SAT
(Michigan and Indiana, 1997-2001)

	Participation			Average Scores		
	Michigan	Indiana		Michigan	Indiana	
	ACT	ACT	SAT	ACT	ACT	SAT
Group 1 * 1998	0.23 (1.35)	-1.80 ⁺ (1.04)	-0.23 (1.52)	-0.14 (0.09)	0.20 (0.44)	-1.32 (2.70)
Group 2 * 1998	0.37 (1.21)	-1.51 (0.84)	-0.26 (1.44)	-0.05 (0.09)	0.35 (0.47)	-2.36 (3.25)
Group 1 * 1999	-0.59 (1.41)	1.19 (0.97)	2.12 (1.54)	-0.11 (0.10)	-0.01 (0.43)	-3.63 (3.07)
Group 2 * 1999	-0.44 (1.39)	0.67 (0.95)	1.40 (1.45)	0.16 (0.09)	0.59 (0.48)	0.95 (2.90)
Group 1 * 2000	0.45 (1.35)	3.51 ^{**} (1.17)	0.68 (1.61)	-0.03 (0.10)	0.02 (0.42)	-0.74 (2.99)
Group 2 * 2000	1.44 (1.30)	-0.43 (0.91)	2.08 (1.46)	0.09 (0.09)	0.27 (0.46)	-3.15 (2.95)
Group 1 * 2001	0.40 (1.38)	4.25 ^{**} (1.16)	1.31 (1.46)	-0.01 (0.10)	0.22 (0.44)	0.63 (3.06)
Group 2 * 2001	-0.62 (1.33)	-0.19 (1.17)	0.02 (1.74)	0.05 (0.09)	0.77 (0.49)	0.21 (3.27)

See equation (6) in the text. All the groups are based on base-year spending (1994 for Michigan, 1995 for Indiana) and Group 4 is the omitted category. All regressions are weighted by 12th grade enrollment and control for ethnicity and free lunch. Robust standard errors are in parentheses. ⁺, ^{*}, ^{**} denote significance at the 10, 5, and 1 percent levels respectively.

Table A-1(a): Selected Characteristics of K-12 Schooling in Michigan and five Mid-western States, 2001-2002

	Total Enrollment	Number of Districts	Students per District	Expenditure (1999-2000)	P-T Ratio	Fr./Red. Lunch	Asian	Hisp.	Black	White
Ohio	1,830,985	615	2977	\$7,065	15.0	27.4	1.2	1.8	16.5	78.9
Indiana	996,133	304	3277	\$7,192	16.7	31.1	1.0	3.9	11.8	83.0
Wisconsin	879,361	434	2026	\$7,806	14.4	26.0	3.4	5.0	10.2	80.1
Minnesota	851,384	439	1939	\$7,190	16.0	26.4	5.2	3.8	7.0	82.0
Illinois	2,071,391	1,011	2049	\$7,133	16.0	35.2	3.5	16.2	21.2	59.0
Michigan	1,730,668	573	3020	\$8,110	17.5	31.2	2.0	3.6	20.0	73.4

Source: Common Core of Data, National Center for Education Statistics. P-T ratio stands for the pupil-teacher ratio. The column 'Fr./Red. Lunch' shows the percentage of students eligible for free or reduced lunch. The last four columns show the percentages of students in each ethnic group.

Table A-1(b): Participation and Performance in SAT, ACT and NAEP (4th grade), Michigan and five Mid-western States

	SAT		ACT (2001)		NAEP Reading (1992)			NAEP Math (1996)			NAEP Math (2000)		
	Score	Tested	Score	Tested	Score	Basic	Prof.	Score	Basic	Prof.	Score	Basic	Prof.
Ohio	1073	26	21.4	63	217	63	27				231	73	26
Indiana	1000	60	21.4	20	221	68	30	229	72	24	234	78	31
Wisconsin	1180	6	22.2	68	224	71	33	231	74	27			
Minnesota	1169	9	22.1	66	221	68	31	232	76	29	235	78	34
Illinois	1165	12	21.6	71							225	66	21
Michigan	1133	11	21.3	69	216	62	26	226	68	23	231	72	29

Source: The Nation's Report Card, National Center for Education Statistics. The columns under 'Basic' show the percentage of test-takers scoring at or above the Basic level, similarly for 'Prof.' (Proficient). See the source for details about the categories.

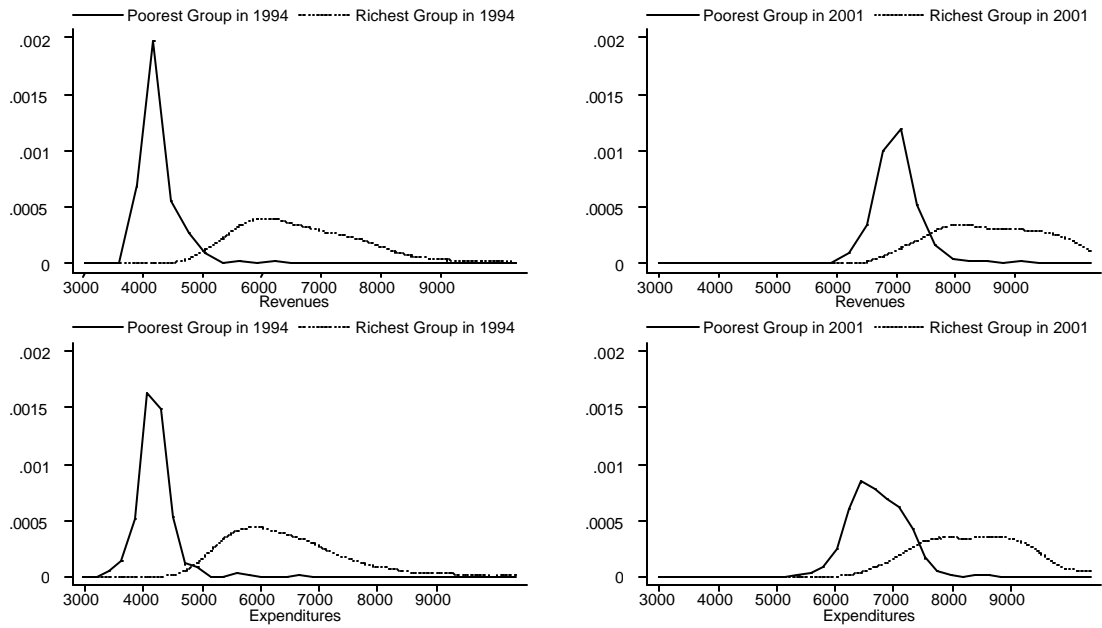


Figure 1. Spending in 1994 and 2001, Groups 1 and 5

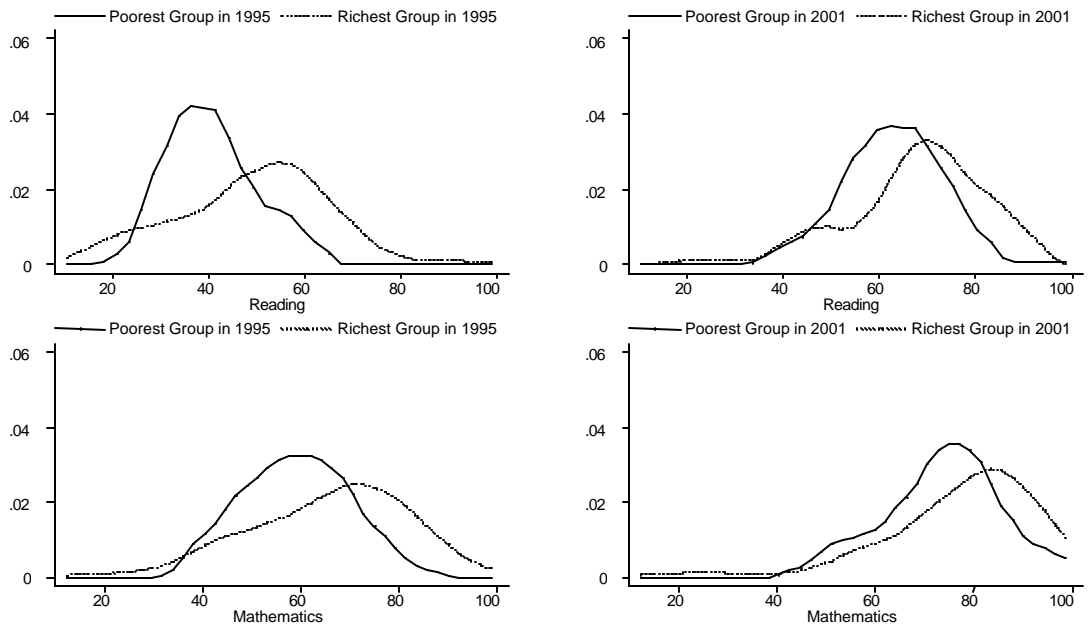


Figure 2. Test Scores in 1995 and 2001, Groups 1 and 5

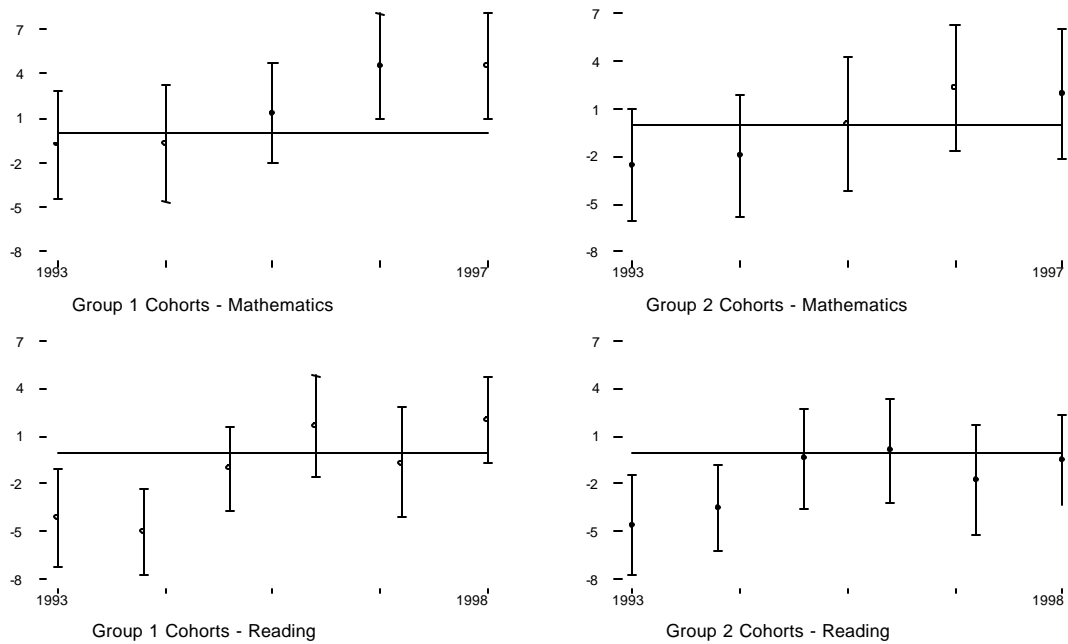


Figure 3: Gain between 4th and 7th grades for Groups 1 and 2 Cohorts

Notes: The figure shows the estimated coefficients on Groups 1 and 2 from regressions of equation (4), as reported in Table 7. The small circles represent the estimated values, while the lines show the respective 95% confidence interval bands. Cohorts are identified by the year they were in 4th grade, e.g. the 1993 cohort is the one that was in 4th grade in 1993 and in 7th grade in 1996.