NBER WORKING PAPER SERIES

PUBLICLY PROVIDED EDUCATION

Eric A. Hanushek

Working Paper 8799
http://www.nber.org/papers/w8799

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 2002

This paper was prepared for the Handbook of Public Economics (Alan Auerbach and Martin Feldstein,Editors).Helpful comments and suggestions were provided by Alan Auerbach,John Kain,Susanna Loeb,Derek Neal,Tom Nechyba,and Steve Rivkin. The views expressed herein are those of the authors and not necessarily those of the National Bureau of Economic Research.

Publicly Provided Education
Eric A. Hanushek
NBER Working Paper No. 8799
February 2002
JEL No. H4, I2

ABSTRACT

Historically, most attention in public programs has been given to the resources devoted to the activity, and resources have been used to index both commitment and quality. Education differs from other areas of public expenditure because direct measures of outcomes are available, making it is possible to consider results and, by implication, to consider the efficiency of provision. Early interpretations of the evidence, emanating from popular interpretations of the Coleman Report that "schools do not make a difference," are incorrect, but the basic evidence behind the statement suggests serious performance problems of government supply, because purchased inputs to schools are not closely related to outcomes. This paper reviews that evidence along with providing an evaluation of the various controversial aspects including issues of causality, consumer behavior, and estimation approaches. Two detailed policy areas are discussed in terms of the evidence on performance: public versus private provision and the financing of schools.

Eric A. Hanushek
Hoover Institution
Stanford University
Stanford, CA 94305-6010
and NBER
hanushek@stanford.edu

Contents

ABSTRACT i
Contents ii
Scope and Roadmap for Discussion 2
Overview of Schooling Issues 5
  Quantity of Schooling 5
  Quality Considerations 9
  Aggregate Resources and School Outcomes 14
Role of Government 16
Efficiency of Production 21
  General structure 23
  Effects of teacher and schools 25
  Components of school inputs 27
  Results of Production Function Estimation for the United States 29
    Basic Results 31
    Components of results 32
    The role of peers 33
    Racial integration 36
    Family inputs 38
  The importance of study quality 40
  International Evidence 43
Interpretation of Evidence on School Performance 46
  Controversies 48
    Causality 48
    Labor Market Outcomes 52
    Meta-Analysis and the Summary of Results 54
    Experimental Evidence (STAR Experiment) 56
    Consumer behavior 58
    Inefficiency and estimation approach 60
    Multiple outputs 62
  Cost Considerations 64
  Institutional Structure and Incentives 66
Public versus Private Provision 70
  Unsubsidized choice 71
  Subsidized choice 75
  International evidence 79
Financing of Schools 79
  Alternative funding schemes 82
  Fiscal effects of alternative plans 86
  Effect of equalization on equity 88
  Mobility, Distribution, and Equilibrium 91
Some Underexplored Topics 94
References 108

The provision of education is a major public sector activity around the world, and both developed and developing nations frequently act as if ensuring an appropriate education ranks close in priority to providing for the safety and security of their citizens. And, much like these other fundamental areas, governments not only provide a majority of the funding for schools but also typically operate the schools. The objectives of this essay include not only a consideration of the purposes and rationale for governmental involvement in schooling but also an evaluation of performance in addressing the underlying goals.

Even though government's presence in education is commonly accepted without much question, the degree and form of involvement warrants attention. The amount and quality of education is known to contribute to the income and well-being of individuals. Recent analyses also suggest that education has a powerful effect on the strength and growth of national economies. These factors do not, however, necessarily justify the extent or manner of governmental involvement.

Regardless of one's opinion about the involvement of government, it is valuable to assess how well government does at providing education. Historically, most attention in public programs has been given to the resources devoted to the activity, and resources have been used to index both commitment and quality. Such a perspective, while forced on many areas of governmental activity by lack of good measures of outcomes, is obviously quite limited. Importantly, education increasingly differs from other areas of public expenditure because direct measures of outcomes are becoming more readily available. Thus, instead of measuring governmental human capital investments just by expenditure, it is possible to consider results and, by implication, to consider the efficiency of provision. Such consideration reveals a complicated picture of expenditure patterns that are not matched by performance, although the conclusions and policy implications to be derived from this general finding are not straightforward. Because of the ability to consider performance directly, the study of publicly provided educational services also provides a possible window on one of the critical policy questions that continuously arises in mixed economies: how well does government do in directing society's resources to meet its public goals? Nonetheless, given our current

state of knowledge, considerable uncertainty necessarily remains about whether or not any inefficiencies of the educational system are typical of governmentally provided services.

The distribution of educational outcomes is also an element of any evaluation of educational performance. Education is not a pure public good, available without diminution to all. Instead, it is a complicated intermediate good that is partially produced by government through provision that varies across local jurisdictions and that that interacts with the endowments and actions of students and families. Substantial portions of the rewards of more schooling accrue directly to the individual. Thus, education has more of the character of a publicly provided private good. As such, the distribution of outcomes is not only an important concern in judging the performance of government provision but also an object of policy.

Education clearly encompasses a wider variety of things than can be readily covered here. The central issues analyzed are the organization, financing, and production of educational services. This discussion is placed within the context of the role of government in intervening in the market for education. The following section more fully delineates the scope of this investigation and provides a roadmap through the subsequent discussions.

Scope and Roadmap for Discussion

This essay comes in the midst of a rapidly expanding analytical base on the economics of education. The importance attached to human capital in many areas of economic research plus the intense policy concerns about schooling have heightened interest in scholarly exploration of the education sector per se. A primary purpose of this overall discussion is to highlight the most promising lines of research and to project future areas of productive research. In doing this, some limits on the scope of this endeavor are necessary.

Much of the discussion centers on experiences in the United States with the provision of primary and secondary education. This focus permits a clear development of the issues of service delivery that can be based on a very extensive analytical base. It is, nonetheless, a somewhat arbitrary delineation of the overall set of potential topics. Although the provision of formal schooling by government includes extensive interventions in tertiary education, higher education involves quite different institutional

structures. More importantly, little progress has been made in the measurement of outcomes of higher education, thus precluding direct analysis of the financing and production issues that are central to this analysis.[1]

The restriction to formal schooling situations excludes important aspects of human capital investment. Most countries have governmental involvement in various aspects of job training, including both formal vocational training and work with private employers. This involvement ranges from governmentally provided training programs to interventions in apprentice programs arranged directly with firms. The institutional structure of such programs differs widely across countries, but it is frequently more closely linked to the labor ministries of government than to the education ministries. This separation of function, while perhaps unfortunate from an overall policy perspective, implies that the issues and analyses of training programs have developed very differently from those related to formal schooling.[2] This analysis also generally follows the American custom of separating preschool programs from formal schooling. Again, however, a full treatment of human capital investment policies should span this period, because there is some indication that investments early have high pay-offs, and some discussion is provided below.[3]

The restriction generally to U.S. experience is an unfortunate limitation forced by the availability of comparable studies from other countries and other institutional experiences. The wide international variation in governmental institutions should be and is an advantage for some analyses. While the

---

[1] As developed below, the largest issues in primary and secondary education are whether or not various structures and financing schemes lead to qualitative differences in the performance of schools. For primary and secondary education, a variety of readily available measures of student outcomes, including test scores and differential labor market performance, permits direct analyses. But, no similar measures of higher education outcomes – beyond pure quantity differences – are available and accepted. A few attempts to look at labor market outcomes of higher educational quality are available (e.g., Solmon (1973); Sewell and Hauser (1975); James and Alsalam (1993); Behrman, Rosenzweig, and Taubman (1996); Behrman et al. (1998); Eide, Brewer, and Ehrenberg (1998); Dale and Krueger (1999)), but the field remains largely undeveloped.

[2] A review of materials on job training can be found in Heckman, LaLonde, and Smith (1999). For a general consideration of optimal human capital investment across the life cycle, see Heckman, Lochner, and Taber (2000).

[3] See, for example, Gramlich (1986), Barnett (1992), Heckman (2000), Currie and Thomas (1995, 1999), U.S. General Accounting Office (1997), Currie (2001).

limitation is rapidly disappearing with significant data development and analytic efforts around the world, the more limited range of analyses that are currently available makes in-depth consideration of the international similarities and differences impossible to develop in depth. To the extent possible, comparisons with both developed and developing countries are made throughout this discussion, but they generally cannot be summarized and organized as clearly as the United States experiences.

The discussion begins with an overview of stylized facts about schooling and with a discussion of how human capital enters into overall economic output and performance. A consideration of the conceptual basis for governmental involvement in education follows this.

The performance of government in providing education is a central element to the entire interpretation of both research and policy in this area. Beginning with a general conceptual model of the educational production process, detailed attention is given to what research has said about the effectiveness of inputs to education. This summary of analytical studies leads naturally to discussion of potential analytical issues that arise in the area and to how one might interpret the range of results.

The study of school performance is related to a variety of organizational issues that are key elements of governmental intervention and participation in education. Two of the most significant organizational aspects are the relationship between public schools and private schools and the form and means of financing education. The substantial literatures on each provide insights into these significant policy choices.

Finally, even though there is rapid expansion of research in this area, a number of areas remain underexplored. The last section identifies a series of issues that appear to be productive areas for the continual development of the area.

Overview of Schooling Issues

Economists generally view schooling as an investment both by students and by the society at large. Each incurs costs, and each reap rewards. For a student, the costs of education include the direct costs of tuition, books, and other school-related expenditures as well as the income that the student forgoes when attending school instead of taking a paying job. Similarly society incurs direct costs in subsidizing a school system that provides free or heavily subsidized education to its citizens. It also forgoes the opportunity to devote the skills, people and resources that are engaged in education to other projects. This viewpoint – regarding education as an investment – dates back to the 17th century with the writings of Sir William Petty and includes work by Adam Smith and other influential economists (see Kiker (1968)). It was brought into mainstream economics, however, by Schultz (1961, 1963), Becker (1993), and Mincer (1970, 1974) and has become the basis of a steady stream of subsequent theoretical and empirical analyses.

Quantity of Schooling

A look at the history of the twentieth century suggests that schooling has generally been a good investment, buoyed by steady increases in the demand for skilled workers. Individuals have dramatically increased their own investments in education.

In the United States, at the beginning of the twentieth century, only six percent of the adult population had finished high school. After the first world war, high school graduation rates began to increase rapidly. But changes in education work their way only slowly through the overall population. By 1940, only half of Americans aged 25 or older had completed more than eight years of school, that is, had had any high-school education at all. Not until 1967 did attainment of the median adult aged 25 or over exceed high school.[4] Since 1967, however, the increase in the number of years of schooling completed by Americans has slowed. The young adult population, aged 25 to 29, has had stable completion rates for

[4]See U.S. Bureau of the Census (1975, 2000) and Goldin (1998).

almost two decades. At the turn of the 21st century, over 80 percent of Americans over 25 had completed high school or more.

The changes in other nations have been even more dramatic. Table 1 shows the percentages of different age groups completing upper secondary schools for a sample of more developed countries.[5] The different age groups effectively trace the normal schooling in different decades in the past, so that the changes with age show the rate of increase in schooling. While the United States has been stable since the 1960s, most of the other countries have undergone massive increases in high school completion – mirroring the historical developments in the U.S. before and immediately after World War II (Goldin (1998)).

The benefits of education to individuals are also clear. The average incomes of workers with a high school education remain significantly above those of the less educated, and the average income of workers with a college education now dwarf those of the high-school educated. In the U.S., the rapidly increasing earnings of college-educated workers during the past two decades currently provides them with a premium of more than 70 percent higher earnings than a high school graduate with similar job experience.[6]

The earnings patterns elsewhere in the world appear quantitatively more varied, but there is a strong similarity in the earnings effects associated with more schooling. Table 2 shows earnings distributions by level of schooling and by gender across the adult labor force, again for a sample of developed countries. While the earning distribution is more compressed in some countries than others – probably reflecting characteristics of labor markets – invariably there are obvious gains in earnings to more schooling. Not only are wages higher for the better educated, but they also tend to enjoy greater job opportunities and suffer less unemployment (U.S. Department of Education (1996a); Organisation for Economic Co-operation and Development (2001)).

---

[5] A comprehensive comparison of schooling across nations can be found in Barro and Lee (2001).

[6] More detail on the patterns of earnings can be found in Murphy and Welch (1989, 1992), Kosters (1991), Pierce and Welch (1996), and Deere (2001). McMahon (1991) reports slightly lower private rates of return for high school completion than for college completion, although they remain substantial. These calculations all rely on just salary differentials, and greater equality in the provision of fringe benefits may act to compress the differences for total compensation. However, no analysis of schooling returns in terms of total compensation is available.

For individuals the increased relative incomes of more educated people have been sufficient to offset the costs. An individual can expect significant financial benefit from extended schooling, even after appropriately considering costs.[7] Individuals also gain non-financial benefits from education. For example, there is evidence that more educated people make better choices concerning health, so they tend to live longer and to have healthier lives. There is also evidence that the children of more educated parents get more out of school. They attend longer and learn more. Such benefits of schooling simply reinforce those from the labor market.[8]

The common interpretation of the overall returns is that high technology economies produce large demands for skilled workers, workers who can adapt to new technologies and manage complicated production processes effectively. Formal models with this character are developed in Nelson and Phelps (1966) and Welch (1970) and summarized in the ideas of dealing with disequilibrium in Schultz (1975).

Society as a whole also benefits from education. National income rises directly with earnings from workers with more and better skills. The more educated are more prone to be civically involved, to vote in

---

[7] While most economists think of schooling as involving the production of human capital in individuals, the screening or signaling perspective is a clear alternative (e.g., Spence (1973), Wolpin (1977), Weiss (1995)). The screening model in the extreme suggests that individuals begin schooling with differing abilities and that schooling merely allows employers to identify those with more ability. From the individual's viewpoint, it does not matter what the source of earnings enhancement is, be it production by schools or screening. The individual will be equally induced to make schooling investments based on the comparison of returns and costs. The two may, however, yield quite different incentives to governments to invest, because signaling may lead to different social and private returns to schooling. As a general matter, these models are not identified with just labor market outcome data. A variety of specialized tests under different maintained assumptions about individual motivations and firm behavior have been conducted but have not provided clear support for screening. These tests include looking for "sheepskin effects," particularly high returns to completing given institutional levels, as in Layard and Psacharopoulos (1974). Some support of screening does come from analysis of incentives to complete high school when there are fewer college graduates Bedard (2001). See Riley (2001) for a review of general theoretical and empirical work. The key difficulty with these tests, however, remains that they focus on labor market outcomes, where the private returns to schooling are generally expected to exist independent of the underlying causal mechanism. The analysis below concentrates importantly on outcomes that relate directly to the schooling process (the point where the two models are hypothesized to differ significantly).

[8] See, for example, Michael (1982); Haveman and Wolfe (1984); Wolfe and Zuvekas (1995); and Leibowitz (1974). Many factors are unclear, however, because of questions of causality; see, for example, Farrell and Fuchs (1982).

local and national elections, and to be a better informed and more responsible electorate.[9] Increases in the level of education are associated with reductions in crime (e.g., Ehrlich (1975), Lochner and Moretti (2001)).

Recent economic studies argue that education may provide economic benefits to society greater than the sum of its benefits to individuals – by providing a rich environment for innovation, scientific discovery, education can accelerate the growth rate of the economy; see, for example, the analyses of growth by Lucas (1988), Romer (1990a), Barro (1991), Jorgenson and Fraumeni (1992), and Barro and Sala-I-Martin (1995). The growth effects depending on the aggregate level of education in the economy enter as an externality to the individual. (Estimation by Acemoglu and Angrist (2000), however, questions this effect, at least at the state level).

Education appears also to have helped to achieve both greater social equality and greater equity in the distribution of economic resources. Schooling was a centerpiece of the U.S. War on Poverty in the 1960s, and the benefits of improved schooling are demonstrated in comparisons of the earnings of different social and ethnic groups. Earnings by blacks and whites have converged noticeably since the Second World War, and much of this convergence is attributable to improved educational opportunities for African-Americans (see Smith and Welch (1989); Jaynes and Williams (1989)). However, as discussed below, that convergence slowed down noticeably in the 1980s with skill differences being cited as a prime determinant (Juhn, Murphy, and Pierce (1993)).

Nonetheless, while there are many well-documented associations between amount of schooling – either individually or in the aggregate – and desirable economic outcomes, significant questions remain about the magnitude and interpretation of these relationships. First, the association may misstate the causal

[9]The pattern of U.S. voting over time can be found inStanley and Niemi (2000). An analysis of the partial effects of educational attainment (which are positive in the face of overall declines in voter turnout over time) is presented in Teixeira (1992).

impact of changes in schooling for individuals and the aggregate.[10] Second, the average effects may not correspond to the marginal effects. Third, in general externalities have been notoriously elusive and difficult to estimate convincingly, and education proves to be no exception. Finally, the measurement issues, as highlighted in the next section, are significant. Each of these topics (with the possible exception of the last) has received surprisingly limited research and is a fertile area for future work. In many contexts, they are key to both analytical and policy concerns.

Quality Considerations

For most of the 20th century, the U.S. debate over the economic consequences of schooling concentrated on the amount of school attained or, simply, the quantity of schooling of the population. Policy deliberations focused on school completion rates, on the proportion of the population attending postsecondary schooling, and the like. And analyses of the benefits of schooling were most concerned with the effects of quantity of schooling—whether benefits are seen in terms of individual incomes or social benefits like improved voting behavior of citizens.

Most policy and analytical attention has now switched to quality dimensions of schooling. In the United States, with the slowing of individual income growth[11] and of income convergence by race,[12] improving the quality of schooling, or how much is learned for each year, has been seen as the natural policy focus. Similar concerns, albeit generally with a lag, have diffused to other developed and developing countries.

---

[10] For example, Bils and Klenow (2000) question the importance of education as a cause of growth, as opposed to the relationship going the other way around. See also the perspectives in Mankiw, Romer, and Weil (1992) and Benhabib and Spiegel (1994). At the individual level, see Card (1999).

[11] See, for example, Levy and Murnane (1992) for a review of U.S. earnings patterns. See also Welch (1999) for an update and interpretation of distributional patterns.

[12] Discussion of distributional issues including earnings differences by race can be found in Smith and Welch (1989), O'Neill (1990), Card and Krueger (1992b), Levy and Murnane (1992), Bound and Freeman (1992), Boozer, Krueger, and Wolkon (1992),Juhn, Murphy, and Pierce (1993), Hauser (1993), Kane (1994), Grogger (1996), Welch (1999), and Deere (2001).

The economic effects of differences in the quality of graduates of our elementary and secondary schools are much less understood than the effects of quantity, particularly with regard to the performance of the aggregate economy. The incomplete understanding of the effects of educational quality clearly reflects difficulties in measurement. Although quality of education is hard to define precisely, it is natural to use the term quality to refer to the knowledge base and analytical skills that are the focal point of schools. Moreover, to add concreteness to this discussion, much of the discussion will rely on information provided by standardized tests of academic achievement and ability.

Relying on standardized tests to provide measures of quality is controversial—in part because of gaps in available evidence and in part because of the conclusions that tend to follow (as discussed below). The contrasting view emphasizes measuring "quality" by the resources (i.e., inputs) going into schooling. Early investigations include Wachtel (1975), Akin and Garfinkel (1977), and Rizzuto and Wachtel (1980). Most recent along this line is Card and Krueger (1992a); see also the reviews of the discussion in Burtless (1996b) and Betts (1996). In the context of developing countries, where these issues might be more important, see Behrman and Birdsall (1983). A substantial part of the controversy relates to the adequacy or effectiveness of expenditure or resource measures as a proxy for worker skills (as discussed below). In the end, cognitive skill measures appear to be the best available indicators of quality and do relate to outcomes that we care about, where resource measures are quite inadequate.

A variety of studies of the labor market have been concerned about how individual differences in cognitive ability affect earnings and specifically modify the estimated returns to quantity. The early work was subsumed under the general topic of "ability bias" in the returns to schooling. In that, the simple question was whether the tendency of more able individuals to continue in school led to an upward bias in the estimated returns to school (because of a straightforward omitted variables problem).[13] These studies

[13]See, for example, Griliches (1974). More recently, see Taber (2001). Discussions of alternative approaches to dealing with ability bias can be found in Card (1999). That discussion in general does not consider school quality, although some of the formulations could be recast in that way.

have focused on the estimated returns to years of schooling, although that seems to be a badly formulated question. The correction most commonly employed was the inclusion of a cognitive ability or cognitive achievement measure in the earnings function estimates. In interpreting that work, one must believe that quantity of schooling is uncorrelated with quality as measured by tests of cognitive ability and achievement.

These studies, nonetheless, provide insight into quality measurement issues through their common control for cognitive effects on earnings. The results of the early work generally indicated relatively modest impacts of variations in cognitive ability after holding constant quantity of schooling.[14] In this work, there was no real discussion of what led to any observed cognitive differences, although much of the work implicitly treated it as innate, and not very related to variations in schooling.[15] Further, all of the early work relied on generally small and nonrepresentative samples of the population.

The most recent direct investigations of cognitive achievement, however, have generally suggested larger labor market returns to measured differences in cognitive achievement. For example, Bishop (1989, 1991), O'Neill (1990), Grogger and Eide (1993), Murnane, Willett, and Levy (1995), Neal and Johnson (1996), Currie and Thomas (2000), Murnane et al. (2000), and Murnane et al. (2001) each find that the earnings advantages to higher achievement on standardized tests are quite substantial.[16] These results are derived from quite different approaches. Bishop (1989) considers the measurement errors that are inherent in most testing situation and demonstrates that careful treatment of that problem has a dramatic effect on the

---

[14]This limited impact of cognitive achievement was also central to a variety of direct analyses of schooling that reformulated the earnings determination process such as Jencks et al. (1972) and Bowles and Gintis (1976) and Bowles, Gintis, and Osborne (2001).

[15]Manski (1993a) represents one strand more recent work with this same general thrust. He recasts the issue as a selection problem and considers how ability or quality interacts with earnings expectations to determine continuation in schooling. Currently, however, no empirical work along these lines identifies the quantitative importance of selection or the interaction of school quality and earnings in such models.

[16]Outside of the United States, few studies are available. One exception for developing countries that finds significant effects of cognitive skills on income is Boissiere, Knight, and Sabot (1985). The NAS/NRC study on employment tests Hartigan and Wigdor (1989) also supports the view of a significant relationship of tests and employment outcomes, although the strength of the relationship appears somewhat less strong than that in the direct earnings investigations. Nonetheless, it seems likely that, for the purposes here, the GATB may not be a good measure of the cognitive outcomes of schools and may not correspond well to standard measures of cognitive

estimated importance of test differences. O'Neill (1990), Bishop (1991), Grogger and Eide (1993), and Neal and Johnson (1996) on the other hand, simply rely upon more recent labor market data along with more representative sampling and suggest that the earnings advantage to measured skill differences is larger than that found in earlier time periods and in earlier studies (even without correcting for test reliability). Currie and Thomas (2000) provide evidence for a sample of British youth and rely on a long panel of representative data. Murnane et al. (2001), considering a comparison over time, demonstrate that the results of increased returns to measured skills hold regardless of the methodology (i.e., whether simple analysis or error-corrected estimation). Murnane et al. (2000) and Murnane et al. (2001) provide further evidence of the effects of cognitive skills (although offering some caution in the interpretation of strength of cognitive effects versus other traits). Ultimately, the difficulty of separating cognitive skills from pure schooling has made this estimation very difficult (Cawley et al. (2000), Heckman and Vytlacil (2001)) and leaves ambiguity about the exact magnitude of effects.

An additional part of the return to school quality comes through continuation in school. There is substantial evidence that students who do better in school, either through grades or scores on standardized achievement tests, tend to go farther in school (see, for example, Dugan (1976); Manski and Wise (1983). Rivkin (1995) finds that variations in test scores capture a considerable proportion of the systematic variation in high school completion and in college continuation. Indeed, Rivkin (1995) finds that test score differences fully explain black-white differences in schooling. Bishop (1991) and Hanushek, Rivkin, and Taylor (1996)find that individual achievement scores are highly correlated with school attendance. Behrman et al. (1998) find strong achievement effects on both continuation into college and quality of college; moreover, the effects are larger when proper account is taken of the endogeneity of achievement. Hanushek and Pace (1995), using the High School and Beyond data, find that college completion is significantly related to higher test scores at the end of high school.

This work, while less complete than might be desired, leads to a conclusion that variations in

---

achievement.

cognitive ability, as measured by standardized tests, are important in career success. Variation in measured cognitive ability is far from everything that is important, but it is significant in a statistical and quantitative sense.

The linkage of individual cognitive skills to aggregate productivity growth is more difficult to establish. There is no clear consensus on the underlying causes of improvements in the overall productivity of the United States economy, nor on how the quality of workers interacts with economic growth.[17] The analysis of the impact of schooling quality on cross-country differences in growth by Hanushek and Kimko (2000), however, suggests that quality may be very important and could even dominate effects of the quantity of schooling differences across countries. They develop measures of labor force quality based on several different international mathematics and science tests and then find these to be highly correlated with international differences in growth rates. The concern in such work is the direction of causality. While a series of specification tests in Hanushek and Kimko (2000) indicates that there is a causal relationship between quality and growth, the exact magnitude of the effect is open to question.

Parallel to the work on individual wage determination, a number of studies have also pursued how school resource differences correlate with economic growth. These differences, however, have not shown a close relationship with international growth (see Barro and Sala-I-Martin (1995), Hanushek and Kimko (2000), and Barro and Lee (2001).

[17]One observation is useful, however. When looking at the history of productivity increase in the United States economy, several distinct time periods stand out. Productivity growth continued at some two percent per year through the 1960s, but fell off subsequently - first to one percent in the 1970s and then to virtually zero in the 1980s. It subsequently rebounded in the 1990s. Noting that productivity changes in these time periods through the 1980s mirror the aggregate pattern of the Scholastic Aptitude Test (SAT) scores, which fell dramatically from 1964 through 1980 before partially recovering, some have gone on to presume that the test scores are driving the productivity changes. Such could not, however, be the case -- since, as Bishop (1989) makes clear, the test takers with lower scores remained a small proportion of the total labor force through the 1980s. Lower test scores in the 1980s may signal forthcoming problems, but they cannot be an explanation for past changes in the economy.

Aggregate Resources and School Outcomes

School policy in the United States and elsewhere has focused attention on quality issues and desires to improve student outcomes. It is useful in this light then to consider briefly how this policy attention has shown up and what the results have been. The simplest picture comes from the aggregate data.

The concern in the United States about the quality of schooling has undoubtedly contributed to the growth in spending on schools. The U.S. has increased the resources devoted to students dramatically over the entire 20th century, with per pupil spending rising at 3½ percent per year in real terms for a 100-year period (Hanushek and Rivkin (1997)). Importantly, between 1960 and 1995 when performance become available, real spending per pupil tripled.[18] Clearly some of this recent expenditure was required simply to compete with other sectors for highly educated women and does not represent an expansion of school activities, but, even allowing for this, expenditure shows a strong trend.[19]

While U.S. spending on education has increased significantly during the last quarter of century, quality of student performance measured by test performance has remained roughly constant. Beginning in 1970, the United States embarked on an extensive testing program for students, the National Assessment of Educational Progress (NAEP), that was designed to track performance over time. It appears that the performance of U.S. 17-year-olds has remained roughly constant over the thirty-year period of observation.

---

[18]The measurement of real increases in resources has been the subject of some controversy and is difficult to do with precision. The preceding calculations deflate nominal spending by the Consumer Price Index (CPI). This is an output price index and is likely to diverge from an appropriate input price indices. Education is a labor intensive industry, which historically has shown little productivity improvement. The consumer price, or alternatively the GDP deflator, indicate how much of society's resources are being devoted to schooling. But, because of real wage increases in the economy, input costs in the education sector are likely to rise more rapidly than the CPI, so that the CPI-adjusted increases will overstate the increases in real inputs to education (e.g., Scitovsky and Scitovsky (1959) or Baumol (1967)). Rothstein and Miles (1995) suggest an alternative approach of using a modified service-sector CPI. This approach– based on a different measure of output prices emphasizing the service sector–cannot, however, solve the problem of obtaining more accurate measures of input prices, although it can provide a means of developing comparisons of productivity change across labor-intensive sectors (Hanushek (1997b)). While development of accurate input price indices is difficult because of issues of quality measurement for teachers, the difference between CPI-adjusted and input-adjusted measures is important to keep in mind. Use of a simple wage index or of a measure of increases in salaries for college graduates nonetheless still shows large real resource increases.

[19] See Flyer and Rosen (1997) for a discussion of the competing forces on teacher labor markets.

The simple comparison of NAEP scores from the early 1970s through the 1990s, shows lower science scores, roughly the same reading scores, and higher math scores.[20] Obviously, a variety of factors could influence the aggregate pattern of performance and costs, including for example changes in the student population or the structure of schools–topics discussed below. Nonetheless, the aggregate comparison of resources and performance creates a *prima facie* case that performance of public schools warrants careful consideration.

While the United States remains near the top of all nations in terms of spending per pupil, a number of other countries now spend similar amounts or even greater amounts. Table 3 displays estimates of both the absolute levels of spending and the proportion of GDP per capita devoted to primary and secondary schooling.[21] The comparison of spending patterns across countries shows considerable variation, part of which might be accounted for by higher proportions of private spending for schooling.[22]

Interestingly, the amount of spending internationally does not have a very close relationship to the performance of students. Table 4 displays national scores of eighth graders and twelfth graders on the Third International Mathematics and Science Study (TIMSS), conducted in 1995. Countries are rank ordered from highest to lowest in each and compared to performance in the United States. (Bold indicates

---

[20]The earliest NAEP test date differs by subject area with the first science test in 1969, first reading test in 1971, and first math test in 1973. Tests have been given approximately every four years and also involve earlier ages. Each of the subject areas has exhibited some change over the entire time period, and, while only endpoints are reported, it is also true that each of the tests has been roughly flat during the 1990s. See U.S. Department of Education (2000).

[21]Such comparisons across countries are clearly difficult to do with any precision. The absolute spending patterns require an international exchange rate, but even then are prone to inaccuracies because of differences in teacher labor markets. Further, countries differ in what is included in statistics for spending on schools. The GDP comparisons get around the exchange rate issues but suffer from lack of any models of how spending should change with national income. See U.S. Department of Education (1996b) and Organisation for Economic Co-operation and Development (2001).

[22] Data on private expenditures are difficult to find on a consistent and complete basis. The Organisation for Economic Co-operation and Development (2001) tabulations for developed countries display private spending in the form of tuition and other expenditures on private schools. But they leave out private tutoring which appears to represent a significant investment in a variety of countries – mostly notably the East Asian countries. Moreover, countries use different definitions of what is included in school expenditures, of the age period for schooling, and the like. The OECD analysis attempts to standardize data collection, but this is obviously difficult to do completely.

significantly above the U.S., and italics indicates significantly below the U.S.). Clearly, national scores are not closely related to the spending rankings in the previous table. More systematic investigation reveals the same results: Performance on the international tests is not closely related to resources of the school systems in these countries.[23]

The overview of education attainment, spending, and performance demonstrates the importance of schooling to individuals along with the commitments of governments to the provision of schooling. This provides a backdrop for consideration of government's involvement.

Role of Government

As is well-known, the existence of large returns to quantity or quality of schooling does not by itself warrant large scale governmental involvement. Large returns imply that individuals have strong incentives to obtain schooling, without the intervention of government. There are several primary justifications generally given for the level of governmental involvement in education: externalities, economies of scale, market failures in general, and redistributive motives. In the presence of these, purely private decisions are unlikely to lead to optimal social decisions.

Externalities are frequently proposed as central to government's interest in education. In general, activities that are perceived to have significant externalities are prime candidates for increased governmental support. With positive externalities, Pigouvian subsidies can be used so that individuals make decisions in line with the appropriate social calculus. Or other governmental interventions might be used to move toward a social optimum. But, as is also well recognized, externalities are noticeably elusive, and, while optimal tax and subsidy policies in the face of externalities are well understood conceptually, few estimates of the magnitude of externalities exist anywhere. Nevertheless, economists and citizens, if polled on

[23]See Hanushek and Kimko (2000) for analysis of results on international mathematics and science tests through 1990; see Woessman (2000, 2001)) for consideration of the TIMSS scores across nations.

externalities in education, would probably support the view that education involves extensive externalities (Cohn and Geske (1990)).

Leading candidates for areas of external benefits involve citizen involvement in the community and government, crime reduction, family decision making and child upbringing, and economic growth. There is evidence that more schooling does have a positive impact in each of these areas.

In each area, a significant portion of the beneficial effect of education appears to come from comparing very low levels of school attainment with significantly higher levels. Thus, extensive discussions of the social benefits of schooling in developing countries would seem both warranted and correct.[24] It is difficult to have, for example, a well-informed citizenry when most of the population is illiterate. It may also be difficult to introduce advanced production technologies, at least in a timely manner, if workers cannot be expected to read the accompanying technical manuals.

On the other hand, even if accepting the importance of externalities at minimal levels of schooling, there is little reason to believe that there are constant marginal externalities when expanded on both the extensive and intensive margins.[25] Specifically, arguments about the social benefits of expanded education seem much stronger in the case of developing countries of Africa than in the case of the United States during the 21st century. Where half of the population has attended some postsecondary schooling, another year of average schooling seems unlikely to change dramatically the political awareness of the U.S. population. Similarly, if the average high school student scores somewhat higher on the National

[24] Interestingly, policy discussions of education in developing countries tend to concentrate most on private rates of return and the market outcomes of added schooling, even if they make some reference to other social benefits such as political participation and lower fertility. See, for example, Heyneman and White (1986), Psacharopoulos, Tan, and Jimenez (1986), and Lockheed and Verspoor (1991).

[25] This issue is raised by Friedman (1962) and remains for the most part in the discussions of college education in Hartman (1973) and Mundel (1973). None of these, however, provides empirical evidence on the existence or magnitude of any externalities. The early primer on externalities in education (Weisbrod (1964)) concentrates chiefly on geographic spillovers and fiscal effects and downplays the issues raised here. A discussion of the magnitude of externalities that is similar to the one here is found in Poterba (1996).

Assessment of Educational progress, it is doubtful that many would expect noticeable changes in the identified extra social benefits of education.

Although education may be associated with a variety of social outcomes, a particularly relevant question is whether there is a causal relationship or not. With very little done on even assessing the magnitude of effects – largely because of poor measurement, progress on understanding the underlying causal structure has been even more limited. For example, one of the few direct investigations of causation indeed opens serious questions about common interpretations. Behrman and Rosenzweig (2002) present evidence on the role of mother's education in the intergenerational transmission of skills. In this, they pay particular attention to identifying the causal impact of mother's education through use of identical twins and conclude that it is not only much smaller than believed but possibly negative.[26]

A leading candidate for potential externalities of education in the United States and other developed countries, however, would revolve around economic growth. If a highly skilled workforce permits entirely different kinds of technologies to be introduced, or to be introduced earlier in a development cycle, expanded education of an individual may indeed affect other workers in the economy. Or, if improved abilities of the best students lead to more rapid invention and development of new technologies, spillovers of educational investments may result. Nevertheless, little evidence exists that distinguishes externalities in economic growth from simply the impact of better workers and more human capital.[27]

Beyond externalities, government also has a natural role when there are other market failures. The most obvious possibility comes through capital market imperfections. If individuals cannot borrow against their human capital – because human capital is embodied in the individual – there may be underinvestment in education (cf. Becker (1993)). This possibility, only observable in postsecondary education when

---

[26]For an early study of family effects, see Leibowitz (1974).

[27] A recent consideration employs cross-sectional wage information to look at productivity spillovers and finds little evidence after considering endogeneity of schooling (Acemoglu and Angrist (2000)). These issues can also be found in discussions of endogenous growth models such as Nelson and Phelps (1966), Romer (1990b), Mankiw, Romer, and Weil (1992), Benhabib and Spiegel (1994), and Hanushek and Kimko (2000).

government freely provides lower education, has not received strong empirical support (Cameron and Heckman (1999)), but the current interventions in the market make it difficult to assess completely the importance of this. Further, if there are economies of scale, say, from some fixed components of school operations, governmental intervention may provide for efficient operations. Nonetheless, while the empirical analysis is thin, little support for pervasive economies of scale exists.

An additional imperfection that deserves mention, and that enters into the discussion later, involves information. Student achievement involves a complex mixture of educational inputs including the student's own abilities, the influence of parents and friends, and the impact of schools. These factors are not easily separated, so that individuals themselves may have trouble assessing the independent influence of schools. If such is the case, informational problems may impede the decision making of individuals in terms of human capital investments. It may be that government can produce superior information about the quality of school inputs than the individual – although, if this is the rationale for governmental involvement, the form of intervention is important. In particular, government would not need to operate schools in order to provide information about their performance.[28]

The second major category of justification for governmental intervention is a redistributive motive. If society has certain goals for the distribution of income and well-being in society, normal market operations are unlikely to achieve those goals. The precise form of societal goals and the relationship to schooling has not been given much attention.[29] (Note also that redistributive goals may also interact with

[28] As mentioned before, a different aspect of information failures would relate to signaling and screening models. A common interpretation of these models is that schooling does not increase productivity, it only identifies more able people through their use of school attendance to convey their ability. In such a case, the social returns to schooling may be considerably less than the private returns – indicating that government should work to lessen the amount of costly schooling. Or, government may also not what to pursue programs designed to reduce school dropouts if it lessens the information on individuals that is available.

[29] Fair (1971) considers optimal policies for income distribution when it is an explicit component of the social welfare function. Becker and Tomes (1976, 1979) concentrate on intergenerational aspects of income transmission and distribution. Behrman, Taubman, and Pollak (1982) pursue intergenerational distribution issues from an alternative model of parental behavior. Hanushek, Leung, and Yilmaz (2001) evaluate using education as a redistributive device compared to the other mechanisms of negative income taxes and wage subsidies.

concerns about capital market constraints, where the desire is to break any linkages of poverty that exist because parents cannot provide appropriate schooling opportunities to their children).

An alternative redistributive motive actually appears to guide much policy and to interact with a range of policy initiatives discussed below.[30] Because housing in the United States and in many other places tends to lead to concentrations of poor people, minorities, and others who traditionally have not faired as well in schools, schools tend also to reinforce these concentrations.[31] To the extent that concentrations of poverty have added effects on schooling over and above individual poverty per se, there is an externality that interacts with any redistribution objectives, government may again have a clear role for correcting a market failure.

Without pursuing the details of any mandate for public action, however, two conclusions are important. First, while various market interventions are frequently employed to justify governmental intervention into education, very little explicit research or consideration has been given to the exact nature of these. For example, are the externalities related to the quality of schooling or just the minimal quantity? Second, as underscored by Poterba (1996), even less attention has been given to the appropriate mechanism for any governmental intervention. For example, if government wished to deal with capital market imperfections, should it provide free or reduced priced schools, make loans to students, operate the schools directly, or give grants to students to attend schools? A simple comparison is useful. The organization of the educational sector has moved toward government financing *and* government provision, similar issues in the health sector have led to very different institutional structures (at least in the United States). With few

[30] These issues arise most significantly when talking about policies that affect peer groups (e.g., desegregation policies) and policies that affect the financing of schools across local educational districts or agencies.

[31] Much of the past work on concentrations of poverty has involved crime, health, and welfare outcomes. As discussed below, the analytical complications of this work are serious. As Brock and Durlauf (forthcoming), Manski (1993b) and Moffitt (2001) point out, the empirical analysis of peer influences has been inhibited by both conceptual and data problems – problems that raise serious questions about interpretation of many existing studies. These critiques, in part precipitated by analyses of neighborhood poverty (e.g. Mayer and Jencks (1989), O'Regan and Quigley (1999), Rosenbaum and Popkin (1991)), point to a number of potentially severe empirical problems that are at least partially present in the recent set of randomized housing experiments aimed at understanding neighborhood effects (e.g., Rosenbaum (1995); Katz, Kling, and Liebman (2001); Ludwig, Duncan, and Hirschfield (2001)).

exceptions, little policy attention is given to any underlying consideration of the scope and form of governmental intervention.

The summary from considering the role of government is that the arguments for the currently large intervention – one quite generally including both financing and provision of services – remains not well analyzed. Thus, the remainder of this essay addresses a more limited issue: how well does government do at what it is trying to do.

Efficiency of Production

Because of the heavy involvement of the public sector in the actual provision of schooling, understanding the efficiency of production becomes an important issue. With competitive, private provision, little attention is given to economic or technical efficiency. Barring obvious market imperfections, there is general faith that market forces will push firms toward efficient use of resources. Even with market imperfections, there is generally little attention given to issues of technical efficiency, because firms are presumed to produce the highest possible levels of output given the chosen inputs–even if the firms are producing at the wrong level or not using the economically best set of inputs. But, the involvement of government in production, frequently in near-monopoly situations, alters the focus considerably. The possibility of inefficient production becomes a much more serious concern.

The attention to performance and efficiency in education began chiefly with *Equality of Educational Opportunity* (the "Coleman Report"), a U.S. government publication that appeared in 1966 (Coleman et al. (1966)). The specific focus of the Coleman Report, mandated by the Civil Rights Act of 1964, was the extent of racial discrimination in U.S. public schools. Two aspects led to the broad attention given to it and contributed to the controversy that has followed. First, it took the position that the central issue was not so much governmental inputs to schooling as it was student outcomes. Second, it is widely interpreted to imply that "schools are not important."[32]

[32] The Coleman Report was heavily criticized on methodological grounds (Bowles and Levin (1968); Cain

Public programs are frequently measured by the magnitude of public spending on them or the array of specific real resources (personnel of various types, capital in buildings or laboratories, etc.) going into a program. An extension of this is that variations in spending or resources indicate varying amounts of public involvement. The presumption behind employing spending measures is that funds are used effectively, implying spending is a simple index of the outcomes. The presumption behind the use of real resource measures is that the specific resources are important components indicating differences in quality, even if input prices vary across schools. The Coleman Report, which was required to look at the extent of racial discrimination in the public provision of schooling, needed a measure of the importance of various inputs to the schooling process. This requirement led the researchers to turn to measuring student outcomes and to relating various inputs directly to outcomes. This focus, which had not been applied very broadly in education or in other areas of government-provided services, dramatically changed the basic form of analysis.

Much discussion of schools tends to use the terms "quality" and "resources" interchangeably, but this usage presumes efficient operations of schools. A central part of the analysis discussed here looks directly at aspects of how effectively public schools use resources – and concludes that considerable inefficiency in resource usage exists. At a minimum, school quality should not be simply measured by expenditure patterns or by specific resources. More importantly, policy should logically reflect this reality.

The attention to the Coleman Report reflected the popular interpretation of the analysis that "schools do not make a difference." That interpretation, as discussed below, is incorrect, but the basic evidence behind the statement suggests serious performance problems of government, because purchased inputs to schools are not closely related to outcomes. Evaluation of the alternative interpretations, nonetheless, requires more general treatment of the educational production process and the empirical results that are available.

---

and Watts (1970); Hanushek and Kain (1972)). The most serious issue was the use of an analysis of variance procedure that biased the findings against any school factors being important and toward family factors. As discussed

General structure

The framework of analysis of educational performance considers a general production function such as:

(1) $\mathbf{O}_{it} = f(\mathbf{F}_i^{(t)}, \mathbf{P}_i^{(t)}, \mathbf{S}_i^{(t)}, \mathbf{A}_i) + \upsilon_{it}$

where O_{it} =performance of student i at time t, $F_i^{(t)}$=family inputs cumulative to time t, $P_i^{(t)}$=cumulative peer inputs, $S_i^{(t)}$=cumulative school inputs, A_i =innate ability, and a stochastic term, υ_{it}.

This general structure has motivated an extensive series of empirical studies. The typical empirical study collects information about student performance and about the various educational inputs and then attempts to estimate the characteristics of the production function using econometric techniques.

Two aspects of this formulation are important to point out. First, a variety of influences outside of schools enter into the production of achievement. Second, the production process for achievement is cumulative, building on a series of inputs over time. Both of these are important in the specification and interpretation of educational production functions.

The relevance of many factors outside schools highlights the necessity of going beyond simple comparisons of student performance across schools. Most of the attention in analytical studies has focused on the measurement of school attributes. This focus seems natural from a policy point of view. It also reflects the common use of administrative data in estimating production functions, because administrative data are frequently short of many measures of family background. Nonetheless, this lack of attention is unfortunate. First, increasing attention has been given to potential policies related to families – such as preschool and daycare programs, after school programs, parent education and the like. Second, because

---

below, this study also confused measurability of inputs with the importance of teachers and schools.

families frequently exert preferences on the schools that their children will attend, incomplete measurement of external influences on performance raise intense issues of selection bias and preclude simple statements about causal influences of schools. Such an observation of course does not seem very profound, but, as discussed below, many empirical studies give little attention to nonschool influences in addressing the impact of school factors. Moreover, public policy debates surprisingly frequently rely on simple accounting of performance across schools. For example, much of the current movement toward increased school accountability often relies on just aggregate student scores for a school.[33] Just the level of student performance is not a reliable indicator of the quality of schools students are attending.

The cumulative nature of achievement, where the learning in any time period builds on prior learning, implies that any analysis must take into account the time path of inputs. This places heavy demands on measurement and data collection, because historical information is frequently difficult to obtain.

The cumulative nature of the production process has been a prime motivation for considering a value-added formulation. At least in a linear version of (1), it is possible to look at the growth in contemporaneous performance over some period of time, instead of the level of performance, and relate that to the flow of specific inputs. The general value-added formulation can be written as:

$$\textbf{(2)} \qquad \mathbf{O}_{it} - \mathbf{O}_{it^*} = f^*(\mathbf{F}_i^{(t-t^*)}, \mathbf{P}_i^{(t-t^*)}, \mathbf{S}_i^{(t-t^*)}) + \upsilon_{it} - \upsilon_{it^*}$$

[33] With the increasing popularity of publishing average performance of students in different schools, the interpretation of scores becomes more important. In fact, without consideration of the various inputs that go beyond just schools, alternative accountability systems can have perverse effects (cf. Hanushek and Raymond (2001)). The integration of the underlying theoretical and empirical analysis of the determination of achievement with accountability and incentive systems is an important but underdeveloped area of investigation.

where outcome changes over the period (t-t*) are related to inputs over the same period. Note that this formulation dramatically lessens the data requirements and eliminates anything that appears as a fixed effect in the level of achievement (eq. 1).[34]

A final key issue is how student performance is measured. A prime justification for the attention to education, as described previously, is its hypothesized effects on labor market outcomes. The question remains about how best to measure educational output for understanding production relationships and policy options. With few exceptions (e.g., Betts (1995), Grogger (1996)), accurate measures of school inputs have not been related to subsequent earnings, making direct analysis impossible.[35] Thus, most analyses have conceptualized this as a two-stage problem: school resources and other things are related to test scores, school completion, or other intermediate outcomes, and these outcomes are related to subsequent success.[36]

Effects of teacher and schools

The most obvious complication of estimating models such as equation 1 is the necessity to specify precisely the various inputs into the production of student achievement. A logical starting place is estimation of the magnitude of differences across teachers and schools.

Consider

---

[34] This formulation presumes that innate abilities are constant and thus fall out of achievement growth. With more information on variations over time, it is also possible to allow for ability differences in growth (Rivkin, Hanushek, and Kain (2001)). Alternative formulations have prior achievement, O_{it*}, on the right hand side, allowing for coefficient different than one (Hanushek (1979)). This latter approach has the advantages of allowing for different scales of measurement in achievement during different years and introducing the possibility that growth in performance differs by starting point. It has the disadvantages of introducing measurement error on the right hand side and of complicating the error structure, particularly in models relying on more than a single year of an individual's achievement growth.

[35] Another class of studies, those aggregated to high levels such as the state level, have also considered labor market outcomes (e.g., Akin and Garfinkel (1977), Card and Krueger (1992a)). These studies, which introduce a wider set of analytical concerns, are discussed below.

[36] In more pragmatic terms, if interested in understanding policy influences on student outcomes, one would not want to wait decades until the ultimate impact in the labor market is observed.

(3) $\mathbf{O}_{it} - \mathbf{O}_{it^*} = f^*(\mathbf{F}_i^{(t-t^*)}, \mathbf{P}_i^{(t-t^*)}) + \sum t_j \mathbf{T}_{ij} + (\upsilon_{it} - \upsilon_{it^*})$

where $\mathbf{T}_{ij}$ is an indicator variable if student has teacher j during the period t-t*. This general covariance, or fixed-effect, formulation identifies teacher and school effects by systematic differences in achievement gains by students. In this formulation, teacher quality is measured implicitly by the average gain in achievement for the students of each teacher (adjusted for other factors influencing achievement).

Such analyses consistently show large and significant differences among teachers (e.g., Hanushek (1971, 1992); Murnane (1975); Murnane and Phillips (1981); Armor et al. (1976); Rivkin, Hanushek, and Kain (2001)). In the general formulation of equation 3, however, identification and interpretation of teacher and school effects is nonetheless complicated, since any factors that are constant across the period t-t* and across the students with teacher j are incorporated in the estimated effect, t_j. Thus, for example, teacher effects, school effects and classroom peer effects are not separately identified if the estimates come from a single cross section of teachers. Hanushek (1992), however, demonstrates the consistency of teacher effects across grades and school years, thus indicating that the estimated differences relate directly to teacher quality and not the specific mix of students and the interaction of teacher and students. Rivkin, Hanushek, and Kain (2001) remove separate school and grade fixed effects and observe the consistency of teacher effects across different cohorts – thus isolating the impact of teachers as opposed to just some combined teacher and classroom effects.[37]

The magnitude of differences in teacher quality is noteworthy. The estimated difference between a "good" and a "bad" teacher in poverty schools of Gary, Indiana, was approximately one grade level per

[37] The approach in Rivkin, Hanushek, and Kain (2001), however, relies on just variations within schools (having removed any fixed school and grade effects on achievement). Therefore, it cannot identify either the magnitude of between school differences in teacher quality or the importance of overall school differences such as that reflecting school leadership, quality of facilities, or the like. Even though for many purposes understanding the magnitude of between school quality differences is also important, an important objective of this paper is eliminating any possibility of selection effects.

academic year; i.e., a student with a good teacher might progress at 1.5 grade equivalents in a school year, while those with a bad teacher might progress at 0.5 grade equivalents (Hanushek (1992)). Alternatively, Rivkin, Hanushek, and Kain (2001) produce lower bounds on estimates of the variance in teacher quality entirely from heterogeneity of teachers within Texas schools. The estimates indicate that one standard deviation in teacher quality is equal to one-fifth of the average gap in performance between low income and higher income students. In other words, contrary to some conclusions emanating from Coleman et al. (1966), schools have the ability to compensate for educational differences arising from family backgrounds. A string of five above average teachers can, by the previous estimates, entirely close the average achievement gap by income level within Texas schools.

The identification of teacher and school effects also relies on the linearity of effects. If teachers, for example, have different effects on certain subgroups of students, the estimates of equation (3) do not separate out pure teacher effects. This problem, which also exists for estimation that relies upon specific measures of teacher characteristics, can be investigated through straightforward extensions of the model to allow t_j to vary across groups or according to other characteristics.

While estimation of this general fixed effect model demonstrates the significant impact of variations in teacher and school quality, it does not pinpoint the underlying characteristics or causes of these differences. Estimation by Murnane (1975) and by Armor et al. (1976) demonstrates that school principals are able to identify variations in teacher quality in the value-added sense of t_j. Thus, evidence suggests that quality variations are observable – an important issue in light of the next section that demonstrates that quality is not captured by measures of common characteristics such as degrees and experience of teachers.

Components of school inputs

The vast majority of analyses of educational production have pursued a different approach: the specification and estimation of hypothesized components of the teacher and school effects. This work returns to the specifications in equations 1 or 2 and attempts to incorporate specific measures of the

components of $\mathbf{S}_{it}$, the relevant inputs from the school. High on the list of characteristics has been a variety of measures of the resources going into schools, since budgeting for added expenditures of various sorts presents a series of obvious hypotheses about which factors influence achievement. This approach has not been very productive from the viewpoint of empirical description of the educational progress. But that in itself is significant because, as discussed below, these measured inputs are frequently the object of governmental decision making and policy.

Studies of educational performance include a variety of different measures of resources devoted to schools. Commonly employed measures include (1) the real resources of the classroom (teacher education, teacher experience, and class size or teacher-pupil ratios); 2) financial aggregates of resources (expenditure per student and teacher salary); and, 3) measures of other resources in schools (specific teacher characteristics, administrative inputs, and facilities).

The real resource category receives the bulk of analytical attention. First, these best summarize variations in resources at the classroom level. Teacher education and teacher experience are the primary determinants of teacher salaries. When combined with teachers per pupil, these variables describe the most significant variations in the instructional resources across classrooms. Second, these measures are readily available and well-measured.[38] Third, they relate to the largest changes in schools over the past three decades. Table 5 displays the dramatic increases in these school inputs for the U.S., with pupil-teacher ratios falling steadily, teacher experience increasing, and the percent of teachers with a masters' degree actually doubling between 1960 and 1990. Similar increases in resources have been well documented in other countries around the world (Organisation for Economic Co-operation and Development (2001)). Fourth, studies of growth in performance at the classroom level like equation 2, commonly thought to be the superior analytical design, frequently have these resource measures available but not the others.

[38] A majority of existing analyses of student performance has relied upon administrative data from schools. The real resource variables are commonly collected and reported within such data, and, because they are frequently the basis of payments or regulations across levels of government, they tend to be reported accurately.

The analytical design of studies of real resources stands in contrast with that for the other resource measures. The financial aggregates, particularly expenditure per pupil, are typically not even calculated for the classroom or the school, but instead are only available for the school district or for entire states or nations. As a result, studies employing these are typically the most aggregated studies, a source of analytical problems discussed below. Moreover, studies focusing on spending are not amenable to value-added specifications, causing the set of specification problems described previously. The study of spending is directly relevant for many policy discussions,[39] but these expenditure studies are noticeably lower quality than the best, and the typical, study investigating real classroom resources. The measures of other school resources typically are measured poorly and tend to be available only at the district level.[40] Since resources such as facilities tend to be relatively smaller in terms of overall spending, one would also expect these factors to be less important in determining student achievement.

Results of Production Function Estimation for the United States

Existing estimation of educational production functions provides considerable information about governmentally provided schooling. The intellectual heritage of this is the Coleman Report (Coleman et al. (1966)). This governmentally sponsored study spawned a large body of subsequent analyses, in large part because of its conclusions that cast doubts on the productivity on public schools. Perhaps more important, it set a standard of studying outcomes of schooling. Because of the common approach to studies conducted

---

[39]Some studies include expenditure per pupil along with measures of the real classroom resources. In such a case, since variations in classroom instructional expenditure are held constant, expenditure per student is interpreted as spending outside of the classroom. If only some of the classroom resources are included, the interpretation is more ambiguous and depends on the specific specification.

[40] For example, policy deliberations often consider the relative proportion of resources going to administration versus instruction. In the United States, the proportion of expenditures at the classroom level has fallen dramatically over time (Hanushek and Rivkin (1997)), leading some to view this as a measure of waste. However, without accounting for the uses of these resources and their effects on achievement, it is difficult to make such efficiency statements. Unfortunately, little work has concentrated on expenditures outside of the classroom. Data are imprecise, because even the definition of what are administrative expenditures is not settled. When available, administrative and other expenditure categories are generally not disaggregated at the school level. Similarly, measures of resources like laboratories or libraries are noted by their existence as opposed to any idea

since the Coleman Report, it is possible to provide a consistent summary of the results of how resources and other inputs affect student performance.

Providing a consistent summary is especially important in this area, because the large number of existing studies, taken individually, appear to support a number of disparate conclusions which do not hold when put in the context of the entire body of evidence. A tendency to focus selectively on a few studies and findings is exacerbated by the relationship between the research and serious policy discussions. The results of this research has been entered into legislative debates, judicial proceedings, and executive proposals – frequently placing heavy weight on the subset of studies that supports a particular position.

This discussion begins with tabulation of all studies meeting rudimentary quality standards (published in a refereed journal or book, including some measure of family background, and presenting information about the statistical properties of estimates). By including the universe of U.S. studies (available through 1994), the initial analysis provides an overview that is not affected by selectivity of results.

Estimates of key production function parameters come from 376 separate published estimates, found in 89 separate articles or books.[41] The estimated relationships vary in a variety of substantive ways (by measure of student performance, by grade, by included measures of resources). These studies also vary widely in quality, as generally captured by methodology and adequacy of data. Subsequent discussion considers how any results might be affected by key dimensions of focus and quality. In particular, the following sections delve into both analytical issues (methodology and data) and substantive specification issues (behavior and organization).

---

extent or quality.

[41] A more complete description of the studies can be found in Hanushek (1997a), which updates the analysis in Hanushek (1986). The tabulations here correct some of the original coding of effects in that publication. They also omit the estimates from Card and Krueger (1992b). In reviewing all of the studies and estimates, it was discovered that these estimates were based on models that did not include any measures of family background differences. This specification requirement is a minimal quality criterion, since omission will almost certainly lead to biased resource estimates. Family backgrounds have been shown to be quite generally correlated with school resources and have been shown to have strong effects on student outcomes.

Basic Results

Table 6 presents the overall summary of results of estimates of the effects of key resources for U.S. public schools. The tabulations note the numbers of separate estimates for each parameter along with their estimated direction of effect and with their statistical significance (5 percent level).

In terms of real classroom resources, only 9 percent of the estimates for the level of teachers education and 14 percent of the estimates for teacher-pupil ratios show positive and statistically significant effects on student performance.[42] These relatively small numbers of statistically significant positive results are balanced by another set finding statistically significant negative results—reaching 14 percent in the case of teacher-pupil ratios or the same percentage as finding the expected positive effect. Most estimates (72 percent of the teacher-pupil ratio estimates and 86 percent of the teacher education estimates) are statistically insignificant and those reporting the sign of insignificant estimated coefficients are split fairly evenly between positive and negative. A higher proportion of estimated effects of teacher experience are positive and statistically significant: 29 percent. Importantly, however, 71 percent still indicate worsening performance with experience or less confidence in any positive effect. And, because more experienced teachers can frequently choose their school and/or students, a portion of the positive effects could actually reflect reverse causation (Greenberg and McCall (1974); Murnane (1981); Hanushek, Kain, and Rivkin (2001c)). In sum, the vast number of estimated real resource effects gives little confidence that just adding more of any of the specific resources to schools will lead to a boost in student achievement. Moreover, this statement does not even get into whether or not any effects are 'large'. Given the small confidence in just

[42]The individual studies tend to measure each of these inputs in different ways. For example, while many studies include an indicator variable for whether or not the teacher has a master's degree, some will include measures of the graduate credits. With teacher-pupil ratio, some measure actual class size, while the majority measure teacher-pupil ratio. A variety of functional forms have been used, ranging from simple linear relationships to different nonlinear forms with thresholds, quadratics, and the like. In all cases, estimated signs are reversed if the measure involves pupil-teacher ratios or class size instead of teacher-pupil ratio. Further, where nonlinearities indicate positive effects over some range but not others, say with ranges of teacher experience, the most favorable for the hypothesis of positive effects is recorded.

getting noticeable improvements, it seems somewhat unimportant to investigate the size of any estimated effects, at least for the aggregation of studies.

The financial aggregates provide a similar picture. There is very weak support for the notion that simply providing higher teacher salaries or greater overall spending will lead to improved student performance. Per pupil expenditure has received the most attention, but only 27 percent of the estimated coefficients are positive and statistically significant. In fact, seven percent even suggestion some confidence in the fact that adding resources would harm student achievement. In reality, as discussed below, studies involving per pupil expenditure tend to be the lowest quality studies, and there is substantial reason to believe that even these results overstate the true effect of added expenditure.

Outside of the basic resource factors, a vast number of specific measures of teachers and schools have been included at one time or another, but few measures have been repeated frequently enough to permit any sort of tabulation. One set of exceptions involves either administrative inputs or facilities. While these categories include a wide range of specific measures, the results of such investigation as tabulated in Table 6 show little consistent effect on student performance.[43] An additional exception is teacher test score, where teachers have been given some sort of achievement or IQ test and their score on that has been related to their students' performance. Table 6 displays the results of the 41 studies that include teacher test scores. Of all of the explicit measures that lend themselves to tabulation, stronger teacher test scores are most consistently related to higher student achievement, even though only 37 percent provide positive and statistically significant effects.

Components of results

The forgoing tabulations combine the available evidence in a variety of ways – across grade levels,

[43]Administrative inputs are measured with such things as overall spending, the salaries of administrators, or the qualifications of administrators. Facilities include expenditures and specific measures such as availability of laboratories, the size and presence of a library, and the property of the school. In all cases, results are tabulated such that more of the measured characteristic means greater resources.

across measures of outcomes, and across studies of varying quality. While study quality is considered in the next section, the other issues do not have a strong influence on the overall findings. Real resources tend to show the same inconsistent pattern with achievement at both elementary and secondary school levels (not shown). Similarly the results of real or financial resources are not very different for studies focusing on test scores as the measure of performance and those focusing on other outcome measures. One specific issue has received extra attention. Do high resource schools encourage students to stay in school longer (which has obvious impacts on earnings)? Answering this question is, perhaps, more difficult than the straight achievement question, because labor market opportunities will affect the school completion decision as will net tuition and parental financial support when contemplating college. That question is a focal point of Hanushek, Rivkin, and Taylor (1996). In that study of school completion, school resources have no significant impact on student behavior once individual achievement and school costs are considered.[44] The issue of performance measurement has arisen specifically within the context of results for achievement tests and results for labor market outcomes (see controversies, below). Nonetheless, except for the differences in aggregation of the underlying estimation, no significant differences in results are found.

The role of peers

Schools are made up of teachers and other personnel but also include other students – peers. In fact, early sociological discussion of schools emphasized peers, and the Coleman Report (Coleman et al. (1966)) is commonly interpreted as arguing that peers played a more significant role in student achievement than did the resources and other formal structure of the schools. The interest in peers and integration of this

[44] Resource effects on college continuation are emphasized in Card and Krueger (1996). The estimation of resource impacts on school continuation is, however, particularly prone to specification problems. One might expect state effects to be particularly important in determining school continuation, since the availability and expense of public colleges and universities and the opportunity costs implied by different local labor markets vary significantly across states. The studies of college attendance in general do not control for interstate differences, leading to serious specification errors. (State policies and their impacts on production function estimates are discussed below). Betts (1996) reviews a number of these studies of educational attainment and does suggest some positive effects of resources. For the studies tabulated here (which differ from those considered by Betts), there tend to be positive effects of expenditure on school attainment, but there are only 25 total studies and only five estimated from within individual states. Thus, the small samples make it difficult to resolve this issue conclusively.

with Tiebout choices of families are set out in an interesting set of papers emphasizing the general equilibrium outcomes including de Bartolome (1990), Manski (1992), Benabou (1993, 1996), Epple and Romano (1998, forthcoming), Nechyba (2000, forthcoming-b), and Caucutt (forthcoming). A central element running through these papers is how financing mechanisms, particularly vouchers, interact with demands of families for different peers. While the structure of the vouchers and the mechanisms for individual behavior affect the solutions, these papers suggest that important peer influences can have direct ramifications for overall welfare and for the distribution of outcomes. These papers consider the outcomes of sorting, assuming that peer groups are important. Unfortunately, the development of empirical work has not kept pace with the theoretical investigations.

The empirical analysis of peers is very difficult to conduct and the results have been hard to interpret because they become confused with issues of the underlying estimation. Manski (1993b, 2000), Moffitt (2001), and Brock and Durlauf (forthcoming) describe a variety of econometric issues surrounding the topic.

Perhaps the most significant issue in estimating peer effects is the presence of omitted or mismeasured variables for schools and peers. Mismeasured individual factors or neglected influences that are common to the peers lead naturally to further identification issues, because mismeasured individual factors may be proxied by aggregate measures for the peers. Importantly, omitted variables bias will under very general circumstances lead to overstating peer influences, so peer effects can appear significant even in the case where they have no true effect Hanushek et al. (forthcoming). As the previous analysis of measured school factors indicates, characterizing school quality has been difficult, and thus it is highly likely that standard estimation of educational production functions with peers will overstate peer influences.

The theoretical econometric literature has concentrated more on the endogeneity of peer influences. Peer effects can be thought of as a simultaneous equation system where each student affects the others in the class and is in turn affected by the others, implying that standard issues of identification arise. These

issues are very difficult to deal with, particularly if the main effect of peers is through contemporaneous interactions. Specifically, if my behavior affects peer behavior and peer behavior affects me, estimation is problematic. Normal exclusion conditions, or even random assignment of students to classrooms, offer little hope in this case. If instead it is the characteristics of peers, such as how prepared they are for the curriculum or their general motivation and outlook, the development of both econometric estimators and the use of sample randomization becomes more feasible. In this case, nonetheless, the pure measurement issues still impinge on the ability to separate individual effects from peer effects.

Early empirical investigations provide mixed results about the importance of peers (Hanushek (1972, 1992)), Henderson, Mieszkowski, and Sauvageau (1976), and Summers and Wolfe (1977)). More recent analysis has taken alternative approaches to identification of peer effects. Hanushek et al. (forthcoming) consider changes in the peer composition that arise from small changes across grades and cohorts in demographics and prior achievement. They are able to eliminate any time invariant effects of schools and grades in schools through a fixed effect strategy and then to identify the effects of peers. They find small but significant differences of having smarter peers, and they find some effects of racial composition.

A separate strand of research, which has generally not been too careful about the structure of peer estimation, has looked at questions of ability tracking, or streaming, in schools. Conventional wisdom has suggested that heterogeneous grouping is good because the higher achievers help the lower achievers but are not affected by having low achieving classmates (Oakes (1985, 1992)). A careful review of this literature along with new estimation of tracking effects is found in Argys, Rees, and Brewer (1996). Nonetheless, this important policy issue deserves more attention, particularly in terms of the underlying methodological difficulties.

Peer effects have been considered in an international context by Zimmer and Toma (2000). Their analysis, using data from the second international mathematics and science tests given in 1981, estimated

achievement models across five countries (Belgium, France, New Zealand, Canada, and the United States) and incorporating public and private schools. They find that peer ability appears important, especially for low achieving students, and that peer effects may be more important in public than in private schools.

Because peer effects enter into a wide variety of other economic analyses and because they are the subject of intense policy deliberation, it remains surprising that there has not been more empirical research into the topic.

Racial integration

Racial integration of schools has been one of the most significant factors in United States public schools over the past fifty years. Yet, while there has been an enormous amount of legislative and judicial attention to racial desegregation of schools, the analysis of its effects is quite limited.

The U.S. Supreme Court ruling in *Brown v. Board of Education* (1954) held that separate but equal was unconstitutional because separate could not be equal.[45] This ruling led to dramatic changes in schools throughout the country. While school integration started slowly, the decade of the 1970s witnessed a substantial reduction in segregation brought about largely through legal pressure on local school districts (Welch and Light (1987)).

Many of the early (post-Brown) analyses, which focused on short run effects of purposefully moving students to less segregated schools, consider a variety of student outcomes and yield mixed effects of desegregation (Crain and Mahard (1978), Cook (1984), Armor (1995)). These studies are, however, plagued by methodological problems – largely related to sample selection issues but also including the heterogeneity of desegregation circumstances – making it difficult to assess the general impact of desegregation efforts.

Another segment of this literature focused on not so much on student outcomes but on the impact of desegregation efforts on the subsequent racial composition of schools. The large-scale exodus of whites

[45] *Brown v. Board of Education*, 347 U.S. 483 (1954).

from many cities and towns clearly dampened the impact of school desegregation on interracial contact, and beginning with analysis of "white flight" in the face of court ordered desegregation by Coleman, Kelley, and Moore (1975), much of the analytical focus shifted to outcomes defined in terms of racial contact (Welch and Light (1987), Clotfelter (2001)).[46]

Finally, a related but distinct strand of research focuses on whether peer racial composition, as opposed to desegregation actions, affects achievement for blacks as well as for other demographic groups. Coleman et al. (1966)) and its offshoots (U.S. Commission on Civil Rights (1967)) provided empirical evidence that racial isolation harms academic achievement.[47] Subsequent work by Crain (1970), Boozer, Krueger, and Wolkon (1992), Grogger (1996), and Hoxby (2000c) finds that school racial composition affected academic, social, or economic outcomes. Hanushek (1972) finds that higher concentrations of blacks hurts both whites and blacks, but is concerned that the racial composition of the school may simply be a proxy for heterogeneity in school quality and other omitted factors. Rivkin (2000) finds no evidence that exposure to whites increases academic attainment or earnings for Black men or women in the high school class of 1982, and Cook and Evans (2000) indicate that little of the black-white difference in National Assessment of Educational Progress scores can be attributed to racial concentration. On the other hand, Hanushek, Kain, and Rivkin (2001b) find that the racial composition of the school has a significant effect on black students, but not whites or Hispanics. Moreover, the negative effect of being in a school with higher concentrations of blacks is highly concentrated on blacks in the upper ability groups. The use of stacked panel data that permits controlling for other inputs in a very general manner suggests that unmeasured school quality can be ruled out as the primary cause of these findings and that specialized peer

---

[46] A related line of inquiry investigates racial composition and private schools (e.g., Clotfelter (1976) or Fairlie and Resch (forthcoming)). Whether or not private schools tend to be more segregated than public schools has also been the subject of considerable policy debate since Coleman, Hoffer, and Kilgore (1982). These issues are, however, beyond the scope of this analysis.

[47] The Coleman Report data, collected in 1965, largely reflect the legal and behavioral equilibrium before court ordered desegregation efforts, because most desegregation plans were instituted in subsequent periods (Welch and Light (1987)).

effects are the most likely explanation.

Most existing estimates have not paid close attention to the methodological issues surrounding identification of peer effects.[48] Moreover, the research has often failed to even separate out the most rudimentary differences in school quality.[49] Further research into aspects of the racial composition of schools, similar to peer research, appears to be a challenging but high value line of inquiry.

Family inputs

The emphasis of the original Coleman Report on the role of families in education has generated little subsequent disagreement. Most empirical studies of school performance include some measure of family background.[50] In fact, having some controls for family differences was a requirement for the previous tabulation of resource results, because ignoring these differences will undoubtedly bias resource coefficients through the correlation of family background and resources. Nonetheless, this research for the most part has not considered the details of what aspects of families are most important or of the causal structure of family effects.

The general argument has been that changing the important underlying characteristics of the family would be extraordinarily expensive even if policy makers wished to consider such an approach. Thus, for example, if mother's education were known to be important in a child's achievement, the best short run policy probably is not to send all mothers back to school. For this reason, researchers have not paid close attention to the precise measurement and specification of family effects.

On the other hand, long run policies may nonetheless reasonably relate to family factors. For

---

[48] A recent comprehensive review finds the evidence on achievement and psychological differences is very mixed Schofield (1995) and attributes the inconclusiveness largely to methodological shortcomings.

[49]See for example the U.S. Commission on Civil Rights (1967), which neglects any consideration of school factors Hanushek (1972). Of course, an alternative argument for policy purposes might be that only the reduced form relationship is needed, because the correlates of racial composition may adjust with desegregation. Moreover, much of the discussion about school segregation goes far beyond the simple discussions of student achievement.

[50] One large exception to this rule is that many studies of labor market effects or other performance measures that are collected after completion of schooling often neglect background information. See the review and discussion

example, arguments for improving women's education in developing countries may reflect the potential impacts on children's achievement more than the normal arguments about the return to the mother of human capital investment, particularly in countries where women do not work much in the formal market.[51]

But, if long run policies are directed at family factors, it is important to know the exact nature and causal impact of families. Is it the education of mothers that is important? Or is it the education of fathers (which often is not given the same analytical attention)? Or is it some other aspect of the family – wealth, attitudes, expectations, or other things – that is truly important and that shows up through its correlation with parental education?

The little work that has been explicitly related to family factors has opened serious question about the underlying causal structure. For example, Behrman and Rosenzweig (2002) suggest that mother's education may not be as important in children's schooling as commonly believed, once proper attention is paid to the possibility of omitted factors.[52] Similarly, Mayer (1997) questions whether income per se affects the kinds of family outcomes that are normally subsumed under the topic of poverty and income. Part of the issue in understanding families effects may arise from heterogeneous incentives within and across families (cf., Becker and Tomes (1976), Weinberg (2001)).

As policy debates venture into policies that are designed to change families, it is important to understand better the underlying structure. For example, as described previously, a variety of policies for developing countries point to family factors. But, in developed countries similar policy initiatives such as those designed to strength the family role or to ameliorate adverse family influences are increasingly discussed. The general problem has been that existing studies lack a convincing identification strategy that

---

in Betts (1996).

[51] These arguments also interrelate with a variety of economic models of fertility that consider trade-offs between child quantity and quality, although most do not directly consider child achievement; see Becker and Lewis (1973), Willis (1973), Hanushek (1992)).

[52] Their analysis relies on a sample of identical twins to separate the effects of mothers from other possible influences. While their cross-sectional estimates reproduce the common positive effects of mother's education, their refinements alter the sign.

distinguishes true structural aspects of families from a variety of correlated proxies.

The importance of study quality

The conclusions that should be drawn from the prior tabulations of results have been questioned because no weighting is used to distinguish among the quality of the studies. While the previous discussion presents the universe of available evidence, the studies of educational performance clearly differ in quality and the potential for yielding biased results.

Two elements of quality, both related to model specification and estimation, are particularly important.[53] First, education policy in the United States is made by the 50 separate states, and the variations in spending, regulations, graduation requirements, testing, labor laws, and teacher certification and hiring policies are large. These potentially important differences – which are also the locus of most policy debates in the states – imply that any studies of student performance across states must include descriptions of the policy environment of schools or else they will be subject to standard omitted variables bias. The misspecification of models that ignore variations in state education policy (and other potential state differences) interacts with the estimation approach. As Hanushek, Rivkin, and Taylor (1996) demonstrate, any bias in the estimated parameters will be exacerbated by aggregation of the estimation sample.[54] Nonetheless, the direction of any bias is an empirical issue, because it depends on the correlation of the

[53] Krueger (2000) introduces a different measure of study quality. His proposed measure is the number of separate parameter estimates in a given published analysis. So, for example, a publication that included estimates from a production function for eighth grade reading and one for high school graduation would necessarily be lower quality than a publication that only reported on third grade mathematics. Because this criterion is not related to conventional statistical arguments about model misspecification, estimation bias, or the quality of the underlying data base, it is not pursued here.

[54] Loeb and Page (2000) argue that in some circumstances, such as their analysis of compensating differentials in teacher wages, state aggregate data will be superior because it averages over choices across districts within states. To deal with state policy variations they concentrate on within state differences with a state specific time trend employing a panel of salaries and school attainment.

omitted state regulatory and finance factors and the included school measures such as class size or spending.

Second, as noted, education is a cumulative process, but a majority of studies are purely cross-sectional with only contemporaneous measures of inputs. In other words, when looking at performance at the end of secondary schooling, many studies measure just the current teachers and school resources and ignore the dozen or more prior years of inputs. Obviously, current school inputs will tend to be a very imperfect measure of the resources that went into producing ending achievement. This mismeasurement is strongest for any children who changed schools over their career (a sizable majority in the U.S.) but, because of the heterogeneity of teachers within schools, also holds for students who do not move (see Hanushek, Kain, and Rivkin (2001a)). Even if contemporaneous measures were reasonable proxies for the stream of cumulative inputs, uncertainty about the interpretation and policy implications would remain. The coefficients would bear an ambiguous relationship to the underlying structure parameters of interest, making policy calculations difficult. But there is little reason to believe that they are good proxies.

While judgments about study quality have a subjective element, it is possible to make a crude cut based on violations of these two central problems. Table 7 provides insight into the pattern and importance of the specific omitted variables bias resulting from lack of information about key educational policy differences across states. This table considers two input measures: teacher-pupil ratio and expenditure per pupil. These inputs, on top of being important for policy, are included in a sufficient number of analyses at various levels of aggregation that they can provide direct information of potential misspecification biases. As discussed previously, the percentage of all estimates of the impact of teacher pupil ratios with significant positive estimates is evenly balanced by those with significant negative estimates. But this is not true for estimates relying upon samples drawn entirely within a single state, where the overall policy environment is constant and thus where any bias from omitting overall state policies is eliminated. For single state studies, the statistically significant effects are disproportionately negative. Yet, as the samples are drawn across

states, the relative proportion positive and statistically significant rises. For those aggregated to the state level, almost two-thirds of the estimates are positive and statistically significant. The pattern of results also holds for estimates of the effects of expenditure differences, which are more likely to come from highly aggregate studies involving multiple states. (Expenditure studies virtually never direct analysis at performance across different classrooms or schools, since expenditure data are typically available only at the district level. Thus, they begin at a more aggregated level than many studies of real resources.)

This pattern of results is consistent with expectations from considering specification biases when favorable, but omitted, state policies tend to be positively correlated with resource usage. As noted, while the direction of any bias depends on the degree of correlation, under quite general circumstances, any bias will tend to be more severe if estimation is conducted at the state level than if conducted at the classroom level (Hanushek, Rivkin, and Taylor (1996)).[55] The initial assessment of effects from Table 6 indicated little reason to be confident about overall resource policies. The refinement on quality in Table 7 indicates that a number of the significant effects may primarily be artifacts of the sampling and methodology.

The second problem is a different variant of model specification. Because education is a cumulative process, relating the level of performance at any point in time just to the current resources is likely to be misleading. The standard approach for dealing with this is the estimation of value-added models such as equation 2 where attention is restricted to the growth of achievement over a limited period of time (when the flow of resources is also observed).

Table 8 displays the results of studies that consider value-added models for individual students. The top panel shows all such results, while the bottom panel follows the earlier approach of concentrating just on studies within a state. With the most refined investigation of quality, the number of studies gets quite small and selective. In these, however, there is no support for systematic improvements through

[55] The discussion of aggregation is part of a broader debate trying to reconcile the findings of Card and Krueger (1992a) with those presented here. See discussion below.

increasing teacher-pupil ratios and hiring teachers with more graduate education.[56] The effects of teacher experience are largely unaffected from those for the universe of studies. Again, because of the small and selective nature of these value-added studies within a single state, uncertainty about the precise effect of ignoring history remains. They do, however, make a *prima facie* case that the prior results about the effects of specific resources were not simply an artifact of study quality.

International Evidence

Analysis of how schools affect student performance has been conducted considerably more in the United States than in other countries. Nonetheless, similar investigations have been conducted in other parts of the world, even though data limitations have tended to be more severe.

The evidence for countries other than the United States is potentially important for a variety of reasons. Other countries have varying institutional structures, so different findings could help to identify the importance of organization and overall incentives. Moreover, other countries frequently have much different levels and exhibit larger variance in resource usage, offering the prospect of understanding better the importance of pure resource differences. For example, one explanation of the lack of relationship between resources and performance in the United States is its schools may be generally operating in an area of severe diminishing marginal productivity where marginal resource effects are small. By observing schools at very different levels of resources, however, it would be possible to distinguish between technological aspects of the production relationship and other possible interpretations of the evidence such as imprecise incentives for students and teachers.

Analysis in less developed countries has shown a similar inconsistency of estimated resource effects

[56] Other possible explanations of the results also exist. For example, as discussed in more detail below, Rivkin, Hanushek, and Kain (2001) are able to find statistically significant but small effects of class size differences, and these effects differ by socioeconomic status of the family. This analysis, based on very large samples, suggests the possibility that other analyses lack sufficient statistical power to detect the small and varying effects of some of the school factors.

across studies. While these estimates typically come from special purpose analyses and are frequently not published in refereed journals, they do provide insights into resource use at very different levels of support. Table 9 provides evidence on resource effects from studies completed by 1990.[57] Two facets of these data compared to the previous United States data stand out: 1. In general, a minority of the available studies suggests much confidence that the identified resources systematically and positively influence student performance; and, 2. There is generally somewhat stronger support for these resource policies than that existing in U.S. studies. Thus, the data hint that the importance of resources may vary with the level of resources, a natural presumption. Nonetheless, the evidence still does not suggest that pure resource policies can be expected to have a significant effect on student outcomes.

There is similar evidence to that in the U.S. that very significant differences exist among teachers (Harbison and Hanushek (1992)). Their analysis in the very poor region of northeast Brazil finds very large differences in total teacher effects. Surprisingly these differences are not related to simple measures such as the amount of teacher education or experience, even though, for example, teacher education can be as little as four years.

The evidence on developed countries outside of the United States is more difficult to compile. The review by Vignoles et al. (2000) points to a small number of studies outside of the U.S. and shows some variation in them similar to that already reported among estimates elsewhere.

A set of consistent estimates for recent periods using the data from the Third International Mathematics and Science Study (TIMSS) is presented in Hanushek and Luque (forthcoming). They employ the data on variations in scores within countries. The 36 countries with complete data are weighted toward more developed countries but do include poor countries. They find little evidence that any of the standard resource measures for schools are related to differences in mathematics and science scores within

[57] This compilation of results from Hanushek (1995) incorporates information from Fuller (1985), Harbison and Hanushek (1992), and a variety of studies during the 1980s.

countries.[58] Moreover, there is no evidence in this consistent work that there are different effects of resources by income level of the country or by level of the resources. Thus, contrary to the conclusions of Heyneman and Loxley (1983), schools do not appear relatively more important for poorer countries.

Woessman (2000, 2001) looks at cross national differences in TIMSS math and science scores and concludes that the institutional structure matters importantly for achievement. By pooling the student test scores across countries and estimating models that include both school and national characteristics, he finds suggestive evidence that the amount of competition from private schools and the amount of decentralization of decision making to schools have significant beneficial impacts, while union strength is detrimental and standard differences in resources across countries are not clearly related to student performance. The limited number of national observations for institutions nevertheless leaves some uncertainty about the estimates and calls for replication in other samples that permit, say, variations within countries in the key institutional features.

While the international evidence has been more limited, this situation is likely to be reversed profitably in the future. A key problem historically has been less available performance data for different countries, but this lack of information is being corrected. As student outcome data become more plentiful – allowing investigation of value added by teachers in schools in different environments, international evidence can be expected to grow in importance.

[58] Estimation considered 9-year-old and 13-year-old students with the most countries available for the older populations (36) and fewer for the younger populations (21). Some countries were omitted because they failed to provide data on student backgrounds. The set of countries potentially offering data is shown in Table 4.

Interpretation of Evidence on School Performance

The previous work on educational production has provided substantial evidence that there are vast differences among teachers and schools. At the same time, these differences are not easily described by the resources employed or by any simple set of programmatic or behavioral descriptions. More importantly, since policy is often directed at the resource levels, questions about the effectiveness of the public provision of schooling naturally arise.

The evidence does not say that money and resources never matter. Nor does it say that money and resources could not matter. It simply describes the central patterns of results given the current organization and incentives in schools. Indeed, a plausible interpretation of the evidence is that some schools in fact use resources effectively but that these schools are counterbalanced by others that do not. At the same time, the expansion of resource usage unaccompanied by performance gains implies a high level of inefficiency in the current operations of schools.

The implications of these results, however, depends fundamentally on how the policy and decision making process is conceived. At one level, these conclusions clearly imply that educational policy making is more difficult than many would like. If resources had a consistent and predictable effect on student performance, policy making would be straightforward. State and national governments could decide how much money to invest in schools and could trust local districts to apply funds in a productive manner. But, the fact that local districts do not use funds effectively complicates this picture. The clearest message of existing research is that uniform resource policies are unlikely to work as intended.

The considerations of overall spending levels, either in legislatures or the courts, largely rest on the premise that local districts are best situated and motivated to use funds wisely and productively. The evidence currently does not support the effectiveness of local decision making in the current environment. There is ample evidence, moreover, that policy makers do not fully believe that local decision makers will do a good job. The extensive bodies of rules and regulations at the federal and state level are mainly

designed to ensure that local districts do not do undesirable things in operating their schools and indicate a considerable distrust of the motivations and/or abilities of local districts. This, of course, runs into similar information problems. To set regulations appropriately, one would need to know how resources or process considerations affect student performance—which we do not know in any way sufficient for designing most regulatory approaches to good schooling.

An alternative perspective is simply that current incentives, within the public provision of schooling, do not motivate schools toward improving student performance (Hanushek and others (1994)). The simple premise is that the unresponsiveness of performance to resources is largely a reflection that very little rests on student performance. Good and bad teachers or good and bad administrators can expect about the same career progressions, pay, and other outcomes—making the choice of programs, organization, and behaviors less dependent on student outcomes than on other things that directly affect the actors in schools.

Such a description is, however, itself much too simple, because we have limited experience with alternative incentive schemes (Hanushek and others (1994)). The alternative incentive structures include a variety of conceptual approaches to providing rewards for improved student performance and range from merit pay for teachers to privatization and vouchers. Performance incentives recognize that there might be varying approaches by teachers and schools that are productive. Thus, they avoid the centralized "command and control" perspective of much current policy. At the same time, they recognize that simply decentralizing decision making is unlikely to work effectively unless there exist clear objectives and unless there is direct accountability.[59] Nonetheless, while some evidence is provided below, limited information is available about the design and impact of alternative incentive schemes in schools.

The existing work does not suggest resources never matter, but it also cannot describe circumstances where resources are used well. It does clearly indicate that the current organization and incentives of schools do little to ensure that any added resources will be used effectively.

---

[59]While the decentralization considered here really refers to pure resource policies and general funding, the evidence supports this conclusion even at the level of school based management; see Summers and Johnson (1996).

Controversies

The preceding interpretations of the general ineffectiveness of school resource policies have been challenged on several grounds. The challenges generally suggest that the evidence and its interpretation may be biased.

Causality

A key issue in considering the results of the educational production function studies is whether they identify causal relationships. Resource allocations are determined by a complicated series of political and behavioral choices by schools and parents. The character of these choices could influence the estimates of the effectiveness of resources. Consider, for example, the result of systematically assigning school resources in a compensatory manner. If low achieving kids are given extra resources – say smaller classes, special remedial instruction, improved technology, and the like – there is an obvious identification problem. Low class size could simply mean kids need more help. Issues of this kind suggest both care in interpretation of results and the possible necessity of alternative approaches.

At the individual student level, correlations with aggregate district resources through either formula allocations or community decisions are not a major cause of concern. The classroom allocations may, however, be a concern. For example, within a school, low achievers may be placed in smaller classes, suggesting the possibility of simultaneity bias. Any such problems should be largely ameliorated by value-added models, which consider the student's prior achievement directly. The only concern then becomes allocations made on the basis of unmeasured achievement influences that are unrelated to prior achievement.

Lazear (2001) develops an optimizing model that provides motivation for the decisions of schools in setting class size. His theoretical model, based on profit maximizing schools, emphasizes the externality related to disruptive students. One key issue in the context of the previous empirical results is whether

variations in disruption probabilities should be thought of as exogenous or whether alternatively they represent components of the teacher's classroom management ability, i.e., elements of teacher skill. Such a distinction is obviously very important in interpreting the model and the implications of it. While his model underscores the ambiguities of estimating reduced form models, the appropriate approach for testing the overall model or for estimating the behavioral equations is not developed.

Particularly in the area of class size analysis, a variety of approaches do go further in attempting to identify causal effects, and the results are quite varied. Hoxby (2000b) used de-trended variations in the size of birth cohorts to identify exogenous changes in class size in small Connecticut towns. Changes in cohort sizes, coupled with the lumpiness of classes in small school districts, can provide variations in class size that are unrelated to other factors. After pursuing her instrument strategy, she finds no significant impact of class size on achievement. Other studies have used aggregation approaches (by school or state) to construct instrumental variables estimators for the effects of class size (Akerhielm (1995); Boozer and Rouse (1995)) and have found more positive effects.

Several international studies have also pursued instrumental variables strategies that rely upon specific institutional structure. Angrist and Lavy (1999) note that Maimonides' Rule requires that Israeli classes cannot exceed forty students, so that, like in Hoxby's analysis, the lumpiness of classrooms may lead to large changes in class size when the numbers of students in a school approaches multiples of forty (and the preferred class size is greater than forty). They formulate a regression discontinuity approach to identify the effects of class size, but many of their estimates also use class size variation other than that generated by the discontinuities. The results provide generally positive but varying support for the effects of class size in different grades and circumstances. Case and Deaton (1999) rely upon the distinct policy regime of apartheid in South Africa to identify the effects of changes in resources and class size for black citizens. They argue that mobility restrictions and white decision making break the possible correlation of resources and individual errors, allowing them to identify the causal impact of resources. They find a significant

positive relationship between pupil-teacher ratios (interpreted more generally as overall resources) and both attainment and achievement. Hanushek and Luque (forthcoming), using the TIMSS data, perform a cross-country analysis restricted just to rural schools where mobility is restricted and schools frequently have a single classroom in the relevant grade, eliminating the normal concern about compensatory decision making. They not only find little evidence of consistent benefits from class size reduction but instead tend to find class size is positively related to achievement.[60] A final alternative is Lindahl (2000), who hypothesizes that differential summer learning across students could mask the importance of differences in school resources. Employing achievement data on a sample of Swedish students, the effects of resources are identified by assuming that parental inputs are the same in the summer and in the school year, while school inputs operate only during the school year. He estimates significant effects of class size differences.[61]

Unfortunately, identification of truly exogenous determinants of class size, or resource allocations more generally, is sufficiently rare that other compromises in the data and modeling are frequently required. These coincidental compromises jeopardize the ability to obtain unbiased estimates of resource effects and may limit the generalizability of any findings. Rivkin, Hanushek, and Kain (2001), employing better data of multiple cohort panels, make use of exogenous variations in class sizes within Texas schools across grades and cohorts to identify the impact of class size. The very large samples in that analysis provide estimates of small but statistically significant effects of class size on performance for earlier grades but not later grades and that the effects vary with student background.[62]

---

[60] A second approach removes a common achievement difference if the student's classroom is below the average for the grade, thus removing an average compensatory effect (if it exists) for each country. This approach similarly leads to little difference from the basic estimation.

[61] This analysis follows prior work on summer learning differentials (e.g., Heyns (1978); [Entwisle and Alexander (1992)), but he also suggests implications for the analysis of school differences. He concludes that the previous value-added specifications may be tainted by differential summer learning.

[62] The large samples in that analysis, approaching one million students, permit very precise estimates of the small effects. Krueger (1999) suggests that one reason many estimates are statistically insignificant is that the underlying parameter is very small and the typical estimation sample is insufficiently large to discriminate between small and zero. Of course, the small size of the parameter is relevant for any policy deliberations, because small

In sum, these alternative approaches yield inconsistent results both in terms of class size effects and in terms of the effects of alternative methodologies. The results in each of these studies tend to be quite sensitive to estimation procedures and to model specification. As a group, nonetheless, the results are more likely to be statistically significant with the expected sign than those presented previously for all studies, but the typical estimate (for statistically significant estimates) tends to be very small in magnitude. At the same time, the results of these approaches are inconsistent in terms of statistical significance, grade pattern, and magnitude of any effects, making it difficult to understand to what circumstances any results might generalize.

Consideration of the determination of school inputs has not been undertaken systematically, although anecdotal evidence suggests that it should be given more attention. For example, many teacher contracts in the United States allow more experienced teachers to choose the schools where they teach, leading to the previous concerns about the underlying behavior behind estimated achievement and experience relationships (Greenberg and McCall (1974); Murnane (1981); Hanushek, Kain, and Rivkin (2001c)). At the same time, little has been done on overall expenditure determination and its possible effects.[63] Similarly, while various policies and incentives drive much of the education of teachers, the effects of this have not generally been integrated into the achievement analysis.

While these topics are reconsidered below, it is clear that further work into the determination of resource patterns and the relationship to production function analysis is an important area for additional research. The difficulty of course is finding ways to identify the separate relationships in a convincing manner.

---

effects must be compared to the costly nature of the interventions.

[63] Some work has considered specific expenditures, such as special education (see, for example, Lankford and Wyckoff (1996); Cullen (forthcoming)). General expenditure determination and its effects on production function estimation has not, however, received much attention.

Labor Market Outcomes

Taken as a group, the production function studies give little indication that variations of resources have anything to do with present variations in student performance. The overall findings related, for example, to teacher-pupil ratios are dominated by estimates for test score measures of outcomes (217 out of 276). Test scores are nonetheless not generally the ultimate object of educational policy, but their use instead reflects growing evidence suggesting variations in test scores have important effects in the labor market. However, the widely-publicized findings of Card and Krueger (1992a) are taken as contrary evidence to the general picture given previously by indicating that variations in school resources are related to earnings differences among workers.[64] Some discussion has suggested that the most important difference between this latter study and the bulk of those reported previously comes from the measurement of outcomes, i.e., labor market experiences versus (typically) test scores (see Burtless (1996a)).

On the other hand, the test score-earnings linkage previously discussed provides a *prima facie* case that outcome measurement is not the source of difference. Moreover, specific reviews of the larger set of estimates in the resource-earnings literature suggest no significant differences from the overall results presented above (Betts (1996); Hanushek (1997a)). Yet, because of frequent citations of the different findings by Card and Krueger, it is useful to investigate the possible sources of the differences.

The divergent findings may simply be explained by variations in the level and environment of schooling across different eras. The Card and Krueger (1992a) analysis begins with samples of adult workers from the 1970 and 1980 Censuses of Population and fills in information about the schooling circumstances of individuals from information about their year and state of birth. The workers in their sample attended schools between the 1920s and the 1970s, implying variations in the level of resources going far beyond what is found today. If added resources have diminishing effects on student achievement,

[64]The Card and Krueger (1992a) analysis of school resources and earnings is the most discussed, but it follows a larger line of research. See, for example, Welch (1966), Johnson and Stafford (1973), and Wachtel (1976). An insightful review of past studies that considers underlying characteristics of the studies is Betts (1996).

current school operations may be largely "on the flat" of the production function, while Card and Krueger observe ranges from the past where resources had stronger effects.[65]

A related possibility might be that the political economy of schools has changed over time. For example, with the rise of teachers unions and the resulting change in bargaining positions, resources might be used in different ways and have different student achievement implications now than in the past (e.g., Borland and Howsen (1992); Peltzman (1993); Hoxby (1996b)). In other words, it is quite possible that the enormous changes in educational resources did have an effect on outcomes in the first half of the 20th century, but that more recent studies are also correct in finding no effect for the sorts of resource changes discussed in current schools.

Nonetheless, the most important set of reasons for the different conclusions likely involves specification issues. To begin with, many of the direct analyses of earnings include just the level of school resources, but none of the other factors that might influence student achievement and skill development. For example, it is plausible that students attending schools with a high level of resources also have parents who contribute more time, energy, and money to their education. If parental inputs are left out of the calculation, any estimated effects of school resources would tend to overstate the true independent effect of resources.[66] Further, as pointed out above, aggregation of school inputs is also likely to exacerbate any biases due to specification issues (Hanushek, Rivkin, and Taylor (1996)). Most of the earnings analyses observe school resources measured only at the aggregate state level. The Card-Krueger estimate comes from resource data

[65]The key element is that at high levels of resources the marginal productivity of added resources may be significantly less than at low levels of resources. While not a direct test of this on-the-flat thesis, the lack of significantly stronger resource effects in developing countries introduces some question about this hypothesis; see Hanushek (1995), or, in a growth context, Hanushek and Kimko (2000).

[66] For example, Card and Krueger (1992b) do not include any family background factors, so – in addition to the general aggregations issues of missing state policies – the effects of pupil-teacher ratios will be confounded with any family factors that are correlated at the state level. They stratify by race but include no measures of family background variations that might permit identification of a pure effect of pupil-teacher ratios.

aggregated to the state level, but no measures of state policy differences are included, so their estimates are subject to this bias.[67]

As critiques by Speakman and Welch (1995) and Heckman, Layne-Farrar, and Todd (1996a, 1996b) show, the Card and Krueger (1992a) estimates are very sensitive to the specific estimation procedure. Moreover, the state earnings differences cannot be interpreted in terms of school quality differences in the way that Card and Krueger interpret them. In order to identify their estimates of school quality, Card and Krueger (1992a) must assume that the migration of people across states is random and not based on differential earnings opportunities. Heckman, Layne-Farrar, and Todd (1996a, 1996b) show that there is selective migration and that this fundamental requirement for their interpretation is untrue.[68]

Thus, while some of the discrepancy in results could be attributed to different changes in schools over time, much of it appears to arise simply from the analytical problems in Card and Krueger (1992a). In more general consideration going beyond this specific paper, there appears to be little difference in results.

Meta-Analysis and the Summary of Results

Alternative procedures exist for summarizing the results of studies. These approaches, more common to other fields of study, sometimes go under the general title of "meta-analysis." A preferred approach to assessing disparate results would involve combining the underlying data of the studies directly to develop statistical inferences and tests of hypotheses across the studies. Unfortunately the original data are seldom available for reanalysis—and even when they are, combining data from different sources can be difficult—which forces a variety of compromises in the aggregation of results. The previous data on studies

---

[67] If, on the other hand, there are important measurement errors in the school resources, aggregation could be beneficial because this would tend to average out any measurement problems. A central concern of Hanushek, Rivkin, and Taylor (1996) is distinguishing between the harmful effects of aggregation and model misspecification and the beneficial effects of aggregation and measurement error. That analysis rejects the hypothesis that measurement error is a primary element in the apparent importance of resources in the more aggregated studies.

[68] They also show that the results differ significantly across time and that they are very sensitive to the precise specification of the models. Speakman and Welch (1995) further show that virtually all of the effects of state school resources work through earnings of college attendees, even though the resource measures relate only to elementary and secondary schools.

in Tables 6-8 represent one approach to the aggregation of results, an approach which relies on the minimal set of factors standardly reported. But, instead of simply reporting the distribution of results, others have attempted to do formal statistical tests.[69] An attempt to apply formal statistical tests to education production function data is found in Hedges, Laine, and Greenwald (1994) and Greenwald, Hedges, and Laine (1996). They wish to do formal hypothesis testing using the available data from essentially the same set of published studies employed here.[70] The most basic problem with their statistical analysis is that it fails to address the fundamental problem of whether resources are consistently related to performance. Instead, the question they pose is whether there is any evidence that resources or expenditure differences *ever* appear to affect student performance. They explicitly test the null hypothesis that all parameters indicating the effect of a specific resource on student performance are simultaneously equal to zero; i.e., H_o: $\beta_1=\beta_2=\ldots=\beta_n=0$, where the β_i are the underlying parameters relating a specific resource to student performance in one of the n available estimates. If any single underlying parameter (i.e., one β_i) for the combined sample of studies across varied schooling circumstances is not zero, the null hypothesis is false (that is, someplace there is an effect on student performance). Their statistical procedures are designed in such a case to reject the null hypothesis, leading to acceptance of the alternative that at least one study indicated the resource was related to performance.[71] They reject the null, but it does not change the overall interpretation of the econometric results.

---

[69]The primary argument against the simple tabulations, or vote-counting, employed here derives from the stylized analysis of combining a series of small experiments employing tests with low power, where more studies can actually lead to false conclusions. These examples have little relevance to the statistical tests developed in a regression framework with the very large samples frequently available.

[70] Some of the problems with doing this are immediately evident. Combining testing information is best motivated from thinking about variations in estimates of a single common parameter, something that hard to define given the variations in underlying model specifications. More importantly, published articles frequently do not (and cannot) provide sufficient information. For example, if parameter estimates are correlated across studies, say because they reflect performance in different grades of one school district, estimation of the combined variance of the estimator would require knowledge of the covariances--something that is never provided. Problems on nonindependence enter into the tabulations previously presented, but they are clearly less central to the interpretation of the results than in the case of combined significance testing.

[71]In discussing precisely the issue of how to interpret rejection of this null hypothesis, Hedges and Olkin

Experimental Evidence (STAR Experiment)

In attempting to understand the effects of specific resources on student performance, an appealing alternative to econometric estimation is the use of random assignment experiments. Such an approach can overcome a variety of concerns raised above about selection, causation, and the like. In the mid1980s, because of ambiguity about the effects of class size on student performance, the State of Tennessee launched a random-assignment experiment in reducing class sizes. The design was heavily influenced by an early summary of research by Glass and Smith (1979) that suggested student achievement was roughly constant across class sizes until the class size got down to approximately 15 to 1. After 15 to 1, reductions in class size appeared to yield gains in student performance. Based on this, a group of kindergarten through third graders in Tennessee was randomly assigned to either large classes (22-24 students) or small classes (14-16 students).[72] Students were followed over time as they progressed from kindergarten through third grade.

The student testing shows that children in smaller classes did better at the end of kindergarten and that this better performance was maintained through the third grade. This evidence is sometimes taken as refuting the econometric evidence on teacher-pupil ratios that was presented earlier, but two aspects are important to consider. First, while the use of experimentation is an important research approach (Hanushek and others (1994); Mosteller (1995)), the actual implementation of this experiment is open to question. Second, the findings, which pertain to a very large policy change in very specific circumstances, yielded small and difficult to interpret results.

A number of questions arise about the quality of the randomization in the STAR experiment. Specifically, because of lack of data, it is difficult to assess the randomization of students or teachers into

---

(1985), state, "It is doubtful if a researcher would regard such a situation as persuasive evidence of the efficacy of a treatment" (p. 45).

[72]The design was actually more complicated. The large classes were broken into two groups, one with teacher aides and one without aides. To be eligible for participating in the experiment, a school also had to be large enough so as to ensure that there was at least one small and large class. For a description the experiment and the basic results, see Word et al. (1990).

the experiment. There was substantial nonrandom attrition from the experiment; 51 percent of students initially in the experiment left before the end of the experiment. Substantial numbers (up to 12 percent by test) did not take annual tests; significant proportions of students changed experimental group (with the largest numbers going from large to small classes) during the experiment. It is difficult to assess the impact of any of these with the existing data (see Hanushek (1999); Krueger (1999)).

The second issue is interpretation of the results. The initial achievement differences found in the year students entered a small class were maintained but did not become wider through the grades (Word et al. (1990); Krueger (1999)). Because they continue to get more resources (smaller classes), these resources should, according to the general hypothesis, keep producing a growing advantage. Figure 1 shows the difference in reading performance in small classes that was observed across grades in Project STAR. (The results for math performance are virtually identical in size and pattern). It also shows how the observed outcomes diverge from what would be expected if the impact in kindergarten were also obtained in later grades.

Even if taken at face value, however, a significant issue remains: To what circumstances should the results be generalized?[73] The Tennessee experiment identified performance differences related to very large changes in class size at the entry into school. No similar experimental evidence looks at schools outside of Tennessee, at later grades, at smaller reductions, or at different absolute levels of class size. The previous econometric evidence found that some introductions of lower class appeared effective, even if across the board results were unlikely. But, like the Tennessee results, there is no indication from the econometric studies of when class size reduction is likely to work.

In sum, the potential application of random assignment experiments to schools is a very important innovation. Indeed, it offers an attractive alternative to the more common event, moving directly to major

[73] A second significant issue, discussed below, is whether the resulting effects large enough to justify further policy initiatives? The magnitude of the difference in performance is very small: a 1/3 reduction in class size over four years produced an average gain of about 0.2 standard deviations of student achievement.

public funding of full-scale programs.[74] The existing results from the one major experiment that exists, however, do not offer any substantial contradiction to the previous econometric findings.

Consumer behavior

The concept of increased consumer choice with respect to schools has been high on a variety of reform agendas. At the same time there is considerable choice of schools that comes through residential choice that is frequently tied to specific schools. While moving residence is clearly costly and clearly provides a blunt instrument for school choice, it nonetheless operates on the margin. Each year one-fifth of American households change location, allowing considerable room for exercising consumer choice of schools. An issue here is whether this choice by consumers affects the results observed about the operations of schools.[75]

A general conceptual argument against the inefficiency in school provision revolves around consumers and keys on notions of Tiebout moving behavior (Tiebout (1956)). Consumers appear to desire higher spending for schools; if wasted, one would expect lessened demand and/or moving of consumers to districts where there was more efficient production. This in turn would put pressure on districts to become more efficient.

The evidence suggests that there is consumer pressure, but that it is insufficient to overcome the existing variations in efficiency. The one direct investigation of jurisdictional competition is Hoxby

---

[74]In 1996, the State of California moved to a statewide program of providing significant additional funds to all schools that lowered class sizes in primary grades to state-prescribed levels. This program appears to have been the policy implementation of perceived results from the STAR experiment. Having done this on a statewide basis, there is also no effective way to evaluate the results of such an initiative, so that neither California nor other states can learn from this program. The existing evidence provides little reason to be optimistic about the future achievement effects of this policy Stecher and Bohrnstedt (1999).

[75] A different perspective on Tiebout effects comes from consideration of estimation of expenditure demand functions, where choice of communities by families will generally lead to overstating the income elasticity of demand Goldstein and Pauly (1981), Bergstrom, Rubinfeld, and Shapiro (1982), Gramlich and Rubinfeld (1982), and

(2000a) that considers how the numbers of districts in a metropolitan area – a measure of the potential for Tiebout competition – affects performance of the public schools.[76] She finds that more competition increases outcomes measured by test scores, attainment, and early career earnings. She also suggests that districts are more efficient when there is more competition among districts.

Most work has not focused directly on the relationship between school efficiency and consumer behavior but instead has investigated pieces of consumer choices. First, consumers have been shown to be willing to pay for schools with higher test performance. In particular, estimation of hedonic price indices for housing, beginning with Kain and Quigley (1975) and Rosen and Fullerton (1977), demonstrate positive effects of school achievement levels on housing demand. More recently, Black (1999) refines this approach by considering houses very close to school district boundaries – where contamination of unmeasured neighborhood quality is likely to be small – and finds that test scores of local schools have a strong effect on house prices. Similarly, Weimer and Wolkoff (2001) confirm such within district capitalization of school performance. Second, consideration of the political economy of school budgeting suggests that budget maximizing bureaucrats may act to push expenditure above the optimum (Romer and Rosenthal (1979); Filimon, Romer, and Rosenthal (1982); Romer, Rosenthal, and Munley (1992); Inman (1987); Rothstein (1992)). By gaming voters in the face of budget reversion levels, bureaucrats are able to push the median voter above her optimum spending level. Third, investigations of demand for spending – which typically assumes that expenditure is the same as quantity demanded – have paid no attention to quality of governmental services (e.g., Borcherding and Deacon (1972); Bergstrom and Goodman (1973); Poterba (1997)). In short, most existing analyses do not consider either consumer awareness of inefficiencies or consumer reactions to any inefficiencies.

Rubinfeld, Shapiro, and Roberts (1987).

[76] The underlying analysis uses instrumental variables (based on local geography) to circumvent endogeneity problems with district choice. The estimates of Tiebout choice effects disappear when endogeneity is not considered. An alternative set of estimates of public school choice on outcomes is found in Hanushek and Rivkin (forthcoming).

An alternative view of the forces behind the inefficiency of schools relates to self-interest of school personnel. Pritchett and Filmer (1999) consider how the interests of school personnel align with the decisions that are made and conclude that such self-interest appears important. Because of the special role of school personnel in setting policy, these forces could distort the preferences and choices of consumers.

A basic problem with all of these indirect approaches is the failure to distinguish school quality, or the value-added by schools, from the level of achievement in a school. Indeed, this is almost certainly the problem facing consumers and offers an explanation for remaining inefficiency. Without information about the contributions of families and of schools, consumers face a critical information problem. This information problem, which by all accounts appears to be serious, will in turn lead consumers to make nonoptimal decisions about location, school spending, and the degree of inefficiency that exists in local schools.

The analysis of school changing by Hanushek, Kain, and Rivkin (2001a) does suggest that parents obtain reasonable (value-added) information about school quality, even though such information is not directly provided by standard accountability systems. A difficulty in analyzing choice through mobility is that families move for a wide variety of reasons including changes in jobs and family circumstances. Some of these underlying causes will directly affect student performance, and the analysis of school choice must take this into account. The analysis in Hanushek, Kain, and Rivkin (2001a) concentrates on long run achievement effects of a move, under the presumption that the effects of other factors die out after a move, and finds that families moving to and within suburban areas tend to increase their children's achievement.

Clearly, the role of consumers within the determination of school policies and outcomes is an area for research, because the role of consumer demand is central to many perspectives on schools. We return to part of this issue below.

Inefficiency and estimation approach

Another approach to estimating production frontiers involves variants of mathematical

programming models or imposition of assumptions about the errors in an econometric formulation to capture the frontier notion. The underlying production theory calls for estimating the maximum feasible output from a set of inputs. If inefficiency exists in some schools, estimation should conceptually take it into account. The approaches follow a variety of estimation strategies with data envelope analysis (DEA) being one of the more common (see, for example, the early work in Charnes, Cooper, and Rhodes (1978) or Ruggiero (1996) and the references in it). Related work along more conventional econometric lines concentrates on the specialized nature of the error term that would be generated in the context of a production frontier (e.g., Aigner and Chu (1968); Aigner, Lovell, and Schmidt (1977). An extension combining some features of both can be found in Grosskopf et al. (1997).

The programming model has been used chiefly to consider both the trade-off between alternative outcomes (e.g., math and reading or test scores and graduation rates) and to uncover the most efficient way to produce these outcomes. The basic idea is that an efficient school is one producing the maximum output for its inputs. The estimation then constrains schools to lie on or below the production frontier. As is typical of these approaches, the number of "efficient" schools will be the same as the number of separate inputs, and all other schools will be compared to the efficient ones. Since this approach provides a direct measure of efficiency, it is conceptually possible to investigate how consumers behave when faced with varying efficiency. These approaches, however, suffer from the presumption that the small number of inputs fully captures the relevant production factors and that no measurement error exists.[77] For computational reasons, only a small number of inputs are considered, and these often do not take into account the influence of family background or other nonschool factors. But to fit the underlying theory, it must be the case that other firms can reproduce the output of the efficient firm – which would not be really possible if performance relied on other unmeasured inputs such as managerial skill or missing dimensions of teacher quality.

The stochastic versions relax these assumptions by explicitly including a one-sided error for

[77] Because they do not begin with a stochastic formulation, the programming approaches typically do not yield information about the statistical significance of any parameter estimates.

efficiency (constrained to be less than zero) along with a symmetric error. That is, considering the error in equation (1).

$$\upsilon_{it} = \varepsilon_{it} + \eta_{jt}$$

where individual i attends school j during year t and school j has a stochastic efficiency term with $\eta_{jt} \leq 0$. If η is uncorrelated with the family and school inputs, then ordinary least squares estimation provides consistent estimation and, with specification of the distributions of ε and η, maximum likelihood provides efficient estimates. But, of course, the key assumption is that inefficiency is uncorrelated with any of the observed inputs – an unlikely event unless there is truly nothing systematic about the efficiency of schools. A variety of approaches expand on these models to incorporate both the notions of production frontiers involving multiple outputs and stochastic elements (see, for example, Grosskopf et al. (1997)), but these will be subject to the same underlying issues.

The investigation of inefficiency by schools is clearly a worthwhile pursuit, especially in light of the previous evidence about performance of schools. Nevertheless, the current limitations of programming solutions and other econometric approaches appear overwhelming.

Multiple outputs

Much of the analysis of school performance treats the production of achievement in a given area as independent of that in other areas. Thus, for example, the determination of mathematics performance is separate from that of reading performance. If, however, schools are actively deciding among outcomes and they compete for resources (generally denominated in terms of time), differences in performance among schools when measured in a single dimension may simply reflect different motivation to produce a given array of outputs. That is, a school emphasizing reading may look inefficient in the production of mathematics if the performance in mathematics is compared to the total set of inputs to the school. A

number of studies have directly investigated varying student performance in a multiple outcome fashion through traditional simultaneous equations estimation (e.g., Levin, 1970; Boardman, Davis, and Sanday 1974). The problem in this estimation is the difficulty in identifying the separate equations for each outcome – because it is generally difficult to find factors that affect one outcome but not the others. One novel approach, however, is Brown and Saks (1975). In their formulation, school districts are portrayed as maximizing a welfare function defined over both the level and the variance in achievement. They then estimate reduced form equations for each. They demonstrate that, while school resources are not systematically related to the level of achievement in a majority of the estimates, their apparent strength is increased by also considering the variance in performance. One difficulty with this approach is that it rests upon a presumed common objective function that is not observable. Thus, much like the problems of basing estimates of consumer demand functions on the researcher's presumption of the utility function, this analysis gets it power from specifying the elements of the common preferences of school districts. But, again, the largest problem is that identification of the underlying structural parameters is not possible, making the application of this less useful for any policy purposes.

The common solution for maximizing multiproduct production functions requires some source of the relative weights for each output – making prices particularly attractive in the analysis of general firm behavior. Such prices are not readily available for school outputs.

The previously discussed data envelope analysis (DEA) lists as one of its virtues the ability to handle multiple outputs. Contrary to the normal production theory optimization, outputs are not aggregated according to any simple value maximization process. A discussion of the criteria employed can be found in Grosskopf et al. (1997). A key element in this aggregation is that any inefficiency is found in the transformation of inputs into aggregate output, as opposed to finding the output combination that maximizes any social welfare function.[78]

---

[78] An alternative statistical approach to dealing with multiple outputs is the estimation of canonical correlation models (e.g., Gyimah-Brempong and Byapong (1991)). These methods use statistical criteria to maximize

Finally, by going to the dual of the production problem it is possible to frame the estimation in terms of cost functions (e.g., Downes and Pogue (1994) or Duncombe, Ruggiero, and Yinger (1996)).[79] With maximization, this approach provides an alternative estimation approach that can characterize directly the relationships among various outputs. But the estimation relies not only on maximization of outcomes but also on having accurate measures of input prices. Input prices are particularly problematic with respect to the primary input – teachers. As discussed below, the supply function for teachers of varying quality is unknown, making it difficult to integrate into the production analysis.

Many policy discussions relate directly to consideration of varying outcomes of the educational process, yet little is known either about the trade-offs that are present or about differential consumer demand for outcomes. As data on multiple outcomes increasingly become available, the possibilities of inferring consumer demand from behavior and choices expand.

Cost Considerations

One remarkable aspect of most considerations of educational performance and school decisions is a general lack of attention to the costs of alternatives. Perhaps because the evidence about systematic effects of different programs, of different attributes of schools, or of various other inputs has been so inconsistent, most attention has focused on whether or not a particular input has a positive and statistically significant effect on achievement. But knowing that something might be expected to improve student performance is insufficient to make it the object of public policy. One would clearly want to know whether there were alternative uses of funds that produced higher achievement, or even higher value outside of schools.

---

the correlation of aggregate outputs and inputs, but the criteria in general is not clearly related to underlying structural models or optimization on the part of schools.

[79] These dual approaches also yield qualitatively the same results as the direct expenditure studies except that the impacts are inverted. An expenditure parameter in a production function estimate that is close to zero translates into a very large cost of obtaining any given student outcome. Thus, for example, Duncombe, Ruggiero, and Yinger (1996) suggest implausibly that it costs two and half times as much to achieve a gain in achievement in some

The policy problem is easy to see in the case of discussions of policies to reduce class size. Most of the debate has centered on whether one can expect any achievement gain from a simple policy of reducing class sizes. But even if one takes the results of the Tennessee STAR experiment as accurate, they suggest a very small magnitude of effect on performance: a 1/3 reduction in class size over four years produced an average gain of about 0.2 standard deviations of student achievement (Krueger (1999)).[80] Any consideration of class size reductions should clearly be put in terms of the cost of the program and alternative uses of funds.

An important part of the problem, however, is that the cost of alternative inputs is not well understood. In order to assess the relative benefits of employing alternative inputs, it would be appropriate to weight by the cost of each input. Yet, particularly when the inputs reflect different attributes of teachers (e.g., experience) or inputs (e.g., science laboratory quality), little information is available about costs of the inputs.[81]

Teachers are perhaps the best example. It is possible to identify how effective different salaries are in attracting a pool of teachers without regard to their characteristics (Murnane et al. (1991)). It is also possible to make some assessment about the relationship between salaries and specific characteristics of teachers (e.g., Hanushek and Pace (1995); Ballou (1996); Ballou and Podgursky (1997)) or labor markets (Toder (1972); Antos and Rosen (1975); Chambers (1977); Levinson (1988)). More generally, compensating differentials for teachers may lead to nominal salary differentials that are misleading in terms of individual decision making (e.g. Loeb and Page (2000)). Finally, by looking across time, one can see how teacher salaries cut the distribution of all college graduates, as a possible indication of the overall

---

downstate New York districts as compared to the average school district in New York State.

[80] These estimates are roughly similar to those for disadvantaged fourth grade students found in Rivkin, Hanushek, and Kain (2001), but considerably larger than those for more advantaged students. Putting these results into cost terms reinforces the small magnitude of the estimates.

[81] Levin and McEwan (2001) describe cost analysis based on adding the average cost of components of programs. This approach moves in the right direction of introducing costs into educational decisions, but it cannot deal with quality concerns and other reasons why the average and marginal cost might differ.

quality of teachers (Hanushek and Rivkin (1997); Flyer and Rosen (1997)). However, even suspending concerns about the value of separate teacher attributes, the existing studies do not give sufficient information to integrate any potential benefits with costs. Little progress has been on characterizing the supply function for quality.

The analysis of teachers has been more detailed than analyses of other school inputs. Even such apparently straightforward calculations as the costs of altering class size have been generally neglected. When attempted, the difficulties become apparent, because the specification of precisely how the policy is implemented along with other complementary changes that would be required has received little attention (see, for example, Stecher and Bohrnstedt (1999)).

In many ways it is surprising that the discussion of costs is so short. Programmatic costs do not seem to enter systematically in much educational decision making, suggesting a clear reason for observed inefficiencies in resource use. It also suggests that investigation of costs, particularly for quality aspects of school inputs, is a fertile area of study.

Institutional Structure and Incentives

The existing evidence is conditioned by the organization and structure of existing schools. Specifically, the prevailing incentives in schools may not create very strong pressures for improving student performance (Hanushek and others (1994)). Altered incentives, on the other hand, could potentially lead to very different outcomes.

At the outset, however, it is important to note that a variety of incentives currently exist and that school personnel respond to them. Traditionally few incentives have existed with respect to improved student performance, but instead have pointed in other directions. For example, financial incentives from state payments to schools for children in special education have distorted classification rates into special education (Cullen (forthcoming). Or, unionization has influenced the efficiency of school and the patterns of their resource usage (Hoxby (1996b)). Public schools also appear to react to consumer choices, perhaps

because of the relationship to property values discussed previously. In any event, schools facing wider competition from other districts show improved performance and efficiency (Hoxby (2000a)). The issue is not the existence of incentives but the direction, force, and focus of incentives. While consideration of incentives derived from competition with public schools is considered next, some altered incentives have been attempted within the public schools, and they are the focus of this section.

Perhaps the most discussed incentive topic is merit pay – providing differential pay for teachers based on judgments of performance. Most people think this is a natural kind of policy, mimicking pay systems found in other industries. The common evaluation of those policies attempted is that they fail (Cohen and Murnane (1986); Hatry, Greiner, and Ashford (1994)). The generalizations are that these systems tend to evolve into systems of relatively uniform pay and that they often tend to end up being extra pay for extra work instead of extra pay for good performance. The lessons from existing attempts at merit pay, however, remain unclear. Because performance contracts exist broadly outside of schools, it seems necessary to describe why education is different from other sectors. For example, other areas, similar to education, have difficulty in measuring the contributions of each individual in the firm and need cooperation among workers – yet they maintain differentiated pay. Indeed, even closely related activities like business schools use substantial merit systems and find considerable success (Brickley and Zimmerman (2001)). Private nonsectarian schools appear to use merit pay more extensively than public schools (Ballou (2001)) and also rely more frequent judgments about teachers and to let go teachers who are not performing well (Ballou and Podgursky (1997)).

Part of the issue in assessing merit pay plans revolves around expectations for the incentives. Most merit pay plans have been evaluated in terms of the performance of the existing teachers. This assessment assumes that the most significant issue is whether or not teachers are trying hard to do well. If they are not, merit pay may induce more and better efforts, leading to improved student performance. An alternative view is that the most significant aspect of merit pay proposals revolves around changes in the stock of

teachers, or the selection issue. Merit pay schemes might provide incentives for better teachers to stay and for poorer teachers to leave and thus may have little to do with variations in effort. Existing evaluations have not focused on the employment aspects.

Schools have not made much use of personnel incentive systems developed in other industries. Nor have they considered the theoretical literature developed for other organizations. A key element of parts of this has to do with partial measurement performance and the potential incentives set up by this within organizations (e.g., Baker (1992, forthcoming)).

A key element in any individual-specific accountability system is accurate and reliable measurement of performance. On this score, a variety of approaches have related student performance directly to teachers. Most notable is the Tennessee Value-Added Assessment System (TVAAS) that is mandated by state law but does not include rewards or sanctions for individual teachers (Sanders and Horn (1994, 1995)). Alternative approaches have been developed in Dallas, Texas (Mendro et al. (1998)) and would include the previously discussed statistical models for estimating teacher effects. A variety of concerns have been raised about relying specifically on student test performance, including undue focus on tests that partially measure performance, but importantly there is evidence that principal ratings are highly correlated with estimates of value-added by teachers (Murnane (1975), Armor et al. (1976)). Nonetheless, analysis of methodology for evaluating teachers on the basis of performance and of incentive systems that utilize such information remains limited.

States have, however, moved away from developing teacher-specific incentive systems and toward group ratings and accountability. A currently popular reform approach – standards based reform and school accountability – can be thought of as providing different incentives, although some incentives in these reform efforts are quite confused (Hanushek and Raymond (2001)).[82] The clearest expression of this is the

[82] The evaluation of behavior under different accountability systems is currently limited, although it is likely to expand rapidly as increased testing and use of varying accountability systems expands. For example, Deere and Strayer (2001) suggest that individuals react to public accountability systems but that some of the reaction may not be desirable. An initial evaluation of the Florida program for using vouchers to deal with failing schools has led to considerable controversy over the interpretation of the incentives; see Greene (2001a, 2001b) and Carnoy (2001).

develop of accountability systems that report the scores of students in a school and that may condition rewards on student performance. The approach highlights student performance, often calling for explicit measurement of outcomes along with mechanisms for aiding school to perform better. For example, a number of states have moved to substantial rewards to schools that perform highly on the state standards.[83] On the other hand, little existing evaluation has been conducted.[84] Much of the discussion has concentrated on the explicit evaluation of outcomes (e.g., Koretz (1996)).

Some innovative theoretical modeling has begun to describe some of the competing forces generated by many of the common accountability systems that have emerged with standards based reform (Betts and Costrell (2001)). They demonstrate some surprising outcomes when accountability systems have multiple roles.

The movement of policy toward accountability system and standards based reform has some important antecedents in the work of John Bishop. In a series of papers he hypothesizes that centralized testing systems have powerful incentive effects on students (for example, Bishop (1996); Bishop et al. (2001)). He argues that the strength of incentives for students differ dramatically when there is a consequential central examination as compared with, say, relative grading by classroom teachers. While evaluating this proposition is difficult since testing regimes tend to coincide with state or national jurisdictions that make isolation of test effects difficult, the line of research highlights the importance of students in the educational process. Most of the other analyses of achievement has stressed the role of schools and has brought in students largely to the extent that they have fixed, but this line of inquiry emphasizes the role of student behavior in affecting achievement and therefore brings in the necessity of

---

[83] The idea of rewarding schools instead of teachers relates to arguments used against merit pay. Because teachers must work with each other in a variety of ways, an incentive system that led to "hoarding" of information by teachers or general lack of cooperation would be unfortunate. There has been little analysis of such approaches.

[84] Conducting evaluations of existing state plans for accountability is very difficult. Because many plans cover entire states, it is necessary to do comparative studies across states. But, many other state policies are simultaneously affecting schools in each state. As discussed previously, problems of misspecification and aggregation to state levels can severely distort any such analysis.

incentives for students as well as school personnel.

The general consideration of incentives provides the contrast between the traditional regulatory approach to running schools and the value of identifying outcomes without detailing the approaches to achieving them. The incentive approach has the advantage of not requiring detailed knowledge about the production process in various local situations. Incentives, however, are not without their own complications. Specifically, little experience indicates how to devise the best incentive structures, and the details within the complex organization of schools are almost certainly very important. Some early experimental work into incentive contracts provides a simple example. The U.S. Office of Economic Opportunity, as part of the War on Poverty in the late 1960s and early 1970s, experimented with performance contracting, i.e., providing firms direct payments based on the achievement of students to whom they offered remedial instruction. In the end, virtually nothing was learned from the experiment because the flawed contract led firms to react in ways that limited what could be learned about basic approaches to remedial instruction (Gramlich and Koshel (1975)). The conclusion from this experiment and other incentive designs is that considerably more research is necessary to understand the best approach to designing incentives. The possibility of such research is likely to increase noticeably as education policy increasingly recognizes the importance of incentives.

Public versus Private Provision

If one distinguishes between the financing and the provision of education, it is possible to consider whether private schools, which perhaps have different incentive structures, perform better than public schools. The importance of competition for public schools was emphasized early by Friedman (1962) and has been discussed and implemented in a wide variety of forms. The investigation of this has nonetheless proved to be difficult and controversial.

Since 1970, about 11 percent of the U.S. student population has attended private schools. The private schools are overwhelmingly schools with a religious affiliation, but the religious affiliation has changed dramatically over time. In 1997, 50 percent are Catholic, 35 percent were other religious schools, and the remaining 15 percent are nonsectarian, but in 1970 three-quarters were Catholic (see Sander (2001)). The private schools are restricted from receiving public funding, so those students attending private schools pay tuition. Additionally, since the taxes to support the public schools – most of which are property taxes at the local level – are paid regardless of attendance, students attending private schools in essence pay for both public and private schooling.

The fundamental analytical question that has dominated the literature is whether or not performance in private schools exceeds that in public schools *ceteris paribus*. In particular, if private school performance exceeds that in the public schools, is it good schools or good kids?

Unsubsidized choice

Uncovering the relative advantage of public and private provision of education is confounded by the underlying choice of schools. Consider the following stylized choice problem where attention is concentrated just on public and Catholic schools for now.

(4) $\mathbf{O}^p = f^p(\mathbf{F}, \mathbf{P}, \mathbf{S}, \mathbf{A}) + \upsilon^p$

(5) $\mathbf{O}^c = f^c(\mathbf{F}, \mathbf{P}, \mathbf{S}, \mathbf{A}) + \upsilon^c$

(6) $\text{Prob}(c) = g(\mathbf{F}, \mathbf{P}, \mathbf{S}, \mathbf{A}, \mathbf{X}) + \varepsilon$

Equations 4 and 5 (which follow directly from equation 1) describe production in the public (p) sector and Catholic (c) sector. Equation 6 describes the selection equation governing choice of Catholic schools based on the production characteristics and other factors, **X**, such as religious preferences.

The fundamental difficulty in uncovering the differences between public and private provision comes from the selection equation. It is clear from the beginning that families that send their children to private schools are different, because they could have attended the public schools but instead paid tuition to attend the private school. Specifically, if parents with the largest interest in schooling tend simultaneously to pick Catholic schooling and to provide more motivation or a better learning environment for their children, υ and ε will be correlated. As is well known, the estimated effect of schools will be biased if the production functions are estimated without taking into account the selection of parents.

The estimation of these relationships takes a variety of forms. The simplest forms of estimation pursue two similar strategies. First, assuming the impact of families, peers, and school factors is the same except for a level difference in performance, a pooled production function combining Equations 4 and 5 is estimated with the addition of an indicator variable for attendance at a Catholic school (Neal (1997)). Second, separate production functions are estimated within the public and Catholic school populations and then the difference in performance is calculated by evaluating the production functions at common levels of all of the inputs and comparing the predicted performance (e.g., Coleman, Hoffer, and Kilgore (1982)). Both of these approaches, however, require attention to selection.[85]

Perhaps the most common approach to selection correction is to consider how own religious preference affects choice of public or Catholic schools. The argument is that a Catholic background will increase the chance that a student attends Catholic schools but will not affect the school outcome. But this

[85] The investigation of relative performance of Catholic schools started intensively with Coleman and Hoffer (1987). That analysis estimated separate public and Catholic school production functions and then standardized for differing levels of inputs (**F**, **P**, and **S**) in deriving the differences in performance of the two types of schools. This analysis started a heated controversy (see, for example, Noell (1982); Murnane, Newstead, and Olsen (1985)) over the importance of selection in Catholic schools. Selection could also be dealt with through estimation of the choice equation (Eq. 6), although it is difficult to do this in a way that also can realistically identify achievement

instrument would not be valid if, for example, Catholics on average had higher or lower abilities or were more supportive of their children than nonCatholics. A better choice, however, is the density of Catholic households or schools in an area (e.g., Neal (1997)). In any event, some variant of religious affiliation is commonly employed, though the validity and usefulness of these instruments remains a matter of concern. The stringent requirements for a valid instrument are generally difficult to meet. An alternative approach is to consider value-added models where variations in early achievement are considered. This approach, found in Coleman and Hoffer (1987), implicitly controls for selection if the outside factors entering into achievement are summarized in the initial levels of achievement. Again, while this estimation strategy has appeal, assessing whether any further selection effects remain is difficult.

The common result from estimation is that Catholic schools tend to achieve higher graduation rates after making selection corrections and holding other input factors constant (see Evans and Schwab (1995); Neal (1997); Grogger and Neal (2000)). At the same time, the estimation of effects on measured test scores is not as clear. Estimation yields varying views of the effects of Catholic schools, and the results are frequently not statistically different from zero (see the summary in Neal (1998)).

A key element of evaluation of relative performance is defining the relevant comparison. As Neal (1997) demonstrates, comparisons of public and private schools differ significantly by the group considered. While suburban public and private schools appear roughly comparable, urban Catholic schools, especially those serving minority populations, appear noticeably superior on average to urban public schools. The reasons for this difference in results have not been fully analyzed, but it is consistent with differential abilities to move and to select schools. If poor and minority populations have limited ability to change residences, and thus to select school districts, opening up choice through Catholic schools offers more potential gain. It is also consistent with the possibility that urban schools are on average inferior to suburban schools (cf. Hanushek, Kain, and Rivkin (2001a)). Sander (1996), however, suggests that

---

relationships. See, on choice, Lankford and Wyckoff (1992).

heterogeneity of the population attending Catholic schools is important, with positive effects isolated in the non-Catholic school population and no gains for Catholic students.

The variation in results by the measure of outcomes, if not a statistical artifact in the small number of available studies, is difficult to interpret. Student achievement and graduation are closely related according to past estimation, so it seems surprising that the influence of Catholic quality shows up in one dimension but not the other. Yet, the differences in school completion continue through college graduation, suggesting that these are significant differences. One possibility is that the statistical problems with the achievement estimation are greater, but more attention is needed to the variations in performance across the different outcomes.[86]

A larger problem with this line of research is that it provides no information on the distribution of results. As discussed above, the bulk of research about educational production is aimed at understanding the wide variation in performance of public schools. It seems natural to believe that Catholic schools also exhibit wide variation in performance, although none of the existing analyses document either the magnitude of differences or the potential causes of such differences. For example, assume that there is a difference in the mean performance of Catholic schools compared to public schools. It is important to understand whether this comes from a simple shifting of the entire quality distribution or from a truncation in the bottom of the quality distribution or from a narrowing of the entire distribution with a loss at both the top and bottom of the quality distribution. With the character of underlying changes in quality for Catholic schools being unknown, there is an ambiguity about the exact lessons to be drawn for any policy purposes.

The comparison of Catholic and public schools is based entirely on performance, but a more complete analysis would also incorporate the costs of production. Many observe that Catholic schools

[86] The High School and Beyond survey, which is used in the Coleman work and others, employed a short test that is likely to be noisy. Further, the sampling design behind the National Educational Longitudinal Survey (NELS) employed in Grogger and Neal (2000) makes selection on test taking important but difficult to deal with.

appear to be run with significantly lower expenditure levels, but accurate cost accounting is not readily available.

One final area deserves attention. The focus on Catholic schools, or private schools more generally, is motivated by the idea that more choice will lead to competition among schools that will then lead to improve either in overall performance or in the effectiveness of resource usage. The competition for students argument is the commonly held explanation for any difference in public-Catholic performance. But, if that competition affects the Catholic schools, it should also affect the public schools. This effect on public schools is in fact the primary interest in private school competition, because even a general widening in the availability and access to private schools would leave a large public sector. Pursuing this theme, Hoxby (1994) suggests that public schools located in areas where there is more competition from Catholic schools perform better than comparable public schools facing less competition. This competition is estimated to affect the level of output measured by educational attainment and graduate rates. Sander (1999), on the other hand, does not find such spillovers, leaving the question of competitive impacts open.

Note, however, that if public schools react to the existence of private schools, the evaluation problem becomes difficult. Any estimate of the performance difference of Catholic schools would understate the true impact.

Subsidized choice

Much of the attention to choice and its impact on results emanates from early arguments by Friedman (1962) about governmental provision of schools. He argued that, even if government had a role in the financing of schools, it did not have to provide the service. Instead, by providing vouchers given to parents, citizens would be able to choose among potential schools – both securing an education that suited them and opening up competitive forces to improve the provision of schooling.

In the United States, with a few exceptions there has historically been limited general, publicly supported choice. A number of states and districts have supported some kind of choice among public schools. Part of this reflects efforts both to accommodate and to subvert the racial desegregation of U.S. schools (Armor (1995)). Another part reflects efforts to provide more options and, implicitly, to broaden competition. The latter efforts have included intradistrict and interdistrict school choice (Nathan (1989); Hanushek and others (1994)) along with new efforts to free individual "charter" schools (Nathan (1996)). In recent years, charter schools have dramatically expanded. [87] Hoxby (forthcoming)Subsidized choice plans are to date considerably more limited: publicly provided vouchers in Milwaukee and Cleveland and privately subsidized vouchers in Dayton, New York City, and Washington.[88]

The wide variety of choice options implies that the incentives under each can be quite different. For example, the impact of a plan permitting students to choose any public school in the district would have very different incentives depending on how any excess demand for a school is dealt with, what happens to teachers and administrators in schools with large number of vacancies, how transportation of students is handled, and the like. Similarly, when students are permitted to attend public schools outside of the district, issues about fiscal transfers, the handling of excess demand, and transportation become critical. For voucher plans, the size of the voucher, the eligibility of potential schools, the treatment of excess demand, the range of programs that are permissible, the legitimate grounds that schools can use to select students, and more will determine the character and results of the program. A variety of underlying theoretical models have identified specific elements of importance. For example, Manski (1992) concentrates on the

[87] Charter schools provide provisions for individual schools to opt out of district regulations and to provide an alternative program. The rules governing these schools vary widely by state, but the common element is that they are supported by public funds and remain officially part of the public school system. These schools must attract students who choose them over the regular public schools. While most states have less than one percent of their students in charter schools in 2000, some such as Arizona (3.7 percent) and Michigan (2.8 percent) have developed substantial competitive public sector. The District of Columbia has over 8 percent in charters. The start of this expansion along with a description of the kinds of schools in the charter sector is described in Finn, Manno, and Vanourek (2000)).

[88] Note that privately provided scholarship plans for attending private schools are actually much more extensive, but the individual plans generally have been small and have not been evaluated.

role of peers under different choice plans, while Chubb and Moe (1990) emphasize school decision making. Little empirical work has been done on identifying or assessing the key elements of choice plans.

The existing public school choice plans have not received thorough analysis in terms of student outcomes. A primary deterrent to analysis is disentangling selection from program effects. The issues are identical to those sketch for understanding the impact of Catholic schools, but the development of suitable instruments for the selection equation is much more problematic. One exception is an analysis of open enrollment plans in Chicago public schools that suggests most observed better performance by choice students comes from selection (Cullen, Jacob, and Levitt (2000)).

The voucher programs are different. The most thorough analysis of the voucher programs has occurred in Milwaukee. This publicly supported voucher, while not an experiment but instead an on-going program, had an evaluation plan set up at the outset (Witte (1999)). Subsequently, a number of studies, taking different approaches and reaching somewhat different conclusions, have looked at the same impact of vouchers in Milwaukee (Peterson, Greene, and Noyes (1996); Greene, Peterson, and Du (1998); Rouse (1998); Witte (1999)).

Three issues are important in the analysis: how to define outcomes, how long to wait to observe results, and what the comparison group of students should be. In terms of the outcomes of choice, the available analyses from the Milwaukee program suggest that parents choosing voucher schools are happier with them than with the public schools and that the choice schools generally spend noticeably less than the public schools. Thus, from a simple consumer viewpoint, the vouchers seem useful. Nonetheless, the majority of attention has focused on measured student achievement. The analyses of student achievement show no real gains in vouchers schools during the first years, but by the fourth year of operation voucher schools are doing as well as or better than the Milwaukee public schools, depending upon the precise performance measure. The findings on achievement also depend upon the precise comparison groups. Comparisons to people applying but not enrolling in the voucher program or to the low income eligible

population show the voucher schools on the whole as doing better, while comparisons to the entire public student population in Milwaukee show less gain.

The Milwaukee program is one specific set of rules and institutions. It is not a general test of vouchers versus public schools. At its inception in 1990, participating schools had to be already existing nonsectarian schools that also served a substantial unsubsidized population (Witte (1999)). Participating students had to come from low income families.[89] The number of participants was initially restricted to less than 1 percent of the public school population but grew to 15 percent by 1995. Therefore, this experiment does not represent a general test of vouchers, but instead a very specific kind of application of the ideas. For example, an important element of the original Friedman (1962) proposals is the induced supply response which engages markets in innovating in ways that appeal to consumers – something that is really not a part of the Milwaukee experience.

The voluntary, privately-funded voucher programs offer a separate view of the impact of vouchers. While these have only been operating for short periods of time, some preliminary results are available. Peterson et al. (forthcoming) suggest that these programs have had positive effects for some minority students but not for others. The variations in effectiveness thus leave open many questions that are central to many of the policy discussions.

The short run evaluations of performance in voucher programs, while interesting and useful, necessarily provide a limited picture. The primary arguments for introducing competition involve incentives for new private schools to develop and incentives for the public schools to improve. These are, however, long run effects, making the evaluation problems and perspectives on these programs analogous to evaluating merit pay plans for teachers.[90]

---

[89] Family incomes had to be less than 175 percent of the federal poverty level. The official poverty line for a family of four was approximately $17,500 in 1990.

[90] Note, however, that Hoxby (forthcoming) provides early results that suggest public schools respond quickly when they face direct competitive pressures.

Throughout most of the discussion, very little mention is made of costs. As operated, the existing voucher programs, including that in Milwaukee, appear to be running at noticeably lower expenditure than the comparable Milwaukee public school system. Part of this is easily explicable by the public schools' responsibility for taking all students, including high needs students such as those with handicapping conditions. But part also seems to be simple differences in cost structures. It is important to know f the voucher programs are truly operating at lower costs while producing the same (or better results).

International evidence

Other countries have of course operated much broader choice systems than found in the United States. While it is possible to find descriptive information and information about the flows of students, little is available about the relative performance of private schools under different institutional structures (see, for example, James (1987, 1993); Glenn (1989); West (1996); Cohn (1997), and the case studies in it; Fiske and Ladd (2000); and Jimenez and Sawada (2001)). Again, as with the analysis of production functions, developing information about the performance has been hindered by measures of student outcomes. Nonetheless, evidence on choice is beginning to develop, including McEwan and Carnoy (2000) and McEwan (forthcoming) on the Chilean experience, and Angrist et al. (2001) on Columbia.

International evidence currently offers the best chances for understanding the impacts of voucher systems, particularly the longer run implications. Not only is choice better established in many countries than in the U.S. but also the institutional structure for choice shows much more variation.

Financing of Schools

State governments in the United States have had the primary responsibility of the provision of schooling. The federal government has provided between 5-7 percent of the total funding for schools, and this funding has been focused on special populations – with economically disadvantaged and handicapped

students receiving the bulk of the funding.[91] The federal government has also promulgated laws with major effect on schools. The most important are Constitutional prohibitions against segregation of schools and federal legislation about service requirements and standards for handicapped students (so called "special education"). Because of the limited role of the federal government, however, this discussion will concentrate on state and local spending.

State constitutions invariably identify the role of the state government in establishing and running a public school system that is free to all students. While the exact arrangements differ, state governments have historically set the general operating structure for schools in terms of regulations and laws governing the schools and then given local jurisdictions the responsibility for running the schools. Although there is wide variation across states, local governments typically share funding responsibilities with the states. The determination of funding levels is the complicated outcome of local decisions, voter and political outcomes, and state rules and regulations.

Perhaps the principal policy lever of the states has been the funding level for local education. States have provided grants to local school districts to cover special purpose, or categorical, items, but most spending goes for general operating expenditure. The decisions facing state legislatures concern both the level and the distribution of spending. All states use their funding mechanisms to offset different abilities of school districts to raise funds locally. Particularly as the value of high quality education has become more apparent, states have turned their funding toward the goal of improving student achievement. The problem facing the legislatures has been an inability to turn funding into student achievement with any assurance – the general issue discussed earlier. One result has been an increasing tendency to add further regulations on activities along with further funding. This discussion, however, concentrates on the funding aspect and the implications of state fiscal federalism.

---

[91] Funding for primary and secondary schooling (and higher education) comes through the U.S. Department of Education. Funding for Head Start, another large education-related program for disadvantaged 4-5 year olds, has come through the Department of Health and Human Services and generally is not included in the usual accounting of total spending for education.

The discussion of state funding is heavily influenced by legal cases. School finance in the United States was relatively stable through much of the 20th century. Early developments emphasized the role of states in ameliorating the largest differences in local wealth (Strayer and Haig (1923)). In the late 1960s, the case of *Serrano v. Priest* was entered into California state courts and changed the stability in financing.[92] This case alleged that the state funding of schools violated the 14th amendment of the U.S. constitution that requires equal treatment of individuals under the law. The state funding provided a portion of the funds needed by local districts, and districts were expected to raise the remainder. California like a majority of other states permitted local jurisdictions to use the property tax as its major revenue source. The *Serrano* suit argued that, because local property tax bases differ, this funding arrangement discriminated against students in poor districts by making the funding dependent upon the wealth of others in the community (Coons, Clune, and Sugarman (1970)). While the U.S. Supreme Court ultimately ruled in 1973 that the state funding formula did not violate the federal constitution,[93] the *Serrano* case and similar subsequent ones in a majority of the states were argued on the grounds that the method of funding violated the individual state constitutions. While court cases in states continue, few general conclusions can be drawn about the legal decisions that have variously upheld and struck down existing state funding formulae.

Similar policy dilemmas face the courts in school finance cases as face the state legislatures that are formally charged with funding decisions. The courts have entered into education decision making in deciding on suits brought by people who argue that their districts are receiving insufficient state funding. While frequently motivated by concerns about student achievement, in reality both the judicial statement of the issue and the proposed remedies invariably revolve around the level and distribution of resources.

An important issue in these discussions is the effect of financing on student outcomes. While explicitly considered below, the prior discussions on the limited evidence about any consistent relationship

---

[92] *Serrano v. Priest*, 5 Cal. 3d 584,589, 487 P.2d 1241, 1244, 96 Cal. Rptr. 601, 604 (1971).

[93] *San Antonio Independent School District v. Rodriguez*, 411 U.S. 1 (1973).

between resources and student performance should provide an early note of caution about the potential results of resource based policy decisions.

Alternative funding schemes

With a few exceptions, states and local school districts share the financing of local schools. This sharing, however, varies in several important dimensions. State financing is frequently characterized in terms of a formula of identifying state funding per pupil in relationship to local capacity, local effort, and characteristics of the local student body. The details, the methods of describing the allocation mechanisms, and even the nomenclature differ across states. Nonetheless, three basic funding mechanisms characterize the options – and most states use a combination of them. Categorical aid funds districts based upon specific identified needs; foundation aid compensates for different tax capacity of the local district; and variable matching aid adjusts state support for both different tax capacity and the taxing decisions of the local district.

To understand the funds available to a district, it is easiest to ignore federal subsidies and to concentrate on the state and local funding. For simplicity the funding formulae are written in terms of a property tax base, reflecting the fact that tax on real property is the most frequently employed local revenue raising instrument. With number of students N_j, tax base per student B_j , and local tax rate r, school district j can raise L_j in local funds:

$$(7) \qquad L_j = N_j \times rB_j$$

Districts with larger tax bases clearly can raise funds more easily if they choose to increase revenues. Note, however, that the tax base is actually endogenous. Because the tax base includes not only residential property but also commercial and industrial property, local residents in districts with a smaller

proportion of the tax base from residential property pay less of the total local tax bill. In other words, the "tax price" of raising an extra dollar of local revenue is less than one dollar for the residents (Ladd (1975)). Other things being equal, then, districts with a low tax price are more attractive, and demand for housing in them will be high – leading to housing prices being bid up and to the tax base being larger (Oates (1969)). This endogeneity obviously also enters into policy actions. For example, if the state prohibited local districts from taxing certain kinds of commercial property, some districts would lose two ways: B, the tax base, would fall with the loss of the commercial property from its tax base and the district would likely become relatively less attractive so that housing values fell, bringing B down further.

State governments routinely act to ameliorate the largest disparities in revenue abilities of districts and to even out cost differences across districts. The simplest kinds of state subsidy formulae provide fixed payments for categorical purposes such as:

(8) $C_j = c \times N_j$

Where the subsidy from the central authority (state) to district j is a flat amount, c, times the number of eligible students (N). Thus, for example, the state might provide a flat amount for each student classified as gifted or talented. Categorical aid is identified with specific purposes and often requires specific expenditures.[94] Categorical aid flows for a variety of purposes including student transportation, books and supplies, special education, and bilingual education programs. Districts, however, cannot be assumed to be impervious of incentives. As described below, there is evidence that the number of "categorical" students is endogenous, related to the fiscal incentives that are presented.[95]

---

[94] Of course, as is generally true, funds are fungible – implying that requirements to spend categorical funds would not be binding unless the local education authority wished to spend less than c per eligiblestudent. Some evidence, however, suggests that this is not the case. Evidence of the "flypaper effect" suggests that funds tend to stick where they are provided (e.g., see Gramlich (1977)and the discussion in Oates (1994)).

[95] Another element of categorical programs is identifying and adjusting for program costs. For example,

Most attention, however, has focused on alternative ways of providing general revenue – revenue with few specific use restrictions. By far the most common aid formula is foundation aid where the state subsidy varies inversely with the local property tax base per student such as:

$$(9) \qquad C_j = N_j \times (F - r_o B_j)$$

F is the "foundation," or the amount the district would get per student if it had no ability to tax itself, r_o is the fixed tax rate used to calculate what the local district might be expected to raise itself. Because local revenue capacity rises with the taxable base (equation 7), states typically reduce their subsidy for districts with higher B.

Several aspects of foundation aid are important, although to varying degrees across different states. First, the foundation aid formula can clearly lead to negative subsidies for districts with large tax bases, i.e., when $r_o B_j > F$. States have generally had mixed reactions to whether or not they should recapture any excess funds that a district may be able to raise. Without recapture, districts with large amounts of property wealth continue to have more revenue raising opportunities than those with lower tax bases.

Second, if a district chooses a tax rate, r_j, that is greater than r_o, state subsidies are unaffected and jurisdictions with larger tax bases can capture the full benefit of their tax base.[96] Total education spending (E_j) under a foundation plan combines local and state funding as in:

---

special education funding is frequently adjusted for different handicapping conditions of students. Existing estimates of costs that are built into many funding formula begin with the expenditure choices of schools, making them somewhat problematic. See Chaikind, Danielson, and Brauen (1993)and Lankford and Wyckoff (1996). Other attempts to deal with local cost differences involve both choices of what costs are important and estimation of the their magnitude – but this again is difficult when based largely on the choices of districts which also involve varying preferences about teacher quality (Ballou and Podgursky (1997)).

[96] Many states require that the district levy some minimum tax rate, often r_o, as a way of ensuring that districts bear a minimum portion of the total taxes for schools. Some states also cap the tax rate that local jurisdictions can apply (r_j) in an effort to limit variations in local spending.

(10) $E_j = L_j + C_j$

$= N_j \times [(F) + (r_j - r_o)B_j]$

The incentive to raise the local tax rate of course depends on the desired spending of the community. Thus, as the foundation level – that "guaranteed" by the state formula – increases, local communities have less incentive to spend more than F per student and tend to choose tax rates closer to the nominal rate of the state, r_o.

Third, the number of students is frequently adjusted to take into account special factors. For example, special education students, instead of being paid simple categorical funds as in (eq. 8), may be credited as if they are additional effective students. Therefore, for example, a disabled student who has an identified learning disability may count as, say, 1.6 students. In such a case, the state subsidy for categorical aid depends upon the local tax capacity. But, of course, this also raises the possibility that N_j is really endogenous to the extent that districts can vary populations. For example, Cullen (forthcoming) shows that school districts adjust the numbers identified as requiring special education according to their fiscal incentive;[97] Cullen and Rivkin (forthcoming) also show that further reactions to the provision of special education services by family choices of residence are possible.

Finally, a variable matching grant by states takes into account the actual tax choices of the district. While this formulation comes under many different names (and equivalent ways of writing the formula), perhaps the simplest is to think of this as a "guaranteed tax base" as in:

(11) $C_j = N_j \times r_j (B_o - B_j)$

where B_o is the tax base that is guaranteed to the district. (Other common descriptions include wealth

[97] An important feature of the analysis is separating out the causal impact. In this, Cullen develops a series of plausible instruments that permit identification of the impact on classification.

neutralizing, power equalizing, and percentage equalizing). For any district with an actual tax base less than the guaranteed base, the state supplements the local tax revenue. Again, a key element is what happens when $B_j > B_o$. If recapture is required, the local district sends the excess revenues to the state. In such a case, the effective revenue schedule for all districts (combining equations 7 and 11) is simply:

$$(12) \qquad E_j = N_j \times r_j B_o$$

and districts differ in spending just to the extent that they choose varying local tax rates. A variant of this formulation was first popularized by Coons, Clune, and Sugarman (1970) and is the intellectual backdrop to the original school finance law suits from *Serrano v. Priest* through recent suits. Again, however, recapture provisions have generally proved unpopular and are generally only imposed in response to court orders.

An important aspect of the fiscal consideration is that the formulae provide an indication of the *opportunities* that face a district. They do not necessarily foretell the distribution of spending outcomes that will occur, because these require knowing more about the behavioral responses of individuals. The behavioral responses occur in two primary ways: the choice of districts by individuals and the expression of preferences within each district.

Fiscal effects of alternative plans

While courts and legislatures have significantly changed the fiscal responsibilities for schools since the early 1970s, surprisingly little is known about the effects of these alterations. Some efforts have been made to track the spending patterns of these changes, but less attention has gone to consideration of the overall level.

Wyckoff (1992) traces the within-state variations in spending over a substantial period of change. By calculating alternative measures of spending variations, he concludes that state court action has tended to

lead to some narrowing in spending variations, but this is not uniform.[98] In a comprehensive study of existing court and legislative actions, Murray, Evans, and Schwab (1998) estimate models of spending variation which incorporate information about the timing of court actions. They also conclude that court action has led to significant narrowing of the spending distribution. However, the state court actions directly consider just variations within states, while they find that the most significant narrowing of the spending distribution has been between states.

The exact nature of the changes in the distribution has been the focus of a number of different analyses. In an early analysis, Stern (1973) focuses on how changes in the financing formula would affect equalization of funding within Massachusetts. His simulations indicate that a variable matching subsidy would reduce disparities due to fiscal capacity differences of states but not those due to socioeconomic differences. The key element, understanding the underlying preferences of individuals, was extended by Feldstein (1975). He specifically considered how wealth neutrality approaches might lead to unexpected spending outcomes where spending was inversely related to local wealth. Craig and Inman (1982, 1986) consider spending patterns in light of reforms that involve moving spending responsibilities to lower levels of government. Their work builds upon an explicit model of decision making at different levels of government and highlights the importance of the fiscal federal system. They find surprisingly strong reactions to changes in funding patterns.[99]

Court rules may have effects on the level of spending in addition to any effects on the distribution. Most court rulings do not specify how the state should distribute funds, only that an existing funding system does not meet constitutional requirements. This typically requires that the state legislature alter its approach to funding. A variety of political economy models could describe the resultant changes. The simplest

---

[98] An early predecessor to the systematic consideration of state funding variations along with discussion of alternative measures of the distribution of spending is Berne and Stiefel (1984).

[99] While not emphasized here, they also demonstrate the importance of institutional structure including how grants often have surprising effects. For example, because of the "flypaper effect," tax reductions exhibit different effects on spending than unconditional grants (cf. Oates (1994)).

model recognizes that pure distributional changes dictate a large number of losing districts that find their state funding being decreased. To offset these changes, political incentives push toward increasing the overall level of spending at the same time that there is a redistribution of spending. Downes and Shah (1996) provide some evidence for this effect, although it is relatively small.

One natural interpretation of the results is that there are simply too many different circumstances and court rulings to imply uniform effects on spending. In fact, Hoxby (2001) pursues exactly the characterization of alternative plans. By separating the key parameters and assessing how these differ by states, she is able to estimate the effects of legislative and court actions on overall spending levels. Following the earlier analyses of behavioral responses to altered incentives, she shows convincingly that the naïve, no-behavioral-response analysis that accompanies many plan changes is very misleading.

Effect of equalization on equity

If resource availability is not a good index of educational outcomes or if providing for overall resource levels does not ensure a desired level of performance, the courts face the same dilemma as legislatures. Simply providing more funding or a different distribution of funding is unlikely to improve student achievement (even though it does affect the distribution of tax burdens for school financing across the citizens of a state).

A related issue—one highlighted in some recent school finance court cases—centers on whether funding for schools is "adequate" (cf. Clune (1994)) The idea of adequacy has been to provide sufficient funds to ensure some chosen level of achievement for all students in a state.[100] Such concepts may have popular appeal, but they have no policy superiority to traditional district equity arguments when translated into resource requirements. First, what is adequate is a purely political and economic issue that it likely to

[100] An important aspect of the legal and legislative discussions of adequacy is translation into expenditure. The situation might be very different if equity and adequacy were truly defined in terms of student outcomes, but in such a situation there would be little way of ensuring any given distribution of performance – making judicial or legislative mandates generally impossible.

change both with the demands of the economy and with political views on appropriate levels of government support of programs. Second, and more important, the previous conclusion that resources are not a good index of student performance holds no matter what goals are placed for student achievement or how these goals are arrived at. Thus, there is no objective method of indicating what resources are required for an 'adequate' level of student performance.[101]

A variation of this general adequacy theme is to argue that, while resources alone may not be sufficient to guarantee achievement, adequate resources are surely necessary. Undoubtedly, this is an accurate statement at some level, because a school with no funds would not be expected to add anything to student achievement. Nonetheless, as shown in Table 5 for the United States, real spending per student rose dramatically between 1900 and 1995, even though U.S. student performance appears to have been essentially unchanged. Further, nothing in the previous analytical results about the effects of resources suggests that there is a level below which resources have clear and powerful effects on achievement—which would be a demonstration that some schools are below the threshold of 'necessity.' Just asserting that there is some level of necessary expenditure does not make the case for pure resource policies in today's schooling environment.

Analyses considering spending variations do not, for the reasons previously described, give direct information about any effects on student performance. Perhaps surprisingly given the extent of legal actions during the 1970s, 1980s, and 1990s, there has been little direct analysis of performance differences related to funding equalization. Downes (1992) investigates whether the early alterations in California arising from the *Serrano v. Priest* case had any impact on variations in student test performance. He finds no significant impact on performance when he looked at variations before and after equalization of local spending. Hanushek and Somers (2001) pursue a different tact. They consider whether variations in spending within a state are ultimately reflected in earnings variations when the students progress into the labor market. This

[101] See, for example, the difficulty in defining adequacy in Ladd and Hansen (1999).

analysis, a generalization of that in Card and Krueger (1992a), finds little evidence that earnings variations correspond to earlier spending variations.[102] These findings are consistent with the earlier evidence that suggested no consistent improvement in student performance from changes in funding.

One of the impacts of court cases has been increased centralization of funding. Particularly during the 1970s, the average share of funding coming from the state level rose and surpassed the share at the local level. Centralization of funding itself could have implications for both funding levels and performance. As funding responsibility moves from the locality to higher levels of government, individuals are more detached from the taxing and spending decisions and standard free rider problems increase. The arguments for this position are largely conceptual (e.g., Hoxby (1996a)). The direct empirical analysis typically must confront the concerns about aggregation bias raised previously when they rely on evidence across states (e.g., Walberg and Walberg III (1994); Peltzman (1993)). Indirect evidence on reactions to changes over time, however, provides additional support. Specifically, movements toward more centralization of state finance have been related to efforts to control overall levels of spending by Fischel (1989), Silva and Sonstelie (1995), and Courant and Loeb (1997). The argument is that, in reaction to state court decisions to equalize spending, California voters supported Proposition 13 that placed strict limits on levels of taxes and spending.[103] A similar analysis by Loeb (2001), simulating the effects different financing systems, shows that the amount of local control of supplementation can have dramatic impacts on the level and distribution of spending across districts and that these effects are very dependent on the distribution of local preferences.[104]

---

[102] The Hanushek and Somers analysis concentrates on the within state variance in funding and the subsequent variations in earnings of people educated in a given state. Thus, it avoids the largest problems associated with differences in state policy environments, as discussed previously.

[103] Related work on the causes and consequences of tax limitation initiatives have also had a component of looking at student performance in addition to the more common study of the effects on spending and its composition. See Figlio (1997) and the discussion there.

[104] An innovation in this work is relaxing a common assumption that there is perfect sorting by income, which allows her to base simulations on the observed wealth distributions of school districts.

Mobility, Distribution, and Equilibrium

One very significant problem in considering state financing of schools is that the analysis generally assumes that districts and their residents are fixed. Additionally, it is often assumed that property wealth – the common basis for local taxation – is an acceptable index of the economic well-being of a jurisdiction's population. These assumptions prove to be very problematic but difficult to deal with in any complete fashion (Hanushek (1991)).

If the property tax base just incorporated residential property, the tax base would be a reasonable measure of the economic well-being of the population, but the inclusion of commercial and industrial property in the tax base breaks this relationship. Thus, for example, some cities have large proportions of their students in poverty even though they have large tax bases (e.g., New York City) while other cities have large poverty populations and small tax bases (e.g., Newark, New Jersey). Most of the funding formulae emphasize the local ability of districts to raise funds, but often incorrectly suggest that this is synonymous with poverty. Moreover, the tax advantage of having large amounts of commercial and industrial property in a city will, to the extent that potential tax revenues exceed the costs of providing public services to these properties, be capitalized into the value of residential properties. Thus, two identical properties located in jurisdictions with different tax capacities will sell for different amounts.

Additionally, as discussed previously, communities that have good schools will generally be more desirable than schools with bad schools, and housing prices will be bid up in places with good schools (Kain and Quigley (1975), Rosen and Fullerton (1977), Black (1999), Weimer and Wolkoff (2001)).[105] In other words, residents in school districts with larger capitalization of school quality are actually paying more for their schools than is apparent by a simple comparison of tax rates to support schools. This variation in capitalization also has apparent feedbacks to the quality of schools, since schools in more competitive areas tend to have more effective schools (Hoxby (2000a), Hanushek and Rivkin (forthcoming)). The overall

[105] The capitalization relationships are complicated by variations in the efficiency of schools in producing

endogeneity of the tax base is ignored in most considerations of school finance policy – both legislatively and judicially.

Behind the capitalization of aspects of tax base and schools into housing prices is the choice of communities and mobility of residents. The high level of mobility in school age population underscores the levels of potential adjustment to school cost and quality that occurs regularly.[106] If thought of in terms of the different possible margins of adjustment, as in the exit and voice of Hirschman (1970), it also highlights the equilibrium nature of school finance policy and changes in such policy. Thus, for example, a change in school funding by a state would set in motion a pattern of changed housing values, altered residential and school choices, and adjusted spending and performance patterns of schools. The resulting equilibrium outcomes in terms of the distribution of expenditure and performance patterns across the population are not easy to project.

Several research programs have addressed various aspects of this general equilibrium problem and have provide insights into the effects of different fiscal policies. An early and interesting line of work, although following a different path than here, involves understanding the effects of financing mechanisms on economic growth (Glomm and Ravikumar (1992, 1998)). This work concentrates on how public and private schooling interact with economic growth. Their analysis of endogenous growth models suggests a trade-off between higher per capita incomes with private schools and more equality with public schools. By explicitly considers voting and public choice mechanisms, they also suggest that voters are likely to choose public schools.

Fernandez and Rogerson (1996, 1997, 1998)develop calibrated models of schooling and labor market outcomes that incorporate different jurisdictions. This apparatus allows investigation of different

---

any quality of outcomes; see Somers (1998).

[106] Hanushek, Kain, and Rivkin (2001a) describe school mobility and its effects on achievement. Interestingly, the detailed consideration of moving causes in the Current Population Survey does not attempt to investigate whether moving behavior is related to public services or schools; see Schacter (2001a, 2001b). For within county movers, 26 percent move for family-related reasons, 6 percent for work-related reasons, and 65 percent for housing-related reasons.

financing schemes, and the complete characterization of individuals in the economy permits welfare comparisons derived from different financing schemes. While the exact welfare comparisons appear dependent upon the key parameters of their economy, they develop a structure that identifies how individual behavior responds to changing in financing schemes. A key insight from their work is that it is difficult to predict the impact of different financing schemes on the level of educational spending from a partial equilibrium setting. In their models, however, the schooling sector itself is very simple (efficient provision of public education depending just on spending), making it impossible to trace out implications of policy for the schooling sector.

Epple and Romano (1996, 1998, forthcoming) develop more detailed models of household location that interact consumer choices and schooling outcomes. A central element of their analyses is variations in school quality that depend on peers in the school, leading to an externality from high ability students. Households of different incomes with children of different abilities seek out communities that maximize utility. The presence of a private sector supported by vouchers permits the private schools to internalize the externality of high ability students, but the exact solution depends on the size of the voucher. Epple and Romano (1998) sets out the basic properties of equilibrium including the key predictions that private schools will trade off income and ability, private tuitions declining in ability, and a strict hierarchy of schools. Epple and Romano (forthcoming) turn to the alternative of public school choice and traces out equilibrium models of neighborhood choice, showing that residential stratification is a likely result that interacts with tax and expenditure policies.

In a series of papers, Nechyba (1999, 2000, 2001, forthcoming-a, forthcoming-b) develops rich models of the interplay of residential location, school policies, and segregation. Much of the discussion of private schools, for example, suggests that segregation of schools is a likely result, but these discussions ignore the current incentives to segregate in communities. Nechyba (2000) shows that vouchers work in part to severe the link between residential location and schools. As a result, the pressures to segregate in

residential communities are lessened, and vouchers produce less segregation than a pure public school system. Nechyba (forthcoming-b) introduces private schools in the absence of vouchers, mirroring the observed U.S. equilibrium, and – even in the absence of vouchers – a private school alternative is shown to reduce segregation more than a completely centralized (and equalized) system. Nechyba (forthcoming-a) further explores the interaction of public school efficiency, private schools, and schooling choice, again showing that the general equilibrium responses differ markedly from simple partial equilibrium predictions.

Each of these approaches necessarily concentrates on some specific issues, while highly simplifying others. At the same time, each demonstrates vividly that the general equilibrium can differ quite dramatically from the partial equilibrium results. As more experience is gained in these models, they can enrich their descriptions of behavior and can provide important insights about central issues in the financing of schools.

Some Underexplored Topics

While much has been learned about publicly provided education, a variety of important issues remain underresearched. Most of these are apparent from replaying the prior discussions of what is currently known. Therefore, this discussion is designed to highlight a few areas rather than to develop the areas in any depth.

Perhaps at the top of the list is the behavior of consumers in the face of public provision of schools. If schools are inefficient in their use of public funds, why is consumer and voter behavior so ineffective? Attention to consumer behavior and to underlying issues of political economy has focused almost exclusively on issues of the level of spending, ignoring questions of the results of spending.[107] Understanding the elements of citizen behavior is important for both positive theories of spending and for consideration of the feasibility of alternative policy regimes.

---

[107] Exceptions include Somers (1998) and Pritchett and Filmer (1999). Somers considers how school efficiency and consumer demands interact in determining spending levels and patterns of districts. Pritchett and Filmer consider how self-interest of school decision makers affects patterns of inefficiency.

The incentive structure of schools is the keystone to educational policy, but little headway has been made in understanding this more generally. Many reform proposals – from expanded consumer choice to heightened accountability systems – involve direct attempts to change incentives. Nonetheless, existing research, which has generally relied upon the few natural experiments with different incentives, remains very rudimentary. The empirical analyses typically relate to very specific programs whose generalizability is not well understood. The conceptual work is even farther behind, failing in general to develop much of a vocabulary or taxonomy for incentive schemes and thus leaving open a wide range of important measurement issues. In a larger political view, current participants clearly have preferences over a variety of the potential incentive schemes and exert influence on both their adoption and their evolution – and understanding the political aspects of program development is important. Pressures for changes in incentives currently being contemplated or implemented, however, open large new areas for research.

As discussed throughout, the analysis has been heavily weighted toward studies of the United States experience. This slant to research is particularly unfortunate, because international schooling offers a very wide range of experiences. The variations in organization, funding, and institutions that are found internationally offer promise for resolving some of the questions that have not been fully addressed within the more limited U.S. experience.

The analysis of schooling policies generally pays little attention to the array of policies and influences outside of school: daycare, preschool, after school programs, extracurricular activities, and the like. These facets of children's development are typically treated independent of schools in both analytical and policy discussions, even though they undoubtedly interact with schools and should be considered in a larger optimizing framework for governmental programs.

Table 1. Percentage of Population Attaining Upper Secondary Education or More, by country: 1999

	Ages 25-64	Ages 25-34	Ages 35-44	Ages 45-54	Ages 55-64
OECD countries					
Australia	57	65	59	55	44
Austria[a]	74	83	78	69	59
Belgium	57	73	61	50	36
Canada	79	87	83	78	62
Czech Republic	86	93	89	85	75
Denmark	80	87	80	79	70
Finland	72	86	82	67	46
France[a]	62	76	65	57	42
Germany	81	85	85	81	73
Greece	50	71	58	42	24
Hungary	67	80	76	70	36
Iceland	56	64	59	53	40
Ireland[a]	51	67	56	41	31
Italy	42	55	50	37	21
Japan	81	93	92	79	60
Korea	66	93	72	47	28
Luxembourg	56	61	57	52	41
Mexico	20	25	22	16	9
Netherlands	m	m	m	m	m
New Zealand	74	79	77	71	60
Norway[a]	85	94	89	79	68
Poland[a]	54	62	59	53	37
Portugal	21	30	21	15	11
Spain	35	55	41	25	13
Sweden	77	87	81	74	61
Switzerland	82	89	84	79	72
Turkey	22	26	23	18	12
United Kingdom[a]	62	66	63	60	53
United States	87	88	88	88	81
OECD mean	**62**	**72**	**66**	**58**	**45**
World Education Indicators participants					
Brazil[a]	24	29	27	21	12
Chile[a]	43	55	45	35	24
Indonesia	22	33	21	15	9
Jordan	51	55	55	43	25

Malaysia[a]	35	50	35	20	10
Peru[a]	46	58	48	35	24
Philippines	44	55	45	34	24
Sri Lanka[a]	36	46	36	31	21
Thailand[a]	16	23	17	9	6
Tunisia	8	11	9	6	3
Uruguay[a]	32	39	34	28	20
Zimbabwe	29	51	19	11	7

Note: a. Year of reference is 1998.

Source: Organisation for Economic Co-operation and Development (2001)

Table 2. Relative earnings of the population with income from employment
By level of educational attainment for the population 25 to 64 years of age
(upper secondary education =100)

	year	Male			Female		
		Lower secondary and below	Higher education (nonuniversity)	Higher education (university)	Lower secondary and below	Higher education (nonuniversity)	Higher education (university)
Australia	1997	87	120	144	85	113	154
Canada	1997	84	109	148	76	116	164
Czech Republic	1999	75	177	178	72	127	172
Denmark	1998	87	122	148	89	118	144
Finland	1997	94	128	186	100	122	176
France	1999	88	128	178	79	131	158
Germany	1998	77	105	149	85	104	160
Hungary	1999	72	240	218	67	138	159
Ireland*	1997	72	100	149	57	129	171
Italy	1998	54	n.a.	n.a.	61	n.a.	n.a.
Korea	1998	88	105	143	69	118	160
Netherlands	1997	86	142	138	71	128	145
New Zealand	1999	76	n.a.	n.a.	74	n.a.	n.a.
Norway	1998	85	125	133	84	142	136
Portugal	1998	61	149	188	62	131	190
Spain	1996	75	96	178	68	82	155
Sweden	1998	87	n.a.	n.a.	89	n.a.	n.a.
Switzerland	1999	81	122	144	73	131	154
United Kingdom	1999	73	126	159	68	139	193
United States	1999	65	119	183	63	120	170
Country mean		**78**	**130**	**163**	**75**	**123**	**162**

n.a. – not available.

Source: Organisation for Economic Co-operation and Development (2001)

Table 3. Expenditure per student and Spending Relative to GDP (1998)

	Expenditure[a]		Expenditure relative to GDP[b]
	Primary	Secondary	
	(U.S. dollars)		(percentage)
OECD countries			
Australia	3981	5830	3.8
Austria[c]	6065	8163	4.2
Belgium[d]	3743	5970	3.5
Belgium (Fl.) [d]	3799	6238	3.6
Canada	n.a.	n.a.	4.1
Czech Republic	1645	3182	3.1
Denmark	6713	7200	4.3
Finland	4641	5111	3.7
France	3752	6605	4.4
Germany	3531	6209	3.7
Greece[d]	2368	3287	3.5
Hungary	2028	2140	3.1
Iceland[c]	n.a.	n.a.	n.a.
Ireland	2745	3934	3.3
Italy[c]	5653	6458	3.5
Japan	5075	5890	3.0
Korea	2838	3544	4.0
Luxembourg	n.a.	n.a.	n.a.
Mexico	863	1586	3.5
Netherlands	3795	5304	3.1
New Zealand	n.a.	n.a.	n.a.
Norway[c]	5761	7343	4.4
Poland	1496	1438	n.a.
Portugal	3121	4636	4.2
Spain	3267	4274	3.7
Sweden	5579	5648	4.5
Switzerland[c]	6470	9348	4.5
Turkey[c]	n.a.	n.a.	2.3
United Kingdom[d]	3329	5230	n.a.
United States	6043	7764	3.7
Country mean	**3940**	**5294**	**3.7**
OECD total	**3915**	**5625**	**3.6**
World Education Indicators participants			
Argentina[c]	1389	1860	3.1
Brazil[c,d]	837	1076	n.a.
Chile	1500	1713	3.9
Indonesia[c]	116	497	1.4
Israel	4135	5115	5.5

Malaysia[c]	919	1469	n.a.
Paraguay[c]	572	948	n.a.
Peru	479	671	3.3
Philippines[c]	689	726	4.9
Thailand	1048	1177	3.8
Tunisia[c]	891	1633	n.a.
Uruguay[c]	971	1246	n.a.
Zimbabwe	768	1179	n.a.

n.a. – not available

Notes:

a. Expenditure per student in US dollars converted using PPPs on public and private institutions, by level of education, based on full-time equivalents
b. Direct and indirect expenditure on educational institutions from public and private sources, by level of education, source of fund and year
c. Expenditure amounts for public institutions only.
d. Expenditure amounts for public and government-dependent private institutions only.
e. Year of reference for expenditure amounts is1997
f. Year of reference for expenditure amounts is1999.

Source: Organisation for Economic Co-operation and Development (2001)

Table 4. Country Ranking of Performance on Third International Mathematics and Science Study (TIMSS), 1995

8th Grade Performance		12th Grade Performance	
Mathematics	Science	Mathematics	Science
Singapore	**Singapore**	**Netherlands**	**Sweden**
Korea	**Czech Republic**	**Sweden**	**Netherlands**
Japan	**Japan**	**Denmark**	**Iceland**
Hong Kong	**Korea**	**Switzerland**	**Norway**
Belgium-Flemish	**Bulgaria**	**Iceland**	**Canada**
Czech Republic	**Netherlands**	**Norway**	**New Zealand**
Slovak Republic	**Slovenia**	**France**	**Australia**
Switzerland	**Austria**	**New Zealand**	**Switzerland**
Netherlands	**Hungary**	**Australia**	**Austria**
Slovenia	England	**Canada**	**Slovenia**
Bulgaria	Belgium-Flemish	**Austria**	**Denmark**
Austria	Australia	**Slovenia**	Germany
France	Slovak Republic	**Germany**	France
Hungary	Russian Federation	**Hungary**	Czech Republic
Russian Federation	Ireland	Italy	Russian Federation
Australia	Sweden	Russian Federation	UNITED STATES
Ireland	UNITED STATES	Lithuania	Italy
Canada	Germany	Czech Republic	Hungary
Belgium-French	Canada	UNITED STATES	Lithuania
Sweden	Norway	*Cyprus*	*Cyprus*
Thailand	New Zealand	*South Africa*	*South Africa*
Israel	Thailand		
Germany	Israel		
New Zealand	Hong Kong		
England	Switzerland		
Norway	Scotland		
Denmark	*Spain*		
UNITED STATES	*France*		
Scotland	*Greece*		
Latvia	*Iceland*		
Spain	*Romania*		
Iceland	*Latvia*		
Greece	*Portugal*		
Romania	*Denmark*		
Lithuania	*Lithuania*		
Cyprus	*Belgium-French*		
Portugal	*Iran*		
Iran	*Cyprus*		
Kuwait	*Kuwait*		
Columbia	*Columbia*		
South Africa	*South Africa*		

Note: **bold** – significantly about United States; *italics* – significantly below United States.
Source: U.S. Department of Education (1999)

Table 5. Public School Resources in the United States, 1960-1995

Resource	1960	1970	1980	1990	1995
Pupil-teacher ratio	25.8	22.3	18.7	17.2	17.3
% teachers with master's degree or more	23.5	27.5	49.6	53.1	56.2
median years teacher experience	11	8	12	15	15
current expenditure/ADA (1996-97 $'s)	$2,122	$3,645	$4,589	$6,239	$6,434

Source: U.S. Department of Education (1997)

Table 6. Percentage Distribution of Estimated Effect of Key Resources on Student Performance, Based on 377 Studies

Resources	number of estimates	Statistically significant		Statistically insignificant
		Positive	Negative	
Real classroom resources				
Teacher-pupil ratio	276	14%	14%	72%
Teacher education	170	9	5	86
Teacher experience	206	29	5	66
Financial aggregates				
Teacher salary	118	20	7	25%
Expenditure per pupil	163	27	7	34
Other				
Facilities	91	9	5	86
Administration	75	12	5	83
Teacher test scores	41	37	10	53

Source: Hanushek (1997a), revised.

Table 7. Percentage Distribution of Estimated Effect of Teacher-Pupil Ratio and Expenditure per Pupil by State Sampling Scheme and Aggregation

Level of aggregation of recources	number of estimates	Statistically significant		Statistically insignificant
		Positive	Negative	
A. Teacher-Pupil Ratio				
Total	276	14%	14%	72%
Single state samples[a]	157	11	18	71
Multiple state samples[b]	119	18	8	74
Disaggregated within states[c]	109	14	8	78
State level aggregation[d]	10	60	0	40
B. Expenditure per pupil				
Total	163	27%	7%	66%
Single state samples[a]	89	20	11	69
Multiple state samples[b]	74	35	1	64
Disaggregated within states[c]	46	17	0	83
State level aggregation[d]	28	64	4	32

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.

Source: Hanushek (1997a), revised.

Table 8. Percentage Distribution of Estimated Influences on Student Performance, Based on Value-added Models of Individual Student Performance

Resources	number of estimates	Statistically significant		Statistically insignificant
		Positive	Negative	
A. All studies				
Teacher-pupil ratio	78	12%	8%	80%
Teacher education	40	0	10	90
Teacher experience	61	36	2	62
b. Studies within a single state				
Teacher-pupil ratio	23	4%	13%	83%
Teacher education	33	0	9	91
Teacher experience	36	39	3	58

Source: Hanushek (1997a), revised.

Table 9. Percentage Distribution of Estimated Expenditure Parameter Coefficients from 96 Studies of Educational Production Functions: Developing Countries

Input	Number of estimates	Statistically Significant		Statistically Insignificant
		Positive	Negative	
Teacher/Pupil Ratio	30	27%	27%	46%
Teacher Education	63	56	3	41
Teacher Experience	46	35	4	61
Teacher Salary	13	31	15	54
Expenditure/Pupil	12	50	0	50
Facilities	34	65	9	26

Source: Hanushek (1995)

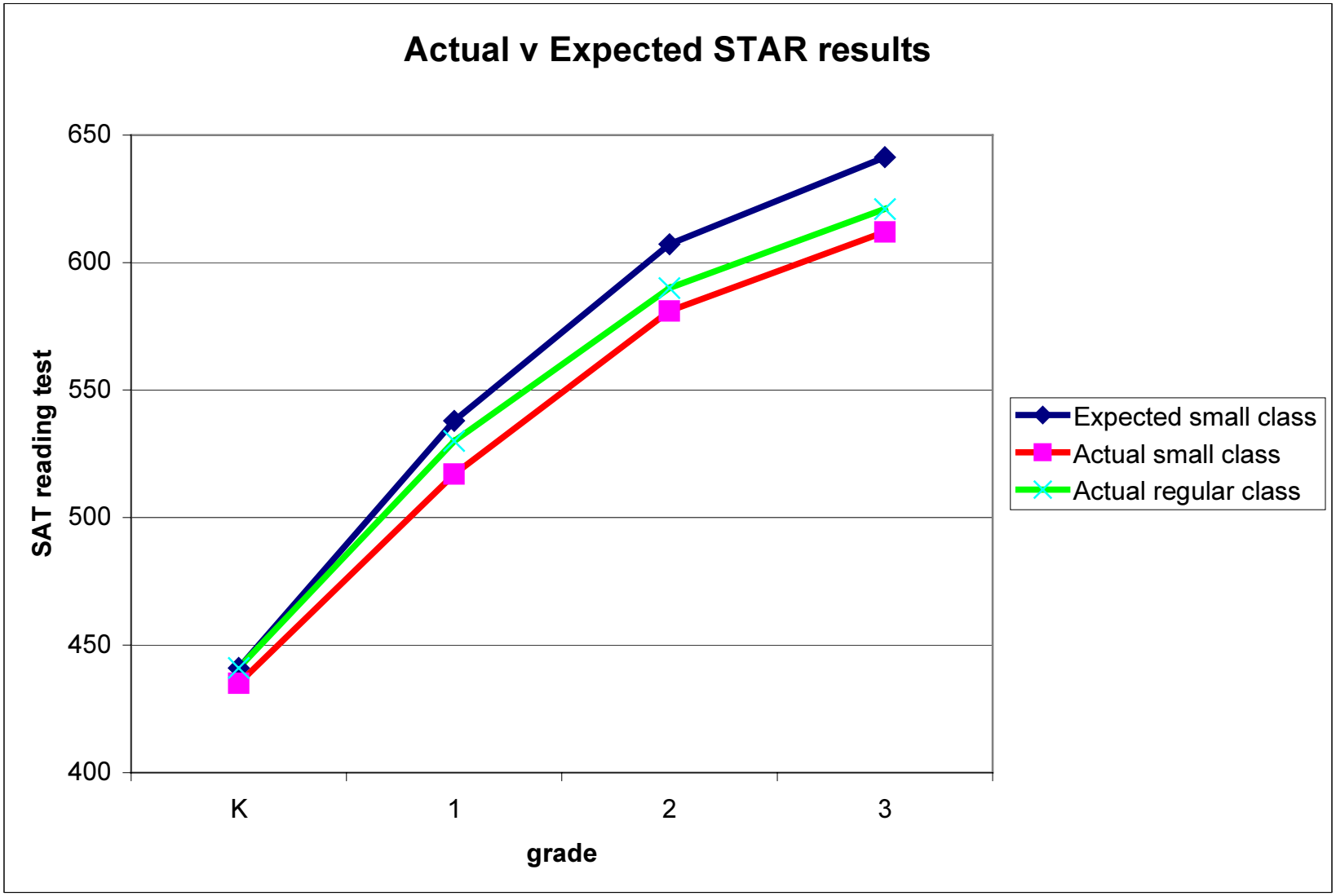

Figure 1. Actual and Expected Student Performance in Reading in Project STAR, by Grade

References

Acemoglu, D., and J. Angrist. 2000. "How large are the social returns to education? Evidence from compulsory schooling laws." In *NBER Macroeconomics Annual 2000*, edited by B.S. Bernanke and K. Rogoff. Cambridge, MA: MIT Press.

Aigner, D.J., and S.F. Chu. 1968. "On estimating the industry production function." *American Economic Review* 58,no.4 (September):826-839.

Aigner, D.J., A.K. Lovell, and P. Schmidt. 1977. "Formulation and estimation of stochastic frontier production function models." *Journal of Econometrics* 6:21-38.

Akerhielm, K. 1995. "Does class size matter?" *Economics of Education Review* 14,no.3 (September):229-241.

Akin, J.S., and I. Garfinkel. 1977. "School expenditures and the economic returns to schooling." *Journal of Human Resources* 12,no.4 (Fall):460-481.

Angrist, J.D., E. Bettinger, E. Bloom, E.N. King, and M. Kremer. 2001. "Vouchers for private schooling in Columbia: Evidence from a randomized natural experiment." Working Paper 8343, Cambridge, MA: National Bureau of Economic Research (June).

Angrist, J.D., and V. Lavy. 1999. "Using Maimondides' rule to estimate the effect of class size on scholastic achievement." *Quarterly Journal of Economics* 114,no.2 (May):533-575.

Antos, J.R., and S. Rosen. 1975. "Discrimination in the market for teachers." *Journal of Econometrics* 2(May):123-150.

Argys, L.M., D.I. Rees, and D.J. Brewer. 1996. "Detracking America's Schools: Equity at zero cost?" *Journal of Policy Analysis and Management* 15,no.4 (Fall):623-645.

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

Armor, D.J., P. Conry-Oseguera, M. Cox, N. King, L. McDonnell, A. Pascal, E. Pauly, and G. Zellman. 1976. *Analysis of the school preferred reading program in selected Los Angeles minority schools*. Santa Monica, CA: Rand Corp.

Baker, G.P. 1992. "Incentive Contracts and Performance Measurement." *Journal of Political Economy* 100,no.3 (June):598-614.

———. forthcoming. "Distortion and risk in optimal incentive contracts." *Journal of Human Resources*.

Ballou, D. 1996. "Do public schools hire the best applicants?" *Quarterly Journal of Economics* 111,no.1 (February):97-133.

———. 2001. "Pay for performance in public and private schools." *Economics of Education Review* 20,no.1 (February):51-61.

Ballou, D., and M. Podgursky. 1997. *Teacher pay and teacher quality*. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.

Barnett, W.S. 1992. "Benefits of compensatory preschool education." *Journal of Human Resources* 27,no.2 (Spring):279-312.

Barro, R.J. 1991. "Economic growth in a cross section of countries." *Quarterly Journal of Economics* 106,no.2 (May):407-443.

Barro, R.J., and J.-w. Lee. 2001. "International data on educational attainment: Updates and implications." *Oxford Economic Papers* 53,no.3 (July):541-563.

Barro, R.J., and X. Sala-I-Martin. 1995. *Economic Growth*. New York: McGraw-Hill, Inc.

Baumol, W.J. 1967. "Macroeconomics of unbalanced growth: The anatomy of urban crisis." *American Economic Review* 57,no.3 (June):415-426.

Becker, G.S. 1993. *Human capital: A theoretical and empirical analysis, with special reference to education*. 3rd ed. Chicago: University of Chicago Press. Original edition, 1964.

Becker, G.S., and H.G. Lewis. 1973. "On the interaction between the quantity and quality of children." *Journal of Political Economy* 81,no.pt. II (March/April):S279-S288.

Becker, G.S., and N. Tomes. 1976. "Child endowments and the quantity and quality of children." *Journal of Political Economy* 84,no.pt. 2 (August):S143-S162.

———. 1979. "An equilibrium theory of the distribution of income and intergenerational mobility." *Journal of Political Economy* 87,no.6 (December):1153-1189.

Bedard, K. 2001. "Human capital versus signaling models: University access and high school dropouts." *Journal of Political Economy* 109,no.4 (August):749-775.

Behrman, J.R., and N. Birdsall. 1983. "The quality of schooling: quantity alone is misleading." *American Economic Review* 73,no.5 (December):928-946.

Behrman, J.R., L.G. Kletzer, M.S. McPherson, and M.O. Schapiro. 1998. "The microeconomics of college choice, careers, and wages: Measuring the impact of higher education." *Annals of the American Academy of Political and Social cience* 559(September):12-23.

Behrman, J.R., and M.R. Rosenzweig. 2002. "Does increasing women's schooling raise the schooling of the next generation?" *American Economic Review*(March).

Behrman, J.R., M.R. Rosenzweig, and P. Taubman. 1996. "College choice and wages: Estimates using data on female twins." *Review of Economics and Statistics* 78,no.4 (November):672-685.

Behrman, J.R., P. Taubman, and R.A. Pollak. 1982. "Parental preferences and provision for progeny." *Journal of Political Economy* 90,no.1 (February):52-73.

Benabou, R. 1993. "Workings of a city: Location, education, and production." *Quarterly Journal of Economics* 434,no.3 (August):619-652.

———. 1996. "Heterogeneity, stratification, and growth: Macroeconomic implications of community structure and school finance." *American Economic Review* 86,no.3 (June):584-609.

Benhabib, J., and M.M. Spiegel. 1994. "The role of human capital in economic development: Evidence from aggregate cross-country data." *Journal of Monetary Economics* 34,no.2:143-174.

Bergstrom, T.C., and R.P. Goodman. 1973. "Private demands for public goods." *American Economic Review* 63,no.3 (June):280-296.

Bergstrom, T.C., D.L. Rubinfeld, and P. Shapiro. 1982. "Micro-based estimates of demand functions for local schools expenditure." *Econometrica* 50,no.5 (September):1183-1206.

Berne, R., and L. Stiefel. 1984. *The measurement of equity in school finance: Conceptual, methodological, and empirical dimensions*. Baltimore, MD: Johns Hopkins University Press.

Betts, J.R. 1995. "Does school quality matter? Evidence from the National Longitudinal Survey of Youth." *Review of Economics and Statistics* 77,no.2 (May):231-247.

———. 1996. "Is there a link between school inputs and earnings? Fresh scrutiny of an old literature." In *Does money matter? The effect of school resources on student achievement and adult success*, edited by G. Burtless. Washington, DC: Brookings.

Betts, J.R., and R.M. Costrell. 2001. "Incentives and equity under standards-based reform." In *Brookings papers on education policy: 2001*, edited by D. Ravitch. Washington, DC: Brookings Press.

Bils, M., and P.J. Klenow. 2000. "Does schooling cause growth?" *American Economic Review* 90,no.5 (December):1160-1183.

Bishop, J. 1989. "Is the test score decline responsible for the productivity growth decline?" *American Economic Review* 79,no.1:178-197.

———. 1991. "Achievement, test scores, and relative wages." In *Workers and their wages*, edited by M.H. Kosters. Washington, DC: The AEI Press.

———. 1996. "Signaling, incentives, and school organization in France, the Netherlands, Britain, and United States." In *Improving America's schools: The role of incentives*, edited by E.A. Hanushek and D.W. Jorgenson. Washington, DC: National Academy Press.

Bishop, J.H., F. Mane, M. Bishop, and J. Moriarty. 2001. "The role of end-of-course exams and minimal competency exams in standards-based reforms." In *Brookings papers in education policy 2001*, edited by D. Ravitch. Washington, DC: Brookings.

Black, S.E. 1999. "Do better schools matter? Parental valuation of elementary education." *Quarterly Journal of Economics* 114,no.2 (May):577-599.

Boissiere, M.X., J.B. Knight, and R.H. Sabot. 1985. "Earnings, schooling, ability, and cognitive skills." *American Economic Review* 75,no.5:1016-1030.

Boozer, M., and C. Rouse. 1995. "Intraschool variation in class size: Patterns and implications." NBER Working Paper 5144, June 1995

Boozer, M.A., A.B. Krueger, and S. Wolkon. 1992. "Race and school quality since Brown v. Board of Education." *Brooking Papers: Microeconomics*:269-338.

Borcherding, T.E., and R.T. Deacon. 1972. "The demand for the services of non-federal governments." *American Economic Review* 62,no.5 (December):891-901.

Borland, M.V., and R.M. Howsen. 1992. "Student academic achievement and the degree of market concentration in education." *Economics of Education Review* 11,no.1 (March):31-39.

Bound, J., and R.B. Freeman. 1992. "What went wrong? The erosion of relative earnings and employment among young black men in the 1980s." *Quarterly Journal of Economics* 107,no.1 (February):201-232.

Bowles, S., and H. Gintis. 1976. *Schooling in capitalist America: Educational reform and the contradictions of economic life*. New York: Basic Books.

Bowles, S., H. Gintis, and M. Osborne. 2001. "The determinants of earnings: A behavioral approach." *Journal of Economic Literature* 39,no.4 (December):1137-1176.

Bowles, S., and H.M. Levin. 1968. "The determinants of scholastic achievement--an appraisal of some recent evidence." *Journal of Human Resources* 3,no.1 (Winter):3-24.

Brickley, J.A., and J.L. Zimmerman. 2001. "Changing incentives in a multitask environment: Evidence from a top-tier business school." *Journal of Corporate Finance* 7:367-396.

Brock, W.A., and S.N. Durlauf. forthcoming. "Interactions-based models." In *Handbook of Econometrics*.

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

———. 1996b. "Introduction and summary." In *Does money matter? The effect of school resources on student achievement and adult success*, edited by G. Burtless. Washington, DC: Brookings.

Cain, G.G., and H.W. Watts. 1970. "Problems in making policy inferences from the Coleman Report." *American Sociological Review* 35,no.2 (April):328-352.

Cameron, S.V., and J.J. Heckman. 1999. "Can tuition policy combat rising wage inequality?" In *Financing college tuition: Government policies and educational priorities*, edited by M.H. Kosters. Washington, DC: The AEI Press.

Card, D. 1999. "Causal effect of education on earnings." In *Handbook of labor economics*, edited by O.C. Ashenfelter and D. Card. Amsterdam: North-Holland.

Card, D., and A.B. Krueger. 1992a. "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):1-40.

———. 1992b. "School quality and black-white relative earnings: A direct assessment." *Quarterly Journal of Economics* 107,no.1 (February):151-200.

———. 1996. "School resources and student outcomes: An overview of the literature and new evidence from North and South Carolina." *Journal of Economic Perspectives* 10,no.4 (Fall):31-50.

Carnoy, M. 2001. *School vouchers: Examining the evidence*. Washington, DC: Economic Policy Institute.

Case, A., and A. Deaton. 1999. "School inputs and educational outcomes in South Africa." *Quarterly Journal of Economics* 114,no.3 (August):1047-1084.

Caucutt, E.M. forthcoming. "Educational policy when there are peer group effects -- size matters." *International Economic Review*.

Cawley, J., J.J. Heckman, L. Lochner, and E. Vytlacil. 2000. "Understanding the role of cognitive ability in accounting for the recent rise in the economic return to education." In *Meritocracy and economic inequality*, edited by K. Arrow, S. Bowles and S. Durlauf. Princeton, NJ: Princeton University Press.

Chaikind, S., L.C. Danielson, and M.L. Brauen. 1993. "What do we know about the costs of special education? A selected review." *Journal of Special Education* 26,no.4:344-370.

Chambers, J.G. 1977. "The impact of collective bargaining for teachers on resource allocation in public school districts." *Journal of Urban Economics* 4,no.3 (July):324-339.

Charnes, A., W.W. Cooper, and E. Rhodes. 1978. "Measuring the efficiency of decision making units." *European Journal of Operational Research* 2(November):429-444.

Chubb, J.E., and T.M. Moe. 1990. *Politics, markets and America's schools*. Washington, D.C.: The Brookings Instituitons.

Clotfelter, C.T. 1976. "School desegregation, "tipping," and private school enrollment." *Journal of Human Resources* 22,no.1 (Winter):29-50.

———. 2001. "Are white still fleeing? Racial patterns and enrollment shifts in urban public schools." *Journal of Policy Analysis and Management* 20,no.2 (Spring):199-221.

Clune, W.H. 1994. "The shift from equity to adequacy in school finance." *Educational Policy* 8,no.4 (December):376-394.

Cohen, D.K., and R.J. Murnane. 1986. "Merit pay and the evaluation problem: Understanding why most merit pay plans fail and a few survive." *Harvard Educational Review* 56,no.1 (February):1-17.

Cohn, E. 1997. *Market approaches to education: Vouchers and school choice*. Oxford, UK: Pergamon.

Cohn, E., and T.G. Geske. 1990. *The economics of education*. 3rd ed. Oxford: Pergamon Press.

Coleman, J.S., E.Q. Campbell, C.J. Hobson, J. McPartland, A.M. Mood, F.D. Weinfeld, and R.L. York. 1966. *Equality of educational opportunity*. Washington, D.C.: U.S. Government Printing Office.

Coleman, J.S., and T. Hoffer. 1987. *Public, Catholic, and private schools: The importance of community*. New York: Basic Books.

Coleman, J.S., T. Hoffer, and S. Kilgore. 1982. *High school achievement: public, Catholic and private schools compared*. NY: Basic Books.

Coleman, J.S., S.D. Kelley, and J.A. Moore. 1975. "Trends in school integration, 1968-73." Paper 722-03-01, Urban Institute

Cook, M.D., and W.N. Evans. 2000. "Families or schools? Explaining the convergence in white and black academic performance." *Journal of Labor Economics* 18,no.4 (October):729-754.

Cook, T.D. 1984. "What have black children gained academically from school desegregation? A review of the metaanalytic evidence." In *School desegregation*. Washington, DC: National Institute of Education.

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

Courant, P.N., and S. Loeb. 1997. "Centralization of school finance in Michigan." *Journal of Policy Analysis and Management* 16,no.1 (Winter):114-136.

Craig, S.G., and R.P. Inman. 1982. "Federal aid and public education: An empirical look at the new fiscal federalism." *Review of Economics and Statistics* 64,no.4 (November):541-552.

———. 1986. "Education, welfare, and the "new" federalism: State budgeting in a federalist public economy." In

Studies in state and local public finance, edited by H.S. Rosen. Chicago: University of Chicago Press.

Crain, R. 1970. "School integration and occupational achievement of Negroes." *American Journal of Sociology* 75,no.4, Part II (January):593-606.

Crain, R.L., and R.E. Mahard. 1978. "Desegregation and black achievement: A review of the research." *Law and Contemporary Problems* 42,no.3:17-53.

Cullen, J.B. forthcoming. "The impact of fiscal incentives on student disability rates." *Journal of Public Economics*.

Cullen, J.B., B. Jacob, and S. Levitt. 2000. "The impact of school choice on student outcomes: An analysis of the Chicago public schools."

Cullen, J.B., and S.G. Rivkin. forthcoming. "The role of special education in school choice." In *The Economics of School Choice*, edited by C.M. Hoxby. Chicago, IL: University of Chicago Press.

Currie, J. 2001. "Early childhood education programs." *Journal of Economic Perspectives* 15,no.2 (Spring):213-238.

Currie, J., and D. Thomas. 1995. "Does Head Start make a difference?" *American Economic Review* 85,no.3 (June):341-364.

———. 1999. "Does Head Start help Hispanic children?" *Journal of Public Economics* 74:235-262.

———. 2000. "Early test scores, socioeconomic status, school quality, and future outcomes.

Dale, S.B., and A.B. Krueger. 1999. ""Estimating the payoff to attending a more selective college: An application of selection on observables and unobservables"."

de Bartolome, C.A.M. 1990. "Equilibrium and inefficiency in a community model with peer group effects." *Journal of Political Economy* 98,no.1 (February):110-133.

Deere, D. 2001. "Trends in wage inequality in the United States." In *The causes and consequences of increasing inequality*, edited by F. Welch. Chicago: University of Chicago Press.

Deere, D., and W. Strayer. 2001. "Putting schools to the test: School accountability, incentives, and behavior." Department of Economics, Texas A&M University (March 2001).

Downes, T.A. 1992. "Evaluating the impact of school finance reform on the provision of public education: The California case." *National Tax Journal* 45,no.4 (December):405-419.

Downes, T.A., and T.F. Pogue. 1994. "Adjusting school aid formulas for the higher cost of educating disadvantaged students." *National Tax Journal* 47,no.1 (March):89-110.

Downes, T.A., and M.P. Shah. 1996. "The effect of school finance reforms on the level and growth of per pupil expenditures." Tufts University Working Paper No. 95-4 (June).

Dugan, D.J. 1976. "Scholastic achievement: its determinants and effects in the education industry." In *Education as an industry*, edited by J.T. Froomkin, D.T. Jamison and R. Radner. Cambridge, MA: Ballinger.

Duncombe, W., J. Ruggiero, and J. Yinger. 1996. "Alternative approaches to measuring the cost of education." In *Holding schools accountable: Performance-based reform in education*, edited by H.F. Ladd. Washington, DC: Brookings.

Ehrlich, I. 1975. "On the relation between education and crime." In *Education, income, and human behavior*, edited

by F.T. Juster. New York: McGraw-Hill.

Eide, E., D.J. Brewer, and R.G. Ehrenberg. 1998. "Does it pay to attend an elite private college? Evidence on the effects of undergraduate college quality on graduate school attendance." *Economics of Education Review* 17,no.4 (October):371-376.

Entwisle, D.R., and K.L. Alexander. 1992. "Summer setback: Race, poverty, school composition, and mathematics achievement in the first two years of school." *American Sociological Review* 57,no.1 (February):72-84.

Epple, D., and R.E. Romano. 1996. "Ends against the middle: Determinating public service provision when there are private alternatives." *Journal of Public Economics* 62,no.3 (November):297-325.

———. 1998. "Competition between private and public schools, vouchers, and peer-group effects." *American Economic Review* 88,no.1 (March):33-62.

———. forthcoming. "Neighborhood schools, choice and the distribution of educational benefits." In *The Economics of School Choice*, edited by C.M. Hoxby. Chicago, IL: University of Chicago Press.

Evans, W.N., and R.M. Schwab. 1995. "Finishing high school and starting college: Do Catholic schools make a difference?" *Quarterly Journal of Economics* 110,no.4 (November):941-974.

Fair, R.C. 1971. "The optimal distribution of income." *Quarterly Journal of Economics* 85,no.4 (November):551-579.

Fairlie, R.W., and A.M. Resch. forthcoming. "Is there "white flight" into private schools? Evidence from the National Educational Longitudinal Survey." *Review of Economics and Statistics*.

Farrell, P., and V.R. Fuchs. 1982. "Schooling and health: The Cigarette Connection." *Journal of Health Economics* 1,no.3 (December):217-230.

Feldstein, M.S. 1975. "Wealth neutrality and local choice in public education." *American Economic Review* 65,no.1 (March):75-89.

Fernandez, R., and R. Rogerson. 1996. "Income distribution, communities, and the quality of public education." *Quarterly Journal of Economics* 111,no.1 (February):135-164.

———. 1997. "Education finance reform: A dynamic perspective." *Journal of Policy Analysis and Management* 16,no.1 (Winter):67-84.

———. 1998. "Public education and income distribution: A dynamic quantitative evaluation of education-finance reform." *American Economic Review* 88,no.4 (September):813-833.

Figlio, D.N. 1997. "Did the "tax revolt" reduce school performance?" *Journal of Public Economics* 65:245-269.

Filimon, R., T. Romer, and H. Rosenthal. 1982. "Asymmetric information and agenda control: The bases of monopoly power in public spending." *Journal of Public Economics* 17:51-70.

Finn, C.E.J., B.V. Manno, and G. Vanourek. 2000. *Charter schools in action*. Princeton, NJ: Princeton University Press.

Fischel, W.A. 1989. "Did Serrano cause Proposition 13?" *National Tax Journal* 42(December):465-474.

Fiske, E.B., and H.F. Ladd. 2000. *When schools compete: A cautionary tale*. Washington, DC: Brookings Institution Press.

Flyer, F., and S. Rosen. 1997. "The new economics of teachers and education." *Journal of Labor Economics* 15,no.1 (part 2) (January):104-139.

Friedman, M. 1962. *Capitalism and freedom*. Chicago: University of Chicago.

Fuller, B. 1985. *Raising school quality in developing countries: what investments boost learning?* Washington, DC: The World Bank.

Glass, G.V., and M.L. Smith. 1979. "Meta-analysis of research on class size and achievement." *Educational Evaluation and Policy Analysis* 1,no.1:2-16.

Glenn, C.L. 1989. *Choice in schools in six nations*. Washington,DC: U.S. Government Printing Office.

Glomm, G., and B. Ravikumar. 1992. "Public vs. private investment in human capital: Endogenous growth and income inequality." *Journal of Political Economy* 100,no.4 (August):818-834.

———. 1998. "Opting out of publicly provided services: A majority voting result." *Social Choice and Welfare* 15:187-199.

Goldin, C. 1998. "America's graduation from high school: The evolution and spread of secondary schooling in the twentieth century." *journal of economic history* 58(June):345-374.

Goldstein, G.S., and M.V. Pauly. 1981. "Tiebout bias on the demand for local public goods." *Journal of Public Economics* 16:131-144.

Gramlich, E.M. 1977. "Intergovernmental grants: A review of the empirical literature." In *The political economy of fiscal federalism*, edited by W.E. Oates. Lexington, MA: Heath-Lexington.

———. 1986. "Evaluation of education projects: The case of the Perry preschool program." *Economics of Education Review* 5,no.1:17-24.

Gramlich, E.M., and P.P. Koshel. 1975. *Educational performance contracting*. Washington, D.C.: The Brookings Institution.

Gramlich, E.M., and D.L. Rubinfeld. 1982. "Micro estimates of public spending demand functions and tests of the Tiebout and median voter hypotheses." *Journal of Political Economy* 90,no.3 (June):536-560.

Greenberg, D., and J. McCall. 1974. "Teacher mobility and allocation." *Journal of Human Resources* 9,no.4 (Fall):480-502.

Greene, J.P. 2001a. "An Evaluation of the Florida A-Plus Accountability and School Choice Program." New York: Center for Civic Innovation, Manhattan Institute (November).

———. 2001b. "The looming shadow: Florida gets its 'F' schools to shape up." *Education Next* 1,no.4 (Winter):76-82.

Greene, J.P., P.E. Peterson, and J. Du. 1998. "School choice in Milwaukee: A randomized experiment." In *Learning from school choice*, edited by P.E. Peterson and B.C. Hassel. Washington, DC: Brookings Institution.

Greenwald, R., L.V. Hedges, and R.D. Laine. 1996. "The effect of school resources on student achievement." *Review of Educational Research* 66,no.3 (Fall):361-396.

Griliches, Z. 1974. "Errors in variables and other unobservables." *Econometrica* 42,no.6 (November):971-998.

Grogger, J., and D. Neal. 2000. "Further evidence on the effects of Catholic secondary schooling." In *Brookings-Wharton papers on urban affairs, 2000*, edited by W.G. Gale and J.R. Pack. Washington, DC: Brookings Institution Press.

Grogger, J.T. 1996. "Does school quality explain the recent black/white wage trend?" *Journal of Labor Economics* 14,no.2 (April):231-253.

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

Grosskopf, S., K.J. Hayes, L.L. Taylor, and W.L. Weber. 1997. "Budget-constrained frontier measures of fiscal equality and efficiency in schooling." *Review of Economics and Statistics* 74,no.1 (February):116-124.

Gyimah-Brempong, K., and A.O. Byapong. 1991. "Characteristics of education production functions: An application of canonical regression analysis." *Economics of Education Review* 10,no.1:7-17.

Hanushek, E.A. 1971. "Teacher characteristics and gains in student achievement: Estimation using micro data." *American Economic Review* 60,no.2 (May):280-288.

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

———. 1979. "Conceptual and empirical issues in the estimation of educational production functions." *Journal of Human Resources* 14,no.3 (Summer):351-388.

———. 1986. "The economics of schooling: Production and efficiency in public schools." *Journal of Economic Literature* 24,no.3 (September):1141-1177.

———. 1991. "When school finance 'reform' may not be good policy." *Harvard Journal on Legislation* 28,no.2 (Summer):423-456.

———. 1992. "The trade-off between child quantity and quality." *Journal of Political Economy* 100,no.1 (February):84-117.

———. 1995. "Interpreting recent research on schooling in developing countries." *World Bank Research Observer* 10,no.2 (August):227-246.

———. 1997a. "Assessing the effects of school resources on student performance: An update." *Educational Evaluation and Policy Analysis* 19,no.2 (Summer):141-164.

———. 1997b. "The productivity collapse in schools." In *Developments in School Finance, 1996*, edited by W.J. Fowler, Jr. Washington, DC: National Center for Education Statistics.

———. 1999. "Some findings from an independent investigation of the Tennessee STAR experiment and from other investigations of class size effects." *Educational Evaluation and Policy Analysis* 21,no.2 (Summer):143-163.

Hanushek, E.A., and J.F. Kain. 1972. "On the value of `equality of educational opportunity' as a guide to public policy." In *On equality of educational opportunity*, edited by F. Mosteller and D.P. Moynihan. New York: Random House.

Hanushek, E.A., J.F. Kain, J.M. Markman, and S.G. Rivkin. forthcoming. "Does peer ability affect student achievement?" *Journal of Applied Econometrics*.

Hanushek, E.A., J.F. Kain, and S.G. Rivkin. 2001a. "Disruption versus Tiebout improvement: The costs and benefits

of switching schools." WP 8479, National Bureau of Economic Research (September).

———. 2001b. "New evidence about *Brown v. Board of Education*: The Complex Effects of School Racial Composition on Achievement." National Bureau of Economic Research (December).

———. 2001c. "Why public schools lose teachers." WP 8599, National Bureau of Economic Research (November).

Hanushek, E.A., and D.D. Kimko. 2000. "Schooling, labor force quality, and the growth of nations." *American Economic Review* 90,no.5 (December):1184-1208.

Hanushek, E.A., C.K.Y. Leung, and K. Yilmaz. 2001. "Redistribution through education and other transfer mechanisms." WP 8588, National Bureau of Economic Research (Novermber).

Hanushek, E.A., and J.A. Luque. forthcoming. "Efficiency and equity in schools around the world." *Economics of Education Review*.

Hanushek, E.A., and w. others. 1994. *Making schools work: Improving performance and controlling costs*. Washington, DC: Brookings Institution.

Hanushek, E.A., and R.R. Pace. 1995. "Who chooses to teach (and why)?" *Economics of Education Review* 14,no.2 (June):101-117.

Hanushek, E.A., and M.E. Raymond. 2001. "The confusing world of educational accountability." *National Tax Journal* 54,no.2 (June):365-384.

Hanushek, E.A., and S.G. Rivkin. 1997. "Understanding the twentieth-century growth in U.S. school spending." *Journal of Human Resources* 32,no.1 (Winter):35-68.

———. forthcoming. "Does public school competition affect teacher quality?" In *The Economics of School Choice*, edited by C.M. Hoxby. Chicago, IL: University of Chicago Press.

Hanushek, E.A., S.G. Rivkin, and L.L. Taylor. 1996. "Aggregation and the estimated effects of school resources." *Review of Economics and Statistics* 78,no.4 (November):611-627.

Hanushek, E.A., and J.A. Somers. 2001. "Schooling, inequality, and the impact of government." In *The causes and consequences of increasing inequality*, edited by F. Welch. Chicago: University of Chicago Press.

Harbison, R.W., and E.A. Hanushek. 1992. *Educational performance of the poor: lessons from rural northeast Brazil*. New York: Oxford University Press.

Hartigan, J.A., and A.K. Wigdor. 1989. *Fairness in employment testing: Validity generalization, minority issues, and the general aptitude test battery*. Washington, DC: National Academy Press.

Hartman, R.W. 1973. "The rationale for federal support for higher education." In *Does college matter? Some evidence on the impacts of higher education*, edited by L.C. Solmon and P.J. Taubman. New York: Academic Press.

Hatry, H.P., J.M. Greiner, and B.G. Ashford. 1994. *Issues and case studies in teacher incentive plans*. 2nd ed. Washington, DC: Urban Institute.

Hauser, R.M. 1993. "The decline in college entry among African Americans: Findings in search of explanations." In *Prejudice, politics, and the American dilemma*, edited by P.M. Sniderman, P.E. Tetlock and E.G. Carmines. Stanford, CA: Stanford University Press.

Haveman, R.H., and B.L. Wolfe. 1984. "Schooling and economic well-being: The role of nonmarket effects." *Journal of Human Resources* 19,no.3 (Summer):377-407.

Heckman, J.J. 2000. "Policies to foster human capital." *Research in Economics* 54:3-56.

Heckman, J.J., R. LaLonde, and J. Smith. 1999. "The economics and econometrics of active labor market programs." In *Handbook of labor economics*, edited by O. Ashenfelter and D. Card. Amsterdam: Elsevier Science.

Heckman, J.J., A. Layne-Farrar, and P. Todd. 1996a. "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 G. Burtless. Washington, DC: Brookings.

———. 1996b. "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):562-610.

Heckman, J.J., L. Lochner, and C. Taber. 2000. "General equilibrium cost benefit analysis of education and tax policies." In *Trade, growth and development: Essays in honor of T.N. Srinivasan*, edited by G. Ranis and L.K. Raut. Amsterdam: Elsevier Science.

Heckman, J.J., and E. Vytlacil. 2001. "Identifying the role of cognitive ability in explaining the level of and change in the return to schooling." *Review of Economics and Statistics* 83,no.1 (February):1-12.

Hedges, L.V., R.D. Laine, and R. Greenwald. 1994. "Does money matter? A meta-analysis of studies of the effects of differential school inputs on student outcomes." *Educational Researcher* 23,no.3 (April):5-14.

Hedges, L.V., and I. Olkin. 1985. *Statistical methods for meta-analysis*. San Diego, CA: Academic Press.

Henderson, V., P. Mieszkowski, and Y. Sauvageau. 1976. *Peer group effects and educational production functions*. Ottawa, Canada: Economic Council of Canada.

Heyneman, S.P., and W. Loxley. 1983. "The effect of primary school quality on academic achievement across twenty-nine high and low income countries." *American Journal of Sociology* 88(May):1162-1194.

Heyneman, S.P., and D.S. White. 1986. *The quality of education and economic development*. Washington, DC: World Bank.

Heyns, B. 1978. *Summer learning and the effects of schooling*. New York: Academic Press.

Hirschman, A.O. 1970. *Exit, voice, and loyalty: Responses to decline in firms, organizations, and states*. Cambridge, MA: Harvard University Press.

Hoxby, C.M. 1994. ""Do private schools provide competition for public schools?"" National Bureau of Economic Research, Working Paper No. 4978, 1994 (December).

———. 1996a. "Are efficiency and equity in school finance substitutes or complements?" *Journal of Economic Perspectives* 10,no.4 (Fall):51-72.

———. 1996b. "How teachers' unions affect education production." *Quarterly Journal of Economics* 111,no.3 (August):671-718.

———. 2000a. "Does competition among public schools benefit students and taxpayers?" *American Economic Review* 90,no.5 (December):1209-1238.

———. 2000b. "The effects of class size on student achievement: New evidence from population variation."

Quarterly Journal of Economics 115,no.3 (November):1239-1285.

———. 2000c. "Peer Effects in the Classroom: Learning from Gender and Race Variation." Working Paper No. 7867, National Bureau of Economic Research (August).

———. 2001. "All school finance equalizations are not created equal." *Quarterly Journal of Economics* 116,no.4 (November):1189-1231.

———. forthcoming. "School choice and school productivity (or could school choice be a tide that lifts all boats?)." In *The Economics of School Choice*, edited by C.M. Hoxby.

Inman, R.P. 1987. "Markets, Governments, and the "new" political economy." In *Handbook of Public Economics*, edited by A.J. Auerbach and M. Feldstein. Amsterdam: North-Holland.

James, E. 1987. "The public/private division of responsibility for education: An international comparison." *Economics of Education Review* 6,no.1:1-14.

———. 1993. "Why do different countries choose a different public-private mix of educational services?" *Journal of Human Resources* 28,no.3 (Summer):571-592.

James, E., and N. Alsalam. 1993. "College choice, academic achievement and future earnings." In *Essays on the economics of education*, edited by E.P. Hoffman. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.

Jaynes, G.D., and R.M.J. Williams. 1989. *A common destiny: Blacks and American society*. Washington, DC: National Academy Press.

Jencks, C., M. Smith, H. Acland, M.J. Bane, D. Cohen, H. Gintis, B. Heyns, and S. Michelson. 1972. *Inequality: A re-assessment of the effects of family and schooling in America*. NY: Basic Books.

Jimenez, E., and Y. Sawada. 2001. "Public for private: The relationship between public and private school enrollment in the Philippines." *Economics of Education Review* 20,no.4 (August):389-399.

Johnson, G.E., and F.P. Stafford. 1973. "Social returns to quantity and quality of schooling." *Journal of Human Resources* 8,no.2 (Spring):139-155.

Jorgenson, D.W., and B.M. Fraumeni. 1992. "Investment in education and U.S. economic growth." *Scandinavian Journal of Economics* 94,no.Supplement:51-70.

Juhn, C., K.M. Murphy, and B. Pierce. 1993. "Wage inequality and the rise in returns to skill." *Journal of Political Economy* 101,no.3 (June):410-442.

Kain, J.F., and J.M. Quigley. 1975. *Housing markets and racial discrimination*. NY: National Bureau of Economic Research.

Kane, T. 1994. "College enrollments by blacks since 1970: The role of college costs, family background, and the returns to education." *Journal of Political Economy* 102,no.5 (October):878-911.

Katz, L.F., J.R. Kling, and J.B. Liebman. 2001. "Moving to Opportunity in Boston: Early results of a randomized mobility experiment." *Quarterly Journal of Economics* 116,no.2 (May):607-654.

Kiker, B.F. 1968. *Human capital: In retrospect*. Columbia, SC: University of South Carolina.

Koretz, D. 1996. "Using student assessments for educational accountability." In *Improving America's schools: The*

role of incentives, edited by E.A. Hanushek and D.W. Jorgenson. Washington, DC: National Academy Press.

Kosters, M.H. 1991. "Wages and demographics." In *Workers and their wages*, edited by M.H. Kosters. Washington, DC: The AEI Press.

Krueger, A.B. 1999. "Experimental estimates of education production functions." *Quarterly Journal of Economics* 114,no.2 (May):497-532.

———. 2000. "An economist's view of class size research." The Class Size Debate, Working Paper, Economic Policy Institute

Ladd, H.F. 1975. "Local educational expenditures, fiscal capacity, and the composition of the property tax base." *National Tax Journal* 28:145-158.

Ladd, H.F., and J.S. Hansen. 1999. *Making money matter: Financing America's schools*. Washington, DC: National Academy Press.

Lankford, H., and J. Wyckoff. 1992. "Primary and secondary school choice among public and religious alternatives." *Economics of Education Review* 11,no.4 (December):317-337.

———. 1996. "The allocation of resources to special education and regular instruction." In *Holding schools accountable: Performance-based reform in education*, edited by H.F. Ladd. Washington, DC: Brookings.

Layard, R., and G. Psacharopoulos. 1974. "The screening hypothesis and the returns to education." *Journal of Political Economy* 82,no.5 (September/October):985-998.

Lazear, E.P. 2001. "Educational production." *Quarterly Journal of Economics* 116,no.3 (August):777-803.

Leibowitz, A. 1974. "Home investments in children." *Journal of Political Economy* 82,no.2, Pt II (March/April):S111-S131.

Levin, H.M., and P.J. McEwan. 2001. *Cost-effectiveness analysis: Methods and applications*. 2nd ed. Thousand Oaks, CA: Sage Publications, Inc.

Levinson, A.M. 1988. "Reexamining teacher preferences and compensating wages." *Economics of Education Review* 7,no.3:357-364.

Levy, F., and R.J. Murnane. 1992. "U.S. earnings levels and earnings inequality: A review of recent trends and proposed explanations." *Journal of Economic Literature* 30,no.3 (September):1333-1381.

Lindahl, M. 2000. "Home versus school learning: A new approach to estimating the effect of class size on achievement."

Lochner, L., and E. Moretti. 2001. "The Effect of Education on Crime: Evidence from Prison Inmates, Arrests, and Self-Reports." WP 8605, Cambridge, MA: National Bureau of Economic Research (November).

Lockheed, M.E., and A. Verspoor. 1991. *Improving primary education in developing countries*. New York: Oxford University Press.

Loeb, S. 2001. "Estimating the effects of school finance reform: A framework for a federalist system." *Journal of Public Economics*(May).

Loeb, S., and M.E. Page. 2000. "Examining the link between teacher wages and student outcomes: The importance of alternative labor market opportunities and non-pecunieary variation." *Review of Economics and Statistics*

82,no.3 (August):393-408.

Lucas, R.E. 1988. "On the mechanics of economic development." *Journal of Monetary Economics* 22(July):3-42.

Ludwig, J., G.J. Duncan, and P. Hirschfield. 2001. "Urban poverty and juvenile crime: Evidence from a randomized housing-mobility experiment." *Quarterly Journal of Economics* 116,no.2 (May):655-679.

Mankiw, N.G., D. Romer, and D. Weil. 1992. "A contribution to the empirics of economic growth." *Quarterly Journal of Economics* 107,no.2 (May):407-437.

Manski, C.F. 1992. "Educational Choice (Vouchers) and Social Mobility." *Economics of Education Review* 11,no.4 (December):351-369.

———. 1993a. "Adolescent econometricians: How do youth infer the returns to schooling?" In *Studies in supply and demand in higher education*, edited by C.T. Clotfelter and M. Rothschild. Chicago: University of Chicago Press.

———. 1993b. "Identification of endogenous social effects: The reflection problem." *Review of Economic Studies* 60(July):531-542.

———. 2000. "Economic analysis of social interactions." *Journal of Economic Perspectives* 14,no.3 (Summer):115-136.

Manski, C.F., and D.A. Wise. 1983. *College choice in America*. Cambridge: Harvard University Press.

Mayer, S.E. 1997. *What money can't buy: Family income and children's life chances*. Cambridge, MA: Harvard University Press.

Mayer, S.E., and C. Jencks. 1989. "Growing Up in Poor Neighborhoods: How Much Does It Matter?" *Science* 243(March 17):1441-1445.

McEwan, P.J. forthcoming. "The effectiveness of public, Catholic, and non-religious private schools in Chile's voucher system." *Education Economics*.

McEwan, P.J., and M. Carnoy. 2000. "The effectiveness and efficiency of private schools in Chile's voucher system." *Educational Evaluation and Policy Analysis* 22,no.3 (Fall):213-240.

McMahon, W.W. 1991. "Relative returns to human and physical capital in the U.S. and efficient investment strategies." *Economics of Education Review* 10,no.4:283-296.

Mendro, R.L., H.R. Jordan, E. Gomez, M.C. Anderson, and K.L. Bembry. 1998. "An Application of Multiple Linear Regression in Determining Longitudinal Teacher Effectiveness." Paper presented at *1998 Annual Meeting of the American Educational Research Association*, April 1998, at San Diego, CA.

Michael, R.T. 1982. "Measuring non-monetary benefits of education: A survey." In *Financing education: Overcoming inefficiency and inequity*, edited by W.W. McMahon and T.G. Geske. Urbana, IL: University of Illinois Press.

Mincer, J. 1970. "The distribution of labor incomes: a survey with special reference to the human capital approach." *Journal of Economic Literature* 8,no.1 (March):1-26.

———. 1974. *Schooling Experience and Earnings*. New York: NBER.

Moffitt, R.A. 2001. "Policy interventions, low-level equilibria, and social interactions." In *Social dynamics*, edited by

S. Durlauf and H.P. Young. Cambridge, MA: MIT Press.

Mosteller, F. 1995. "The Tennessee study of class size in the early school grades." *The Future of Children* 5,no.2 (Summer/Fall):113-127.

Mundel, D.S. 1973. "Whose education should society support?" In *Does college matter? Some evidence on the impacts of higher education*, edited by L.C. Solmon and P.J. Taubman. New York: Academic Press.

Murnane, R.J. 1975. *Impact of school resources on the learning of inner city children*. Cambridge, MA: Ballinger.

———. 1981. "Teacher mobility revisited." *Journal of Human Resources* 16,no.1 (Winter):3-19.

Murnane, R.J., S. Newstead, and R. Olsen. 1985. "Comparing public and private schools: The puzzling role of selectivity bias." *Journal of Business and Economic Statistics* 3,no.1 (January):23-35.

Murnane, R.J., and B. Phillips. 1981. "What do effective teachers of inner-city children have in common?" *Social Science Research* 10,no.1 (March):83-100.

Murnane, R.J., J.D. Singer, J.B. Willett, J.J. Kemple, and R.J. Olsen. 1991. *Who will teach?* Cambridge: Harvard University Press.

Murnane, R.J., J.B. Willett, M.J. Braatz, and Y. Duhaldeborde. 2001. "Do different dimensions of male high school students' skills predict labor market success a decade later? Evidence from the NLSY." *Economics of Education Review* 20,no.4 (August):311-320.

Murnane, R.J., J.B. Willett, Y. Duhaldeborde, and J.H. Tyler. 2000. "How important are the cognitive skills of teenagers in predicting subsequent earnings?" *Journal of Policy Analysis and Management* 19,no.4 (Fall):547-568.

Murnane, R.J., J.B. Willett, and F. Levy. 1995. "The growing importance of cognitive skills in wage determination." *Review of Economics and Statistics* 77,no.2 (May):251-266.

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

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

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

Nathan, J. 1989. "Progress, problems, and prospects with state choice plans." In *Public schools by choice: Expanding opportunities for parents, students, and teachers*, edited by J. Nathan. St. Paul, MN: Institute for Learning and Teaching.

———. 1996. *Charter schools: Creating hope and opportunity for American education*. San Francisco: Jossey-Bass Publishers.

Neal, D. 1997. "The effect of Catholic secondary schooling on educational attainment." *Journal of Labor Economics* 15,no.1, Part 1 (January):98-123.

———. 1998. "What have we learned about the benefits of private schooling?" *FRBNY Economic Policy Review* 4(March):79-86.

Neal, D.A., and W.R. Johnson. 1996. "The role of pre-market factors in black-white differences." *Journal of Political*

Economy 104,no.5 (October):869-895.

Nechyba, T.J. 1999. "School finance induced migration patterns: The impact of private school vouchers." *Journal of Public Economic Theory* 1,no.1:5-50.

———. 2000. "Mobility, targeting, and private-school vouchers." *American Economic Review* 90,no.1 (March):130-146.

———. 2001. "Centralization, segregation, and the nature of communities." Working paper, Department of Economics, Duke University

———. forthcoming-a. "Centralization, fiscal federalism and private school attendance." *International Economic Review*.

———. forthcoming-b. "Introducing school choice into multi-district public school systems." In *The Economics of School Choice*, edited by C.M. Hoxby. Chicago, IL: University of Chicago Press.

Nelson, R.R., and E. Phelps. 1966. "Investment in humans, technology diffusion and economic growth." *American Economic Review* 56,no.2 (May):69-75.

Noell, J. 1982. "Public and Catholic schools: a reanalysis of public and private schools." *Sociology of Education* 55(April/July):123-132.

Oakes, J. 1985. *Keeping track: How schools structure inequality*. New Haven, CT: Yale University Press.

———. 1992. "Can tracking research inform practice? Technical, normative, and political considerations." *Educational Researcher* 21,no.4:12-21.

Oates, W. 1969. "The effects of property taxes and local public spending on property values: an empirical study of tax capitalization and the Tiebout hypothesis." *Journal of Political Economy* 77,no.6 (November/December):957-971.

Oates, W.E. 1994. "Federalism and government finance." In *Modern public finance*, edited by J.M. Quigley and E. Smolensky. Cambridge, MA: Harvard University Press.

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

O'Regan, K.M., and J.M. Quigley. 1999. "Accessibility and economic opportunity." In *Essays in transportation economics*, edited by C. Winston, J.A. Gomez-Ibanez and W. Tye. Washington, DC: Brookings Institution.

Organisation for Economic Co-operation and Development. 2001. *Education at a glance*. Paris: Organization for Economic Co-operation and Development.

Peltzman, S. 1993. "The political economy of the decline of American public education." *Journal of Law and Economics* 36(April):331-370.

Peterson, P.E., J.P. Greene, and C. Noyes. 1996. "School choice in Milwaukee." *Public Interest* 125(Fall):38-56.

Peterson, P.E., W.G. Howell, P.J. Wolf, and D.E. Campbell. forthcoming. "School vouchers: Results from randomized experiments." In *The Economics of School Choice*, edited by C.M. Hoxby. Chicago, IL: University of Chicago Press.

Pierce, B., and F. Welch. 1996. "Changes in the structure of wages." In *Improving America's schools: The role of*

incentives, edited by E.A. Hanushek and D.W. Jorgenson. Washington, DC: National Academy Press.

Poterba, J.M. 1996. "Government intervention in the markets for education and health care: How and why?" In *Individual and social responsibility: Child care, education, medical care, and long-term care in America"*, edited by V.R. Fuchs. Chicago: University of Chicago/National Bureau of Economic Research.

———. 1997. "Demographic structure and the political economy of public education." *Journal of Policy Analysis and Management* 16,no.1 (Winter):48-66.

Pritchett, L., and D. Filmer. 1999. "What education production functions really show: A positive theory of education expenditures." *Economics of Education Review* 18,no.2 (April):223-239.

Psacharopoulos, G., J.-P. Tan, and E. Jimenez. 1986. *Financing education in developing countries: An exploration of policy options*. Washington, DC: World Bank.

Riley, J.G. 2001. "Silver signals: Twenty-five years of screening and signaling." *Journal of Economic Literature* 39,no.2 (June):432-478.

Rivkin, S.G. 1995. "Black/white differences in schooling and employment." *Journal of Human Resources* 30,no.4 (Fall):826-852.

———. 2000. "School desegregation, academic attainment, and earnings." *Journal of Human Resources* 35,no.2 (Spring):333-346.

Rivkin, S.G., E.A. Hanushek, and J.F. Kain. 2001. "Teachers, schools, and academic achievement." Working Paper No. 6691, National Bureau of Economic Research (revised)

Rizzuto, R., and P. Wachtel. 1980. "Further evidence on the returns to school quality." *Journal of Human Resources* 15,no.2 (Winter):240-254.

Romer, P. 1990a. "Endogenous technological change." *Journal of Political Economy* 99,no.5,pt. II:S71-S102.

———. 1990b. "Human capital and growth: Theory and evidence." *Carnegie-Rochester Conference Series on Public Policy* 32:251-286.

Romer, T., and H. Rosenthal. 1979. "Bureaucrats vs. voters: On the political economy of resource allocation by direct democracy." *Quarterly Journal of Economics* 93,no.4 (November):563-587.

Romer, T., H. Rosenthal, and V.G. Munley. 1992. "Economic incentives and political institutions: Spending and voting in school budget referenda." *Journal of Public Economics* 49:1-33.

Rosen, H.S., and D.J. Fullerton. 1977. "A note on local tax rates, public benefit levels and property values." *Journal of Political Economy* 85,no.2 (April):433-440.

Rosenbaum, J.E. 1995. "Changing the geography of opportunity by expanding residential choice: Lessons from the Gautreaux Program." *Housing Policy Debate* 6,no.1:231-269.

Rosenbaum, J.E., and S.J. Popkin. 1991. "Employment and earnings of low-income Blacks who move to middle-class suburbs." In *The urban underclass*, edited by C. Jencks and P.E. Peterson. Washington, DC: Brookings Institution.

Rothstein, P. 1992. "The demand for education with 'power equalizing' aid: Estimation and simulation." *Journal of Public Economics* 49:135-162.

Rothstein, R., and K.H. Miles. 1995. *Where's the money gone? Changes in the level and composition of education spending*. Washington, DC: Economic Policy Institute.

Rouse, C.E. 1998. "Private school vouchers and student achievement: An evaluation of the Milwaukee Parental Choice Program." *Quarterly Journal of Economics* 113,no.2 (May):553-602.

Rubinfeld, D.L., P. Shapiro, and J. Roberts. 1987. "Tiebout bias and the demand for local public schooling." *Review of Economics and Statistics* 69,no.3 (August):426-437.

Ruggiero, J. 1996. "Efficiency of educational production: An analysis of New York school districts." *Review of Economics and Statistics* 78,no.3 (August):499-509.

Sander, W. 1996. "Catholic grade schools and academic achievement." *Journal of Human Resources* 31,no.3 (Summer):540-548.

———. 1999. "Private schools and public school achievement." *Journal of Human Resources* 34,no.4 (Fall):697-709.

———. 2001. *Catholic schools: Private and social effects*. Boston: Kluwer Academic Publishers.

Sanders, W.L., and S.P. Horn. 1994. "The Tennessee value-added assessment system (TVAAS): Mixed-model methodology in educational assessment." *Journal of Personnel Evaluation in Education* 8:299-311.

———. 1995. "The Tennessee Value-Added Assessment System (TVAA): Mixed model methodology in educational assessment." In *Teacher evaluation: Guide to effective practice*, edited by A.J. Shinkfield and D.L. Stufflebeam. Boston: Kluwer Academic Publishers.

Schacter, J. 2001a. "Geographic mobility: March 1999 to March 2000." In *Current Population Reports*. Washington, DC: U.S. Census Bureau.

———. 2001b. "Why people move: Exploring the March 2000 Current Population Series." In *Current Population Reports*. Washington, DC: U.S. Census Bureau.

Schofield, J.W. 1995. "Review of research on school desegregation's impact on elementary and secondary school students." In *Handbook of research on multicultural education*, edited by J.A. Banks and C.A.M. Banks. New York: Macmillan Publishing.

Schultz, T.W. 1961. "Investment in human capital." *American Economic Review* 51,no.1 (March):1-17.

———. 1963. *The economic value of education*. New York: Columbia University Press.

———. 1975. "The value of the ability to deal with disequilibria." *Journal of Economic Literature* 13:827-846.

Scitovsky, T., and A. Scitovsky. 1959. "What price economic progress?" *Yale Law Review* 49(Autumn):95-110.

Sewell, W.H., and R.M. Hauser. 1975. *Education, occupation, and Earnings: Achievement in the early career*. New York: Academic Press.

Silva, F., and J. Sonstelie. 1995. "Did Serrano cause a decline in school spending?" *National Tax Journal* 48,no.2 (June):199-215.

Smith, J.P., and F. Welch. 1989. "Black economic progress after Myrdal." *Journal of Economic Literature* 27,no.2 (June):519-564.

Solmon, L.C. 1973. "The definition and impact of college quality." In *Does college matter?*, edited by L.C. Solmon and P.J. Taubman. New York: Academic Press.

Somers, J.A. 1998. *The connection between educational spending and educational outcomes*. Dissertation, University of Rochester.

Speakman, R., and F. Welch. 1995. "Does school quality matter? --A Reassessment." Texas A&M University (mimeo) (January).

Spence, A.M. 1973. "Job market signalling." *Quarterly Journal of Economics* 87,no.3 (August):355-374.

Stanley, H.W., and R.G. Niemi. 2000. *Vital statistics on American politics*. Washington, DC: CQ Press.

Stecher, B.M., and G.W. Bohrnstedt. 1999. *Class size reduction in California: Early Evaluation Findings, 1996-98*. Palo Alto: American Institutes for Research.

Stern, D. 1973. "Effects of alternative state aid formulas on the distribution of public school expenditures in Massachusetts." *Review of Economics and Statistics* 55,no.1 (February):91-97.

Strayer, G.D., and R.M. Haig. 1923. *The financing of education in the State of New York*. New York: Macmillan.

Summers, A., and B. Wolfe. 1977. "Do schools make a difference?" *American Economic Review* 67,no.4 (September):639-652.

Summers, A.A., and A.W. Johnson. 1996. "The effects of school-based management plans." In *Improving America's schools: The role of incentives*, edited by E.A. Hanushek and D.W. Jorgenson. Washington, DC: National Academy Press.

Taber, C.R. 2001. "The rising college premium in the eighties: Return to college or return to unobserved ability?" *Review of Economic Studies* 68,no.3 (July):665-691.

Teixeira, R.A. 1992. *The disappearing American voter*. Washington, DC: Brookings.

Tiebout, C.M. 1956. "A pure theory of local expenditures." *Journal of Political Economy* 64(October):416-424.

Toder, E.J. 1972. "The supply of public school teachers to an urban metropolitan area: A possible source of discrimination in education." *Review of Economics and Statistics* 54,no.4 (November):439-443.

U.S. Bureau of the Census. 1975. *Historical statistics of the United States, colonial times to 1970, bicentennial edition*. Vol. 1 and 2. Washington, DC: U.S. Government Printing Office.

———. 2000. *Statistical abstract of the United States: 2000*. Washington, DC: U.S. Government Printing Office.

U.S. Commission on Civil Rights. 1967. *Racial isolation in the public schools*. Washington, DC: Government Printing Office.

U.S. Department of Education. 1996a. *Education indicators: An International Perspective*. Washington, DC: National Center for Education Statistics.

———. 1996b. *Pursuing Excellence: A study of U.S. eighth-grade mathematics and science teaching, learning, curriculum, and achievement in international context*. Washington, DC: National Center for Education Statistics.

———. 1997. *Digest of education statistics*. 1997 ed. Washington, DC: National Center for Education Statistics.

———. 1999. *Highlights from TIMSS: The Third International Mathematics and Science Study*. Washington, DC: National Center for Education Statistics.

———. 2000. *Trends in academic progress: Three decades of student performance*. Washington, DC: National Center for Education Statistics.

U.S. General Accounting Office. 1997. "Headstart: Research provides little information on impact of current program. Washington, DC: U.S. General Accounting Office.

Vignoles, A., R. Levacic, J. Walker, S. Machin, and D. Reynolds. 2000. ""The relationship between resource allocation and pupil attainment: A review"."

Wachtel, P. 1975. "The effect of school quality on achievement, attainment levels, and lifetime income." *Explorations in Economic Research* 2,no.4:502-536.

———. 1976. "The effect on earnings of school and college investment expenditures." *Review of Economics and Statistics* 58,no.3 (August):326-331.

Walberg, H.J., and H.J. Walberg III. 1994. "Losing local control." *Educational Researcher* 22,no.5 (June-July):19-26.

Weimer, D.L., and M.J. Wolkoff. 2001. "School performance and housing values: Using non-contiguous district and incorporation boundaries to identify school effects." *National Tax Journal* 54,no.2 (June):231-253.

Weinberg, B.A. 2001. "An incentive model of the effect of parental income on children." *Journal of Political Economy* 109,no.2:266-280.

Weisbrod, B.A. 1964. *External benefits of public education: An economic analysis*. Princeton, NJ: Princeton University.

Weiss, A. 1995. "Human capital vs. signalling explanations of wages." *Journal of Economic Perspectives* 9,no.4 (Fall):133-154.

Welch, F. 1966. "Measurement of the quality of schooling." *American Economic Review* 56,no.2 (May):379-392.

———. 1970. "Education in production." *Journal of Political Economy* 78,no.1 (January/February):35-59.

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

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

West, E.G. 1996. "Education vouchers in practice and principle: A world survey."

Willis, R.J. 1973. "A new approach to the economic theory of fertility behavior." *Journal of Political Economy* 81(Pt. II April/March,):S14-S64.

Witte, J.F., Jr. 1999. *The market approach to education*. Princeton, NJ: Princeton University Press.

Woessman, L. 2000. "Schooling resources, educational institutions, and student performance: The international evidence." Kiel Working Paper No. 983, Kiel Institute of World Economics, Kiel, Germany, December 200

———. 2001. "Why students in some countries do better." *Education Matters* 1,no.2 (Summer):67-74.

Wolfe, B.L., and S. Zuvekas. 1995. "Nonmarket outcomes of schooling."

Wolpin, K.I. 1977. "Education and screening." *American Economic Review* 67,no.5 (December):949-958.

Word, E., J. Johnston, H.P. Bain, B.D. Fulton, J.B. Zaharies, M.N. Lintz, C.M. Achilles, J. Folger, and C. Breda. 1990. *Student/teacher achievement ratio (STAR), Tennessee's K-3 class size study: Final summary report, 1985-1990*. Nashville, TN: Tennessee State Department of Education.

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

Zimmer, R.W., and E.F. Toma. 2000. "Peer effects in private and public schools across countries." *Journal of Policy Analysis and Management* 19,no.1 (Winter):75-92.