NBER WORKING PAPER SERIES

SCHOOLING, INEQUALITY, AND THE IMPACT OF GOVERNMENT

Eric A. Hanushek
Julie A. Somers

Working Paper 7450
http://www.nber.org/papers/w7450

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 1999

Paper prepared for conference on *Increasing Income Inequality in America*, Texas A&M University, March 1999. The views expressed herein are those of the authors and not necessarily those of the National Bureau of Economic Research.

Schooling, Inequality, and the Impact of Government
Eric A. Hanushek and Julie A. Somers
NBER Working Paper No. 7450
December 1999

ABSTRACT

Analyses of income inequality have identified the importance of increased demand for worker skills, but characterizations of worker skills by the amount of schooling attained do not capture important aspects of the widening income distribution and of the stagnating relative wages of black workers. This paper is motivated by the possibility that schooling quality is an important component of the changing income distribution. The central analysis focuses on how governmental schooling policies – particularly those related to the level and distribution of school spending – affect the distribution of worker quality and of income. The substantial differences in spending across states are not significantly related to the variations in achievement growth across states. Further, the three decade old movement toward reducing the variation in school spending within states appears to have done nothing to reduce subsequent income variations of workers. Thus, the direct government policies toward school spending, as carried out in the past, have not ameliorated inequalities in incomes.

Eric A. Hanushek
Department of Economics
University of Rochester
Rochester, NY 14627-0158
and NBER
hanu@troi.cc.rochester.edu

Julie A. Somers
PHB Hagler Bailly, Inc.
One Memorial Drive
Cambridge, Massachusetts 02141-1344

Schooling, Inequality, and the Impact of Government

by Eric A. Hanushek and Julie A. Somers

Much is expected of our schooling system. Schools are expected to provide the preparation that students need for successful entry into the labor market. In the aggregate, schooling is seen as contributing to the growth of the economy and the overall rise in real incomes of workers. Moreover, the school system has become a generally accepted policy instrument to alter the distribution of income in society and to bring about more equality. But the traditional focus – both in research and policy – on the quantity of schooling neglects the rising importance of quality of schooling. Moreover, expansion to consider spending on schools, the most common operational surrogate for quality, does not remedy the distortions in focus. Thus, when specifically considering recent governmental actions that are most related to quality issues, there is little evidence that the direct interventions of government have had much effect on the distribution of outcomes.

Over the course of much of the 20th century, concentration on the quantity of schooling has been reasonable. The central role of government in providing access to schools has ensured a high level of schooling in the population. All states provide for universal and free public schools through grade twelve and most require school attendance at least through age 16.[1] Moreover, states are the primary provider of higher education, covering some 77 percent of all students in higher education with tuitions subsidized by general revenues so that they remain considerably below costs. As a result, the adult population with at least a high school education grew from less than 15 percent in 1910 to over 80 percent in 1990. Virtually all of the this growth, however, occurred before the mid1970s, leading to consideration of other aspects of governmental involvement in education.

[1]All states have compulsory schooling through age 16. Seventeen states mandate age 17 or 18 or high school graduation.

The quantity aspects of education are fairly well understood. They are, for example, featured in most standard decompositions of income inequality. In sorting through the effects of education, however, it has become clear that quality factors may be relevant in explaining the data. The rise of income inequality within schooling groups and the stagnation of black-white relative earnings each open the possibility of important effects of school quality.

This paper focuses on the distributional aspects of the quality of schools. Much of the existing discussion of school policy begins with a statement about the returns to quantity of schooling, notes that these returns are large and growing, and then moves to discussions about quality (Hanushek 1996). One obvious issue is whether quality has the same returns as quantity, but this immediately leads to questions about how to measure quality. In the context of government interventions, it is natural to concentrate on governmental spending as the measure of quality. Nonetheless, past work has suggested that governmental spending is not very closely related to quality (Hanushek 1986, 1997). Indeed, while many are generally satisfied when talking about quantity, there is fairly broad dissatisfaction with U.S. schools, perhaps resulting from their expense but more likely flowing from a general feeling that our schools are not living up to expectations. This paper revisits that question with particular emphasis on the distributional aspects of spending.

Dimensions of inequality

Recent analyses have documented the widening of income distributions. Part of this has flowed from the almost two decade long growth in education premia, i.e., the return to quantity of schooling. The relative earnings of college graduates has moved upwards quite steadily (Murphy and Welch, 1989; Pierce and Welch, 1996). This movement has pushed the distribution of incomes apart, with high school graduates seeing no real growth in earnings over an extended period of time while college graduates have been able to increase their earnings (Levy and Murnane, 1992).

But more than just this divergence has occurred. Within education groups, the distribution of incomes has also expanded (e.g., Levy and Murnane, 1992; Juhn, Murphy, and Pierce, 1991, 1992). Further, the pattern of this within group growth in the variance of income roughly follows the same time pattern as the returns to schooling levels (Juhn, Murphy, and Pierce, 1992; Murphy and Welch, 1992; Katz and Murphy, 1992).

Another dimension of these distributional changes has received societal attention. The distribution of income by race has been a common barometer of equity in society. While there is ambiguity about what is the "right" amount of inequality in the entire population – because of issues about the underlying ability distribution, about the incentive effects, and about other fundamental issues – there is less ambiguity about racial differences. The variations in inequality by racial group over time offer a perspective on changes in equity that is in some ways a cleaner view.

The incomes of blacks relative to whites have changed through time in ways not entirely dissimilar to the changes identified in incomes within and between schooling groups. Relative black-white earnings converged rather steadily from World War II through the 1970s (Smith and Welch, 1989), both because of convergence in schooling levels and because of other factors. But this convergence stopped in roughly 1980, and the earnings gap stood steady through the 1980s (Juhn, Murphy, and Pierce, 1992). Thus, with the growth in educational returns and with the other widening in the income distribution, Blacks found relative progress halted.

Some explanations

While there is controversy about the details (cf Welch 1999), the common interpretation of much of the movement in the various dimensions of inequality relates to an increased demand by the economy for skilled workers. This increased demand has driven the increasing returns to quantity of college education (and to additions of schooling at lower levels). As the return to skill rises, the

income gaps between people of differing schooling levels naturally widens. The movements in black-white earnings also appear closely related to these changes. Throughout the post-World War II period there has been convergence in the quantity of schooling of blacks and whites. For example, while young whites were three times as likely as young black males to complete high school in 1940, the gap was just 93 to 86 percent in 1996.[2] This convergence of schooling levels, with blacks catching up to whites, naturally leads to a convergence in earnings. The rapidly increasing returns to more schooling, however, works in the opposite direction, because whites still retain a clear advantage in overall levels of schooling. Thus, a portion of the overall trend in relative earnings represents the competing forces of converging schooling levels and diverging earnings by school level. But there has also been an increase in the portion of income disparity that is unexplained by schooling. Is this related?

As mentioned before, the income inequality within schooling categories has also been expanding. One possible explanation of at least a portion of the "unexplained" part of the overall income distribution and of the Black-white patterns involves school quality. Over some period of time, there has been a recognition of the heterogeneity of schools. It seems unassailable to assert that not all schools are alike. Simply put, few believe that every year of school – the basic building block of human capital and labor force quality – has the same skill content, regardless of which of the 85,000 elementary and secondary schools or 3,000 colleges produced it. While measurement of quality differences is controversial (and the subject of the subsequent discussion), an expansion in the returns to quality that mirrored the increased returns to quantity could partially explain each of the stylized facts about the changing distribution that were described earlier. If the distribution of quality, as

[2]In 1940, 41 percent of whites age 25-29 had completed at least high school education, while the similar figure for Blacks was 12 percent. The closing of the gap to the 1996 amount was quite steady over the postwar period.

relevant to the labor market, remained constant, a growing return on quality would lead to an expanded variance of earnings.[3] A movement toward improving the relative quality of schooling will push relative black-white earnings closer, while an increase in the returns to quality will lead to divergences when whites begin with higher quality (cf. O'Neill 1990).

We do not have precise measures of quality, but measured cognitive ability provides insights into at least one dimension of quality. The most consistent data on cognitive skills comes from the Scholastic Aptitude, or SAT, tests given to high school juniors and seniors. Figure 1 presents a time series of overall scores and of scores for Blacks and whites. Average test scores declined sharply from 1966 through the 1970s, recovered some in the early 1980s and remained stagnant through the mid1990s. The second element of the figure is the noticeable narrowing of the black-white differential during the early 1980s that essentially stopped in the mid1980s. This pattern of scores has a clear similarity to the changes in earnings inequality, particularly if the returns to skill have risen in recent periods. There is, however, a significant problem with these data. The SAT is voluntarily taken by a select group of students – those wishing to attend highly ranked colleges – and this pattern has been changing over time. Varying numbers of students have taken the exam. The percentage of high school seniors taking the SAT between 1972 and 1995 has varied between 31.0 and 42.4 percent, and, while it has generally risen over the period, it has not been a monotonic increase (U.S. Department of Education, 1966). By common analysis, this increased participation could cause score declines and could distort the Black-white comparisons. It appears that the decline in SAT scores through the 1970s was a combination of changing selection and real performance declines, but it is difficult to be certain about the magnitudes (Congressional Budget Office, 1986).

---

[3]This prediction is complicated, however, if school completion is growing, because the achievement or ability of people defined in terms of completed schooling will tend to be falling (see Welch 1999).

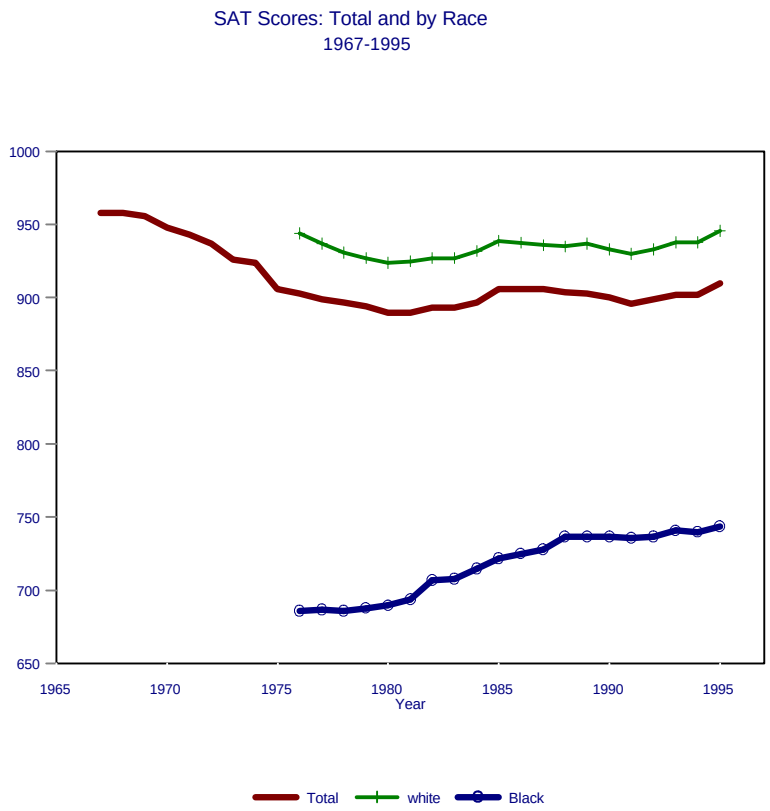


Figure 1. SAT Scores: Total and by Race, 1967-95

The National Assessment of Educational Progress (NAEP) provides a shorter picture of achievement (beginning in the early 1970s) but a picture that is uncontaminated by student selection. NAEP has periodically collected data for a random selection of students in differing subjects. Figure 2 displays the achievement patterns across subjects for 17-year-olds. It also divides performance by race. The patterns of performance mirror that displayed earlier on SAT performance. With the exception of reading performance, aggregate performance dips in the 1970s, recovers in the early 1980s, and then turns flat. For reading, aggregate performance is essentially flat for the entire period. There was also a consistent narrowing of performance between whites and Blacks or Hispanics in the early 1980s, only to stop in the late 1980s. The similarity of the NAEP and SAT trends after 1970 suggests that the precipitous fall in SAT performance during the 1960s may carry over to a representative sample, implying student achievement in the 1990s that is significantly below that in the 1960s.

This paper pursues issues of quality as a complementary element to quantity of schooling. The important matter for the discussion here is the extent of governmental influence on quality movements.

Government Interventions

The federal and state governments have long been involved in education, in part because of its perceived importance in affecting the distribution of income both across individuals and intergenerationally. The clearest statement of governmental motives and involvement is probably that of the War on Poverty, where President Lyndon Johnson highlighted the importance of human capital, broadly defined, in affecting the bottom of the income distribution. Affecting the income distribution and poverty levels through providing increased skills had and has broad appeal, because it offers long term improvement in the opportunities of the poor without excessive intrusion on natural labor market and work incentives.

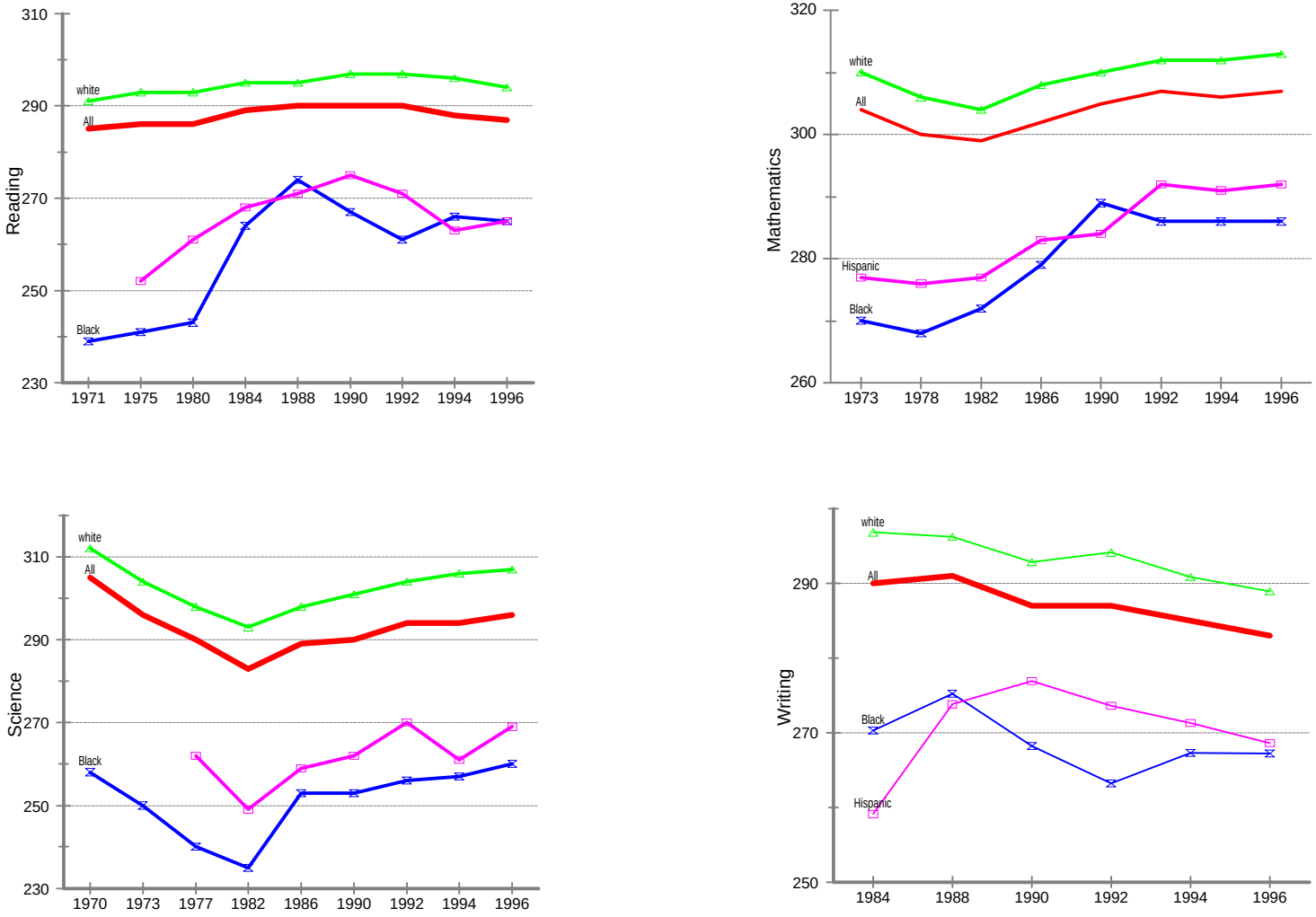


Figure 2. NAEP Scores for 17-year-olds by Subject and Race: 1970-1996

This paper concentrates on the role of skill building that takes place in elementary and secondary schools. A central focus is policy actions that have direct distributional objectives, although this is clearly too narrow. The natural operations of schools, coupled with choice of schools by families, have significant, albeit not necessarily intended, distributional implications. Thus, it will be necessary to decompose the effects of schooling into those affecting the level of skills and those affecting the distribution of skills.

The actual policy interventions have been quite varied. At the federal level, two main features have dominated. First, in the federal courts, the moves to desegregate schools and to promote equal opportunity by race has probably been the largest and most consistent intervention of the past half century. Beginning with *Brown v. Board of Education* in 1954, the federal courts have been important actors in the organization and operation of public schools (Welch and Light 1987; Armor 1995). They have clearly influenced the amount of racial mixing within and across school districts. It also appears that they have directly affected the resources available to Blacks and whites (see Boozer, Krueger, and Wolkon 1992). Little evidence is available to indicate any consistent effects of desegregation programs on student performance or other outcomes (Armor 1995). The evaluation is complicated, however, by the nature of the programs that affects both student distributions and a variety of school resources.[4]

Second, the federal government has funded compensatory programs for schools. These programs include Head Start pre-school programs, Title I compensatory education programs, and special education funding. Table 1 provides some idea of the funding levels that have been involved in the various federal programs. These spending numbers provide information on the real aggregate

[4]Desegregation programs have a variety of effects including changing the racial composition in individual schools, altering the distribution of teachers and resources across schools, and affecting the choice of districts by students. The complex patterns of impact make identification of the impacts on student outcomes difficult. The available studies concentrate on specific aspects but are unable to capture larger effects (cf. Hanushek 1972; Armor 1995).

Table 1. Funding of Major Federal Education Programs (millions of real 1998$)

Federal program	1965	1970	1975	1980	1985	1990	1995	1998
Programs for disadvantaged (Title 1)	$0	$5,625	$5,678	$6,340	$6,373	$5,605	$7,282	$7,811
Special education	$72	$332	$457	$1,626	$1,542	$2,017	$3,398	$3,812
Headstart	$0	$0	$1,224	$1,454	$1,628	$1,806	$3,780	$5,195
Other Department of Education	$2,862	$5,466	$6,387	$5,147	$3,139	$4,452	$4,325	$3,164

funds available.[5] Understanding the impact of these programs is complicated because it is difficult to tell who is receiving services. They are not entitlement programs and do not cover the entire eligible population. Moreover, the numbers of students served by the programs is difficult to measure, in part because of changes in programs. For example, the Title 1 compensatory programs recently changed from being targeted to individual eligible students to being available for entire schools that had concentrations of eligible students – leading to a large jump in the reported numbers of students being served by the program.[6]

Evaluations of the effectiveness of federal programs in improving student performance do not suggest much overall success. Title 1, which has changed form repeatedly over its history, has never indicated success in boosting general performance of disadvantaged students (see, for example, Farkas and Hall, forthcoming). Headstart has historically been dubbed as having limited effectiveness, with any gains in early performance eroding over time (see Barnett 1992). Special education programs have never received any overall evaluation.[7] In sum, there is little reason that federal actions as a whole have had much effect on student achievement.

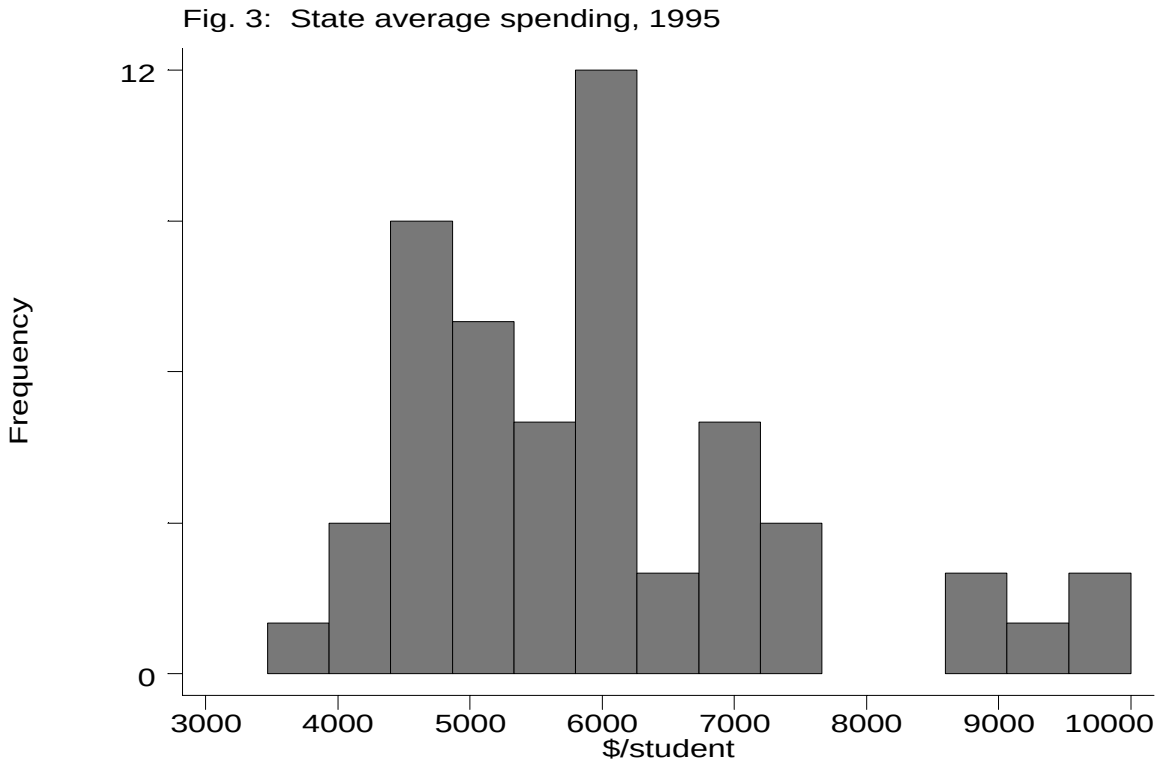
Schools are, nonetheless, the primary responsibility of the states, so the lack of systematic federal impact might not be altogether surprising. The states have pursued a variety of programs that affect equality in schools. Most significantly, states operate independently, implying substantial

[5]All budget numbers have been deflated by the CPI. These numbers present budget authority. Outlays follow a very similar pattern except that values from one year to the next may be distorted by which fiscal year spending is recorded.

[6]See Farkas and Hall (forthcoming) for a discussion of numbers served and of changes in targeting of the program.

[7]Hanushek, Kain, and Rivkin (1998) provided some support that the average program in Texas does boost achievement of special education students.

differences in spending, regulations, and operations across states. While the compensatory federal spending has some equalizing effect, it is small relative to the overall disparities in funding. Figure 3



shows the distribution of mean expenditure across states. The spending data, while unadjusted for any cost of education differences, show a remarkable spread.

The difference in mean spending across states is the largest component of inequality in resources available to students. Murray, Evans, and Schwab (1998) employ various approaches to decompose variations in spending across schools into that lying within states and that which comes from overall spending differences across states. While varying slightly depending upon the measures employed, the general finding is that two-thirds of the differences in school spending come from between state differences.

Figure 4 shows how the portion added by the federal government is related to state and local spending. It is interesting that, while federal spending has focused on purely distributional issues in terms of disadvantaged (low-come) populations, there is little equalization of overall spending across states. Federal spending has done little to disturb existing spending differentials across states, even though this large variation in spending is negatively related to the education and income of the state.

A variety of other programs and financing incentives is designed to promote more equality in schools. Part of this can be seen in Table 2, which charts the financing of schools by the various levels of government. States have increased their share of spending, while local shares have declined. After early growth in programs, the federal share has been virtually flat.[8] The growth in state shares is at least partially related to school funding court cases or attempts of legislatures to deal independently with the issues raised in those cases.

Beginning in the late 1960s, a wave of school finance cases have swept the nation. The origins of these cases are typically traced to the California case of *Seranno v. Priest*. The underlying legal theory was that children in property poor school districts with their commensurately limited taxing power were being discriminated against, because the ability of the school to raise funds depended on

[8]Note that these expenditure calculations do not include Head Start, whose funding does not go through schools (U.S. Department of Education, 1997).

Fig. 4: Federal and nonfederal spending, 1995

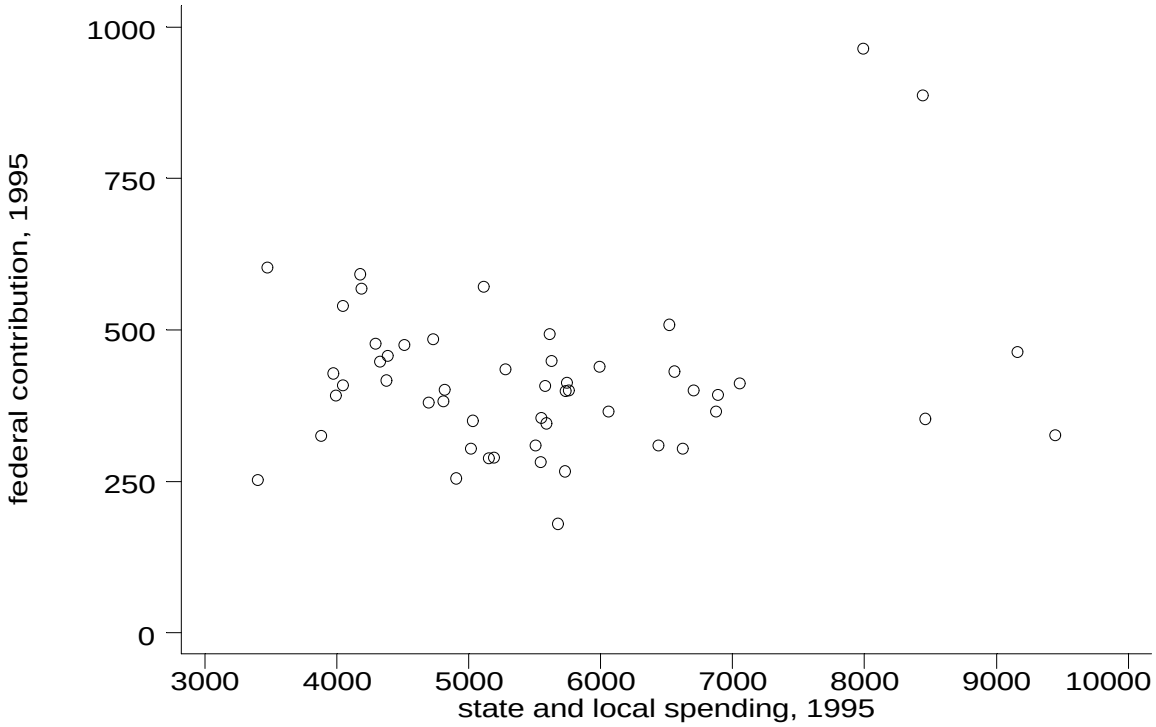


Table 2. Average Shares of Public School Revenues

	1960	1970	1980	1990	1995
Federal	4.4%	8.0%	9.8%	6.1%	6.8%
State	39.1%	39.9%	46.8%	47.1%	46.8%
Local	56.5%	52.1%	43.4%	46.8%	46.4%
Total	100.0%	100.0%	100.0%	100.0%	100.0%

the wealth of the students' neighbors.[9] This suit, originally brought under both state and U.S. constitutions, became the model for similar suits in a majority of the states. While the U.S. Supreme Court ruled that existing state school financing plans did not violate the equal protection clause of the 14th Amendment, most state constitutions explicitly define a state role in the provision of elementary and secondary schooling, and they have been the focus of suits.

State courts have split on whether or not their financing arrangements violate the state constitution, but one overall effect of the court action has been the relative increase in state funding that has come from the state. The general thrust of these suits has been that states should take a larger responsibility in school funding so as to ameliorate if not eliminate the funding advantages that certain districts have. This by itself leads to an increase in state share. Moreover, since there is frequently a significant amount of redistribution of funding called for by court orders and by legislative "equity" initiatives, it appears frequently to be more feasible to increase the total spending while changing the pattern; i.e., it is easier to redistribute a larger pie than a constant pie. Considerable heterogeneity exists across states, however, and such generalizations fit the aggregate better than individual states.

[9]The underlying academic arguments that were important in the development of these cases are found in Coons, Clune, and Sugerman (1970).

The primary focus of the court cases has been equity (although it may be changing recently).[10] If there is a wide disparity in the funding and quality of schools, the argument goes, there will be subsequent disparities in earnings and other outcomes. And, while quality is the general rubric of concern, most of the court cases have focused a majority of attention on purely fiscal and expenditure aspects of schools.

The prevailing evidenced suggests that court cases have tended to promote a more even distribution of spending across districts. Wykoff (1992) analyzed measures of dispersion of spending across states. This work was extended by Murray, Evans, and Schwab (1998). Evans, Murray, and Schwab show that states under court order have moved more toward equality that those not under order, although most states have not made dramatic changes.

The slow changes in the state share should not obscure the fact that overall spending grew dramatically over the past three decades. Table 3 shows the overall growth in spending and in the underlying components of spending. Overall real spending per pupil more than doubled between 1965 and 1995. During the same period pupil-teacher ratios declined dramatically, and the percentage of teachers with a master's degree or more more than doubled so that a majority of teachers in 1995 had graduate degrees. The growth in spending has also been larger in states with court judgments (Downes and Shah 1995). Thus, movements toward equality have also contributed to overall growth (with some potential effects on variations in spending across states).

The important point throughout this discussion is that there has been a consistent push to bring about more equality in the provision of schooling. The federal government has concentrated almost all of its effort on supporting disadvantaged students with supplementary funds and programs. The states

[10]The new version of state school finance cases has focused on "adequacy," or whether state funding is sufficient to meet state educational goals. While ambiguity exists in the exact definition, this set of school finance cases appears to address both the distribution and level of spending across districts.

Table 3. Public School Resources in the United States, 1961-1996

Resource	1960-61	1965-66	1970-71	1975-76	1980-81	1985-86	1990-91	1995-96
Pupil-teacher ratio	25.8	24.7	22.3	20.4	18.7	17.9	17.2	17.3
% teachers with master's degree or more	23.5	23.3	27.5	37.5	49.6	51.4	53.1	56.2
median years teacher experience	11	8	8	8	12	15	15	15
current expenditure/ADA (1996-97 $'s)	$2,122	$2,678	$3,645	$4,308	$4,589	$5,484	$6,239	$6,434

Source: U.S. Department of Education[1997]

provide broader support for schools, but much of it is designed to bring about greater in-state equality in school resources.

Does this movement by governments have a discernible effect on student performance and on subsequent inequality in earnings and well-being? That issue, of course, is central to any assessment of government's contribution to ameliorating societal inequalities.

New Evidence on Government Interventions and Outcomes

The importance of states and localities, plus differences in the availability of data, lead to separate consideration of between state effects and within state effects of government spending and policy. In both between and within analyses, we also consider effects on not only the overall variation in outcomes but also the differences in black and white outcomes. Unfortunately, much of the past evidence relates only tangentially to the kinds of policies that appear directly related to inequality. The past evidence that resources are not closely related to mean student performance suggests that actions aimed at the distribution of resources will also meet with limited success in terms of outcomes. Nonetheless, estimation of the effects of differences in levels of spending across states has been flawed by misspecification problems. And, the effects of changes in the within-state distribution of resources have not been adequately traced to the distribution of student outcomes, the presumed object of equalization policies.

Between state evidence

The between state differences in school quality and student outcomes can be thought of as an extension of existing analyses of educational production functions. In the production function analyses, resources usage (and other factors) are related to student performance. It is easy then to see that these provide some direct information about how resource differences across states would be

expected to affect the overall state distribution of state performance. If the educational production function were known, it would be possible to estimate directly how differences in mean spending translated into marginal differences in mean student performance.

The overall assessment of the findings of educational production function estimation has been presented elsewhere (Hanushek 1986, 1997). The summary in those articles, which collect and tabulate the existing production function estimates of resource effects, is that there does not appear to be a strong or systematic relationship between resources and student performance. This evidence combines the estimated effects of real resources and of spending differences, which both point in the same direction.[11] The estimated effects of variations in resources on student performance are distributed around zero with only a small portion indicating a positive and statistically significant effect of resources on performance. Therefore, the existing econometric studies give little reason to expect that more resources will be translated into outcomes, at least within the current organizational environment. This evidence thus presents a prima facie case that efforts to equalize funding and resources will have little effect on the equality of outcomes.

[11]Teacher salaries are closely related to teacher experience and to teacher degree levels, and variations in pupil-teacher ratios indicate how salaries are distributed across students. Thus, these measures are a good indication of variations in instructional spending per student. Because these real resource measures are more frequently available than spending measures and because they can be taken down to the classroom level, the real resources are overall better indicators of the effects of resources than spending.

Table 4. Percentage Distribution of Estimated Effect of Expenditure per Pupil by State Sampling Scheme and Aggregation

Level of aggregation of recources	number of estimates	Statistically significant		Statistically insignificant		
		Positive	Negative	Positive	Negative	Unknown sign
Total	163	27%	7%	34%	19%	13%
Single state samples[a]	89	20	11	30	26	12
Multiple state samples[b]	74	35	1	39	11	14
With within-state variation[c]	46	17	0	43	18	22
Without within-state variation[d]	28	64	4	32	0	0

Notes: Rows may not add to 100 because of rounding.

a. Estimates from samples drawn within single states.
b. Estimates from samples drawn across multiple states.
c. Resource measures at level of classroom, school, district, or county, allowing for variation within each state.
d. Resource measures aggregated to state level with no variation within each state.

This work is not without controversy, however, and part of the debate relates directly to the discussions here. Card and Krueger (1992) concentrate on how historic differences across states in resources and spending relate to differences in worker earnings.[12] They conclude that interstate differences in spending are directly related to subsequent labor market earnings. Thus, this analysis by Card and Krueger both contradicts the overall findings of the production function analysis and raises the suggestion here that the large interstate differences in school spending directly relate to current issues of income inequality.

The extensive discussions aimed at resolving the differences between the overall production function estimates and the Card-Krueger estimation point to important considerations for this analysis. While the differences have been approached from a variety of perspectives,[13] the key for this work relates to differences in state education policies. No studies have adequately characterized or measured how state policies (other than spending) influence student performance. This omission has clear implications for bias in the estimation of educational production functions across states. Hanushek, Rivkin, and Taylor (1997) note that aggregation of data to the state level will work to exacerbate any regression bias that comes from omitted state-level variables. The theoretical sign for any bias depends on the correlation of state policies and spending and is unknown *a priori*. They provide evidence that demonstrates an aggregation bias that pushes estimated expenditure effects toward showing positive impact on student performance. Specifically, a review of past analyses, reproduced

---

[12]Their approach, described more below, uses census data for workers at the given census year (1970 or 1980) to derive the labor market earnings of workers along with schooling levels and age. They estimate earnings generating functions where the estimated return to schooling varies by workers' state of birth and birth cohort. These returns to schooling are then related to characteristics of schooling systems in each state of birth and time period.

[13]A range of analyses is found in Burtless (1996). The most direct analysis of the Card and Krueger approach is found in Heckman, Layne-Farrar, and Todd (1996a, 1996b). They highlight issues of selective migration of workers across states, functional form of the estimated earnings models, and sensitivity of results to the particular samples.

in Table 4, shows that studies within single states (where any state policies are held constant) find less support for any spending effects than studies that go across states. Moreover, this bias toward positive spending effects is more significant when the across-state analyses rely upon state aggregate data.[14] The essential aspect of this evidence is the strong indication that differences in state policies are important in determining student performance, but these studies offer little guidance about how to specify or measure the key aspects of state differences.

Table 5. Effect of state school spending on growth in NAEP student performance between 4th and 8th grade

	Mathematics (1992-96)	Reading (1994-98)
log (% adults less than high school)	-0.039 (-2.56)	-0.075 (-2.78)
log (school expenditures 1992-96)	-0.0062 (-0.77)	-0.014 (-0.92)
constant	0.427 (4.56)	0.637 (3.70)
R-squared	.18	.22
number of states	36	32

Note: Regression estimates are weighted by the average daily attendance (ADA) in the state in 1995. Test performance is the change in log of average student performance between fourth and eighth grade in the years identified. Education of parents is the percent of the population age 25 and older with 11 or fewer years of education in 1990. School spending is the geometric average real state current expenditure per ADA between 1992 and 1996.

[14]The systematic effects of aggregation could come either from omitted variables bias or from errors in variables. In the later case, aggregation would improve the results, and the stronger effects in the aggregate data would be more accurate. Hanushek, Rivkin, and Taylor (1996) test these alternative explanations and find that misspecification bias and not errors in variables explains the pattern of results.

To consider the impact of between state differences in spending, we begin with consideration of the potential biases from different state policies in the context of a simplified production function:

$$(1) \qquad O_{st} = f(X_{st}, r_s R_{st})$$

where O_{st} is aggregate student outcomes in state s in time t; X_{st} is a vector of current and past family and other exogenous factors affecting performance; and, $\rho_s R_{st}$ is the effective school resources in state s cumulative to time t. The effective school resources are written as spending in the state (R) that is augmented by state policies (ρ_s) which imply that any spending is more or less effective depending upon prior policy choices.[15] Then consider directly estimating an operational version of this such as:

$$(2) \qquad O_{st} = A_t X_{st}^{b} \left(r_s R_{st} \right)^{g} e_{st}$$

where ϵ_{st} is an i.i.d. error term and β and γ are production parameters. Here the problem of state differences in the effectiveness of resources (ρ_s) becomes apparent. In estimation form, we have:

$$(3) \qquad o_{st} = a_t + b x_{st} + g r_{st} + (g \ln r_s + \ln e_{st})$$

where small letters denote logarithms of the previously defined variables. If ρ_s is not measured but included in the composite error term, the parameters of interest will be biased when this omitted variable is correlated either with resources or other determinants of performance. Now consider the situation if student performance and its determinants can be observed at some previous time τ for the same set of students. If state policies (i.e., ρ_s for each state) are constant over time, we can estimate:

$$(4) \qquad \underset{t \to t}{\Delta o_s} = a^* + b \Delta x_s + g \Delta r_s + (\ln e_{st} - \ln e_{st})$$

[15]The state factor incorporates both cost differences across states and policies that affect how resources are translated into student performance. We do not emphasize cost differences here, although they obviously complicate any direct analyses of interstate spending differences.

Estimation of this model with aggregate data, while subject to considerable uncertainty, can provide some clues to how government policy influences student performance. In particular, during the 1990s, the National Assessment of Educational Progress (NAEP) has provided state data on student performance at different grades. The key element of the NAEP testing is that the same cohort of students is tested over time so that estimation of Equation 4 is feasible. In mathematics, state performance for fourth graders in 1992 can be matched with this cohort's eighth grade scores in 1996. Similarly, in reading, state performance for fourth graders in 1994 can be matched with this cohort's eighth grade scores in 1998.[16]

Table 5 presents basic estimates for between-state differences in NAEP performance. These simple models provide no reason to believe that the very large state differences in resources lead to differences in average achievement across states. The estimates of the resource parameter (γ) are significant and negative. State differences in adult education levels do on the other hand significantly affect the growth in student performance.

It is possible, however, that these aggregate performance measures mask an underlying structure of differences. Most importantly, the aggregate performance presented earlier on NAEP over time shows a narrowing of the racial gap, particularly during the 1980s. Grissmer, Flanagan, and Williamson (1998) suggest that this narrowing was largely a function of added resources to the educational system, either reflecting total spending or reductions in pupil-teacher ratios. Cook and Evans (1996) on the other hand find that the narrowing of NAEP scores between blacks and whites cannot be a reflection of school funding levels. Thus, if resources were to be a prime factor behind the

[16]The NAEP is not a longitudinal design with respect to individual students but it includes random samples of the same cohorts in these different time periods. This set of matched cross-sections, when employed in the value-added form of Equation 4, introduces some measurement in background factors, but any bias is proportional to the sampling error to total variance ratio.

narrowing of test score differences by race, it must be the case that minorities are more sensitive to resource differences than white students.

To test for differential sensitivity of minorities to spending, we consider the race-specific value added of schools across states. The same state NAEP data are disaggregated into white, black, and Hispanic scores, and the same log growth models are estimated separately for each group. The results of this estimation are displayed in Table 6. Again, as can be easily seen, state differences in resources have no more effect on black or Hispanic scores than on white scores. These results also hold when (not shown) the racial distribution of students in each state is included in the models.

In sum, the between-state differences in resources appear to have little to do with the distribution of student outcomes. This reinforces the overall findings of the production function analyses, but it also adds important new information. The previous evidence of state differences in spending is confounded with an omitted variables problem. This evidence–which deals with the omitted variables problem through estimation of state value added models–adds further insight into the role of states in affecting inequality. This finding is also important because approximately two-thirds of the variation in school spending is found in variations in mean spending across states (Murray, Evans, and Schwab, 1998).

Table 6. Effect of state school spending on growth in NAEP student performance between 4th and 8th grade, by race and ethnicity

	Mathematics 1992-96			Reading 1994-98		
	White	Black	Hispanic	White	Black	Hispanic
log (% adults less than high school)	-0.036 (-2.24)	-0.095 (-2.20)	-0.197 (-5.73)	-0.082 (-3.10)	-0.134 (-3.00)	-0.327 (-5.15)
log (school expenditures 1992-96)	-0.0039 (-0.45)	-0.138 (-0.77)	-0.0097 (-0.65)	-0.146 (-0.97)	-0.014 (-0.89)	0.0049 (0.17)
constant	0.391 (3.95)	0.725 (2.64)	1.078 (7.23)	0.653 (3.94)	0.869 (3.59)	1.51 (5.40)
R-squared	.14	.16	.58	.26	.28	.54
number of states	36	30	33	32	27	28

Note: Test performance is the change in log of average student performance between fourth and eighth grade in the years identified for each race/ethnic group. Regression estimates are weighted by the average daily attendance (ADA) in each racial/ethnic group in the state in 1995. The education of parents is the percent of the population age 25 and older with 11 or fewer years of education in 1990. School spending is the geometric average real state current expenditure per ADA between 1992 and 1996.

Within state evidence

The previous evidence on how variations in spending across states are unrelated to student performance fails to address two significant issues. First, the measure of student performance is eighth grade test scores, but these may not be a good reflection of outcomes. Second, much of the policy attention given to distributional issues, both in judicial and legislative arenas, has focused on the within state variation in spending, an issue that is ignored in the between state analysis. This section addresses both of these issues.

Test performance of students has been used frequently as a surrogate for longer term outcomes such as labor market success, although there are questions about this. The primary argument for

employing test score measures is that student performance can be related to the specific teachers and programs of students in ways that are generally impossible if one must wait for students to enter the labor market. This approach presumes that test scores are a reasonable index for skill dimensions that ultimately show up in longer term outcomes. While there is evidence that this is the case, it is also clear that much of labor market outcomes is not explained by variations in measured test scores – suggesting the possibility that school resources could still influence students' long run performance.[17] For example, Card and Krueger (1996) suggest that one reason for the divergence of results between the general findings of production function studies of education and their analysis of census earnings patterns is simply the superior measures of long run outcomes that are found in labor market data. By concentrating on labor market outcomes, they assert, a truer picture of the relationship between school resources and student performance can be observed.

The Card-Krueger approach is useful to consider here. They begin with data from the decennial censuses which gives earnings of individuals of varying ages and schooling levels. They then use information about state of birth and age to estimate the average state school spending when each individual was attending schooling. Thus, for example, a 35 year old high school graduate in 1980 who was born in Ohio would be estimated to have the average Ohio school spending between 1951 and 1963. They then relate variations in the estimated return to varying quantity of schooling to school spending for each individual.[18] While their approach was directly related to between-state

---

[17]Early analyses of labor market outcomes tended to show positive but relatively weak relationships between earnings and test scores. More recent analyses, however, have indicated an increased relationship (e.g., O'Neill, 1990; Bishop, 1991; Grogger and Eide, 1993; Murnane, Willett, and Levy, 1995; Neal and Johnson, 1996). The exact reason for the increased relationship is not entirely clear, because more recent studies have improved data sets and analytical methods and because the nature of the labor market itself may have changed. Nonetheless, at least for more recent periods, the two stage model suggested in common production function estimation seems justified.

[18]This description is highly simplified, and the details prove to be important. See Heckman, Layne-Farrar, and Todd (1996b) and other discussions in Burtless (1996).

differences in school resources and earnings, it seems better suited to consideration of within-state differences.

Consider earnings for an individual i that are consistent with equation 2:

$$(5) \qquad o_{sti} = a_t + b x_{sti} + g r_{sti} + (g \ln r_s + \ln e_{sti})$$

Within a state, ρ_s is constant, so the variance in (log) earnings would be:

$$(6) \qquad \text{var}(o_{st}) = b^2 \text{var}(x_{st}) + g^2 \text{var}(r_{st}) + \text{var}(\ln(e_{st})) + 2bg \ \text{cov}(x_{st}, r_{st})$$

under the assumption that the errors (ϵ) are uncorrelated with the exogenous variables. In the subsequent empirical analysis, we do not explicitly measure cov(x,r) but include it as part of the composite error term. This approach presents problems only if there is a systematic relationship across states between the within state variance in resources and family background and the within state covariance of these. We know of no evidence on this.

The empirical analysis matches variations in earnings to variations in family education and state school resources. We concentrate on workers who are age 25-37 in 1990. The bottom end of the range is chosen to begin with a period when most individuals have left school and are in the labor force. The top end of the range reflects availability of school spending data across districts within each state. The *Census of Government School System Finance (F33) File* provides annual data on current expenditures per pupil from 1968-96.[19] Someone who is 37 in 1990 would have been in the ninth grade in 1968. Thus, by beginning with this group, we can calculate the relevant average high school spending (grade 9-12) for each state for the workers in the sample.

[19]In some years there is a full census of districts; in other years there is a sample of districts within each state. We combine these to obtain annual estimates of spending variation.

We divide the age range into four separate cohorts (age 25-27, 28-30, 31-33, and 34-37). For each region of worker residence, we can calculate variations in earnings by cohort, race, gender, years of schooling, and birth state. We then merge information on school spending variations (from the F33 data) and on variations in family characteristics (from the Current Population Survey) within birth state for the relevant cohorts.[20]

We estimate different variants of equation 6 using the combined data set with separate observations for the various groups.[21] Table 7 presents basic estimates for the combined cohorts. These estimates are separately provided by race and gender. While other work (Somers 1998) presents a variety of variants on the basic estimation, here we concentrate simply on the overall patterns.[22]

The key result for the purposes of this investigation is that variations in spending have a significant *negative* relationship to variations in earnings for whites. For black males there is an insignificant positive relationship, while only for black females does a significant positive relationship appear. In other words, except for black females, reduced variation in spending – the object of judicial and legislative policies in many states – is actually associated with higher variation in ultimate labor market outcomes.

---

[20]The quality of family background data from the CPS differs across time, because the specific age of children in families was unavailable before 1967. Here we combine evidence across the different CPS years, but separation does not change the results.

[21]The actual estimation relies upon coefficient of variations for earnings, family background, and spending. Because we use linear rather than logarithmic variances, this approach eliminate differences in cost of living across areas and focuses on the underlying distributions.

[22]Variations include different measures of family background, different sample definitions by cohort and employment status, and different specific empirical specifications that incorporate degree level indicators or not.

Table 7. Estimated Relationship of Variations in School Spending and Variations in Earnings (All workers age 25-37 in 1990)

	White men	White women	Black men	Black women
School spending	-0.051	-0.076	0.014	0.081
(c.v.)	(-6.1)	(-7.2)	(0.6)	(3.5)
Parental education	.077	0.001	0.002	(0.151
(c.v.)	(8.4)	(0.1)	(0.2)	(10.7)
R^2	.52	.64	.43	.51
observations	3950	3903	2166	2190

When other measures of family background are employed, similar results hold. Additionally, because of possible endogeneity concerns, the occurrence of a state court judgment requiring more equalized funding is used as an instrument for spending in years after the judgment. This also fails to change the overall finding of unintended perverse effects of expenditure equalization.

The important aspect of this estimation is that it provides a general way of assessing the impact of variations in within state spending on performance. Even though the modern wave of school finance reform, largely propelled by state actions, has lasted over three decades, little analysis has been done of the impact of these actions.[23] This analysis generalizes the results to incorporate both interstate and intertemporal variations in state spending inequality, and finds that the intended consequences of finance reform have not been realized.

[23]Downes (1992) is an exception. Looking at reform in California following the *Seranno* case, he finds little reduction in the variation of student performance.

Conclusions

Much of the explicit rational for government educational policies has to do with equalizing opportunities for citizens. This rational is clearest at the federal government level but also holds for state governments.

The focus of policy actions toward schools is most often the level of spending in schools. A long history of past analyses of spending variations suggests that spending is a clumsy instrument for changing student performance. So there is question about whether redistributive spending policies translate into more equality of outcomes.

The explicit analysis here concentrates on the two basic components of governmental spending policies. First, wide variation in spending per student exists across states. Past investigation of interstate differences in spending have been confused by misspecification problems and the implications of unmeasured differences in general state schooling policies. Here variations in state spending are related to value-added in student achievement across states. This approach circumvents the most serious biases from state policy differences. The results of between state differences in spending indicate no relationship with student performance on math and reading tests.

Second, variations in spending within states are related to variations in labor market outcomes. This added analysis both relates directly to current judicial and legislative interventions in schools and provides a conceptually superior measure of student outcomes. This analysis provides little confirmation of a powerful equalizing role for states. For both white males and white females, there is a negative relationship between spending variations and variations in ultimate labor market outcomes. Only for black females is there confirmation of an hypothesized positive relationship between variations in school resources and subsequent equality of outcomes.

The results here confirm previous analyses of the determinants of student performance. Resource variations, the most common metric for considering governmental school policies, are not related to the level or variation in student performance.

References

Armor, David J. *Forced justice: School desegregation and the law*. New York: Oxford University Press, 1995.

Bishop, John. "Achievement, test scores, and relative wages." In *Workers and their wages*, edited by Marvin H. Kosters, 146-86. Washington, DC: The AEI Press, 1991.

Boozer, Michael A., Alan B. Krueger, and Shari Wolkon. "Race and school quality since Brown v. Board of Education." *Brooking Papers: Microeconomics* (1992): 269-338.

Burtless, Gary, ed. *Does money matter? The effect of school resources on student achievement and adult success*. Washington, DC: Brookings, 1996.

Card, David, and Alan B. Krueger. "Does school quality matter? Returns to education and the characteristics of public schools in the United States." *Journal of Political Economy* 100, no. 1 (February 1992): 1-40.

Congressional Budget Office. *Trends in educational achievement*. Washington, D.C.: Congressional Budget Office, 1986.

Cook, Michael D., and William N. Evans. "Families or schools?" Explaining the convergence in white and black academic performance". mimeo, 1996.

Coons, John E., William H. Clune, and Stephen D. Sugarman. *Private wealth and public education*. Cambridge, MA: The Belknap Press of Harvard University Press, 1970.

Downes, Thomas A. "Evaluating the impact of school finance reform on the provision of public education: The California case." *National Tax Journal* 45, no. 4 (December 1992): 405-19.

Downes, Thomas A., and Mona P. Shah. The effect of school finance reforms on the level and growth of per pupil expenditures. paper presented at American Economic Association meetings, Washington, DC, January 1995.

Farkas, George, and L. Shane Hall. "Can Tile I attain its goal?" *Brookings Papers on Education Policy* (forthcoming).

Grissmer, David, Ann Flanagan, and Stephanie Williamson. "Why did the Black-white score gap narrow in the 1970s and 1980s?" In *The Black-white test score gap*, edited by Christopher Jencks and Meredith Phillips, 181-226. Washington, DC: Brookings Institution Presss, 1998.

Grogger, Jeffrey T., and Eric Eide. "Changes in college skills and the rise in the college wage premium." *Journal of Human Resources* 30, no. 2 (Spring 1993): 280-310.

Hanushek, Eric A. *Education and race: an analysis of the educational production process*. Cambridge, MA: Health-Lexington, 1972.

---. "The economics of schooling: Production and efficiency in public schools." *Journal of Economic Literature* 24, no. 3 (September 1986): 1141-77.

---. "Rationalizing School Spending: Efficiency, Externalities, and Equity, and their Connection to Rising Costs." In *Individual and Social Responsibility: Child care, education, medical care, and long-term care in America*, edited by Victor R. Fuchs, 59-91. Chicago: University of Chicago Press/NBER, 1996.

---. "Assessing the effects of school resources on student performance: An update." *Educational Evaluation and Policy Analysis* 19, no. 2 (Summer 1997): 141-64.

Hanushek, Eric A., Steven G. Rivkin, and Lori L. Taylor. "Aggregation and the estimated effects of school resources." *Review of Economics and Statistics* 78, no. 4 (November 1996): 611-27.

Hauser, Robert M., and Min-Hsiung Huang. "Trends in black-white test-score differentials". DP # 1110-96, Institute for Research on Poverty, University of Wisconsin, October 1996.

Heckman, James S., Anne Layne-Farrar, and Petra Todd. "Does measured school quality really matter? An examination of the earnings-quality relationship." In *Does money matter? The effect of school resources on student achievement and adult success*, edited by Gary Burtless, 192-289. Washington, DC: Brookings, 1996.

Heckman, James, Anne Layne-Farrar, and Petra Todd. "Human capital pricing equations with an application to estimating the effect of schooling quality on earnings." *Review of Economics and Statistics* 78, no. 4 (November 1996): 562-610.

Hedges, Larry V., and Amy Nowell. "Black-white test score convergence since 1965." In *The Black-white test score gap*, edited by Christopher Jencks and Meredith Phillips, 149-81. Washington, DC: Brookings Institution Presss, 1998.

Juhn, Chinhui, Kevin M. Murphy, and Brooks Pierce. "Accounting for the slowdown in black-white wage convergence." In *Workers and their wages*, edited by Marvin H. Kosters, 107-43. Washington, DC: The AEI Press, 1991.

Katz, Lawrence F., and Kevin M. Murphy. "Changes in relative wages, 1963-1987: Supply and demand factors." *Quarterly Journal of Economics* 107, no. 1 (February 1992): 35-78.

Levy, Frank, and Richard J. Murnane. "U.S. earnings levels and earnings inequality: A review of recent trends and proposed explanations." *Journal of Economic Literature* 30, no. 3 (September 1992): 1333-81.

Murnane, Richard J., John B. Willett, and Frank Levy. "The growing importance of cognitive skills in wage determination." *Review of Economics and Statistics* 77, no. 2 (May 1995): 251-66.

Murphy, Kevin M., and Finis Welch. "Wage premiums for college graduates: Recent growth and possible explanations." *Educational Researcher* 18, no. 4 (May 1989): 17-26.

---. "The structure of wages." *Quarterly Journal of Economics* 107, no. 1 (February 1992): 285-326.

Murray, Sheila E., William N. Evans, and Robert M. Schwab. "Education-finance reform and the distribution of education resources." *American Economic Review* 88, no. 4 (September 1998): 789-812.

Neal, Derek A., and William R. Johnson. "The role of pre-market factors in black-white differences." *Journal of Political Economy* 104, no. 5 (October 1996): 869-95.

O'Neill, June. "The role of human capital in earnings differences between black and white men." *Journal of Economic Perspectives* 4, no. 4 (Fall 1990): 25-46.

Pierce, Brooks, and Finis Welch. "Changes in the structure of wages." In *Improving America's schools: The role of incentives*, edited by Eric A. Hanushek and Dale W. Jorgenson, 53-73. Washington, DC: National Academy Press, 1996.

Smith, James P., and Finis Welch. "Black economic progress after Myrdal." *Journal of Economic Literature* 27, no. 2 (June 1989): 519-64.

Somers, Julie A. "The connection between educational spending and educational outcomes." Ph.D. Dissertation, University of Rochester, 1998.

U.S. Department of Education. *The Condition Of Education 1996*. Washington, D.C.: Education, National Center for Education Statistics, 1996.

Welch, Finis. "In defense of inequality." *American Economic Review* 89, no. 2 (May 1999): 1-17.

Welch, Finis, and Audrey Light. *New evidence on school desegregation*. Washington, D.C.: U. S. Commission on Civil Rights, 1987.

Wyckoff, James H. "The intrastate equality of public primary and secondary education resources in the U.S., 1980-1987." *Economics of Education Review* 11, no. 1 (March 1992): 19-30.