Are private schools better than public schools?:
A critique of the Coleman Report

by Michael Olneck

An associate professor in the Departments of Educational Policy Studies and Sociology at the University of Wisconsin-Madison, Michael Olneck is also on the research staff of the Institute for Research on Poverty.

On April 12 of this year, the New York Times carried an article with the understated headline, "Remarks by Sociologist Stir Debate over Schools." The Washington Post was more blunt: "Private High Schools Are Better Than Public, Study Concludes." The object of their attention was a newly released draft report prepared for the National Center for Education Statistics (NCES) by University of Chicago sociologist James Coleman and two collaborators, Thomas Hoffer and Sally Kilgore. The report, "Public and Private Schools," analyzed 1980 data from a large-scale National Opinion Research Center survey, High School and Beyond, and after comparing students in public, Catholic, and other private schools, concluded that nonpublic schools provided more orderly and demanding learning environments, taught students more, did not aggravate racial segregation, and, in the case of Catholic schools, reduced the strength of the link between family background and academic achievement.

Adding to the controversial nature of "Public and Private Schools" were analyses that the authors said demonstrated that policies facilitating the use of nonpublic schools (e.g., tuition tax credits, vouchers) would favor minorities and the disadvantaged rather than well-to-do whites. Indeed, the authors introduced and concluded their report by assessing the implications of their study for the validity of arguments supporting or opposing the increased use of nonpublic schooling. Their judgment was that it "is hard . . . to avoid the overall conclusion that the factual premises underlying policies that would facilitate use of private schools are much better supported on the whole than those underlying policies that would constrain their use."

In addition to the press attention given to "Public and Private Schools," the National Institute of Education and the National Academy of Sciences recently convened
conferences of experts to assess the merits and validity of the report. The verdicts of scholars are mixed. They range from an eminent econometrician’s assessment that “the quality of documentation, analysis, and interpretation is so defective that it is hard to avoid the overall conclusion that the report reeks with incompetence and irresponsibility,” to the observation by a member of the National Academy of Education that “Coleman, who enjoys an international reputation as a meticulous scholar,” had “dramatically” reversed the conclusion imputed to his 1966 study of school effects that “schools don’t make a difference” and demonstrated that “schools do make a difference, regardless of the family background of students.”

My own judgment is that the report’s data and analyses are inadequate to answer the question of whether nonpublic schools are more successful as educational institutions than are public schools. Coleman and his coauthors tend to exaggerate and place too much confidence in their results, and they did not carry out a number of potentially instructive analyses. Nevertheless, their methods for analyzing achievement outcomes are reasonable, some of their results are plausible even if unconvincing, and the flaws and inadequacies in the report constitute insufficient reason to dismiss its findings out of hand. The report’s conclusions concerning the effects of private schools on segregation, on the other hand, are entirely unwarranted, and rest on questionable empirical findings used for polemical purposes.

Technical limitations in the report

Drawbacks of the sample

The 1980 High School and Beyond Survey collected data on over 58,000 high school sophomores and seniors. But because schools, not individuals, were the primary unit sampled, data were gathered in only 27 private non-Catholic schools and in only 84 Catholic schools. The number of public schools was 894. This means that estimates of parent academic advantage of the nonpublic schools. The most obvious difference is that some parents or students have made a choice to abandon the public schools. Coleman and his coauthors attempt to cope with this problem by statistically holding constant an array of social background measures which are likely to affect both choice of school sector and academic achievement. The trouble is that even when these measures are controlled, it is easy to imagine some characteristic which varies appreciably among students from very similar background, and which can affect both the likelihood of attending a nonpublic school and achievement. Scholastic ability or aptitude at the start of Grade 9 is one obvious characteristic.

In the absence of an ability measure, there do exist some additional analytic strategies which if employed would give us greater confidence in the result than does merely controlling measured family background characteristics. One would be to hold constant the curricular program in which a student was enrolled on the assumption that a student’s track reflects prior differences in ability and learning rather than produces new differences. Two-thirds of the private school students were enrolled in college preparatory programs, while only one-third of the public school students were. College preparatory students in the public schools scored no lower on achievement tests
than did students as a whole in the nonpublic schools. In a New York Times interview on April 26, Coleman rejected the idea of comparing only students in the same tracks. He argued that curricular program “is not a ‘background’ characteristic for which you should control statistically. It has a lot to do with school policies.” His argument is unpersuasive because there is no convincing evidence that curricular track exercises an appreciable effect on achievement growth. Instead, the best evidence is that secondary school curricular placement reflects prior achievement levels rather than determines current achievement. Comparisons among students should take these prior differences into account.

A second strategy would be to use the achievement levels of the sophomores in each of the sampled schools as a measure of a school’s selectivity, and to attempt to explain with this measure any average achievement advantage among seniors not attributable to the background composition of the senior class.

A third alternative for controlling initial differences between private school and public school students would be to apply econometric techniques for eliminating selectivity biases when estimating structural equation models. These techniques rely on a variety of untestable assumptions, and if the data do not conform to these assumptions, inferences based on them can be erroneous. Still, greater confidence could be placed in the report’s conclusions had the authors demonstrated that their results persisted in the face of such tests.

The controversy over whether or not private school students score higher on achievement tests because they know more to begin with or because they learn more will be closer to resolution in 1982, after the 1980 sophomores are retested. Even then, a skeptic could object that private school sophomores with the same test scores as public school sophomores are already learning more or at a faster pace, and that any differences in achievement growth over time reflect only the extrapolation of prior learning patterns. It would be useful if future studies would provide for periodic retesting at the start as well as at the end of school years, so that patterns of learning during the summer might be measured in an effort to compare public and private school students when school influences are presumably less salient and certainly less immediate.

Coleman, Hoffer, and Kilgore bolster their conclusion that initial selectivity does not explain achievement differentials, by showing that measures of student behavior and academic demands can explain achievement differences both among schools within the same sector and across sectors. It is this finding which Coleman has stressed in his public discussion of the report’s implications. But the finding contains a chicken-and-egg dilemma. It raises the question of whether the factors Coleman attributes to school policies are not, in fact, outcomes of unmeasured individual background characteristics or student body composition. Is it reasonable to believe that the relatively low incidence of classes being cut and the greater amount of homework completed in nonpublic schools would be maintained if private schools encountered the same student bodies as are encountered by the public schools? The answer may well be yes, but the data in “Public and Private Schools” cannot demonstrate that this would be true.

Problems in measuring achievement

Like the ill-defined concept “intelligence,” “achievement” exists as no fixed entity nor in natural units. Consequently measuring achievement and its growth is problematic. Tests may not adequately represent objects of interest, and what counts as high levels of or large gains in achievement is not readily obvious. Comparative differences implied by the results of one test can be larger or smaller than those implied by others. On these grounds, some critics have expressed doubt that the tests employed in the High School and Beyond Survey are adequate indices of the differences in achievement among students in public and nonpublic schools.

NCES administered short tests in the areas of reading, vocabulary, and mathematics to both seniors and sophomores. (Tests in other areas, e.g., civics, were given to either seniors or sophomores, but not to both.) Eight items each in the reading and vocabulary tests were common to the senior and sophomore tests, and 18 math items occurred in common. Coleman and his colleagues relied upon these subtests for most of their analyses, and one must wonder how valid and reliable a measure of differences between students in each of the school sectors these few items can provide.

Nor is it obvious that these test items, however adequately they tap what students know or can do, provide a good test of what private and public secondary schools do (or do not do) for their students. The tests may not faithfully reflect what is taught or learned during the secondary school period. Rather, they may reflect learning or nonlearning associated principally, if not exclusively, with the grades prior to high school entrance. Suspicion on this score is prompted by the only description the authors offer of any of the tests: “The mathematics items are all rather elementary, involving basic arithmetic operations, fractions, and only a few hints of algebra and geometry” (p. 159). The “growth” measured by these tests may not correspond well to the growth measured by tests better designed for the task of assessing high school achievement.

To regard the difference between senior and sophomore scores on these particular tests as “two years of achievement,” as the report does, may lead to exaggerated claims about the differences between public and private school
The report's implications for educational inequality

Who benefits from private schools?

Because the student bodies of public and private schools are different, the average differences in achievement among different kinds of schools do not pertain to any particular groups of students. The extent of achievement differences between public and private schools will vary among different kinds of students. To facilitate their discussion, Coleman, Hoffer, and Kilgore pick as their basic point of comparison students with the background characteristics of the average public school sophomore. If we want to predict the benefit an average public school student might realize by transferring to a private school, this is a sensible choice. However, because students in the private schools are disproportionately drawn from advantaged backgrounds, it is difficult to place confidence in this prediction.

Since the vast majority of students from the full spectrum of social backgrounds are in public schools, we are in a better position to ask what a typical private school student would lose by transferring into the public schools than we are in asking Professor Coleman's question. The answer is "not very much." Public school students who are similar to the average private school student are predicted from the report's results to do about as well or better than their private school counterparts. The average private school student appears to gain little or nothing from opting out of the public schools.

If private schools do little to enhance the achievement of advantaged students, how credible is the report's conclusion that less advantaged students may be the principal beneficiaries of Catholic schooling? Noting that among Catholic school students achievement differences between those with college-educated parents and those whose parents only finished high school, between blacks and whites, and between Hispanics and Anglos are smaller than in the public schools, the authors conclude that "the Catholic schools come closer to the American ideal of the 'common school,' educating all alike, than do the public schools" (p. 177). This result also implies that the achievement differences between public and Catholic schools are greatest among students whose parents are less well-educated, and among blacks and Hispanics.

It is tempting to dismiss these findings as the entirely predictable artifacts of a high degree of selectivity governing the entrance of nonwhite or lower socioeconomic students into the Catholic schools. Such students, we might expect, would be atypically successful in their schooling even before entering the Catholic schools. Since half the black students in Catholic schools are Protestants and are therefore more likely to be making a purely educational
Work effort, savings, and income distribution: What are the effects of income transfers?

In recent years, federal government spending for income transfers—social insurance and public assistance—has increased, both in absolute amounts and relative to other public programs. As a result, income transfers have accounted for a growing proportion of total personal income. Poverty has as a consequence declined greatly, income inequality only slightly. Simultaneously, labor force participation rates for men have decreased, and the rate of saving out of disposable income has declined. Some believe that the simultaneous occurrence of these developments is not accidental but causal. They argue that because of the incentives in income support programs, and the taxes required to finance them, work effort is discouraged and savings reduced. In this view, the growth of the income support system has played a significant role in explaining the sluggish performance of the economy, and further expansion would have increasingly negative effects.

Institute researchers Sheldon Danziger, Robert Haveman, and Robert Plotnick have reviewed the existing literature in three particular areas where the nature and magnitude of the effects of income transfer programs are being heavily debated. In their article (see box) they address three main questions: How great are the work disincentives of transfer programs, and how much do they discourage work? To what extent do transfers discourage private saving? What is the magnitude of their effects on poverty and income inequality? The authors begin with labor supply, the area where the impact of transfers has been most fully explored.

Labor supply

Study of the transfer-induced reduction in labor supplied to the marketplace has focused upon the effect of Social Security, because of its universality and its incentives for early withdrawal from the labor force. Other programs, however, have come in for their share of criticism. The four programs discussed in detail by the authors are Old Age and Survivors Insurance, Disability Insurance, Unemployment Insurance, and Aid to Families with Dependent Children. There has been no systematic analysis of the labor supply effects of other transfer programs. Table 1 offers estimates of how much higher total labor supply during the later 1970s would have been if all income transfer benefits were eliminated.

<table>
<thead>
<tr>
<th>Program</th>
<th>Reduction of Work Hours by Transfer Recipients as a Percentage of Total Work Hours of All Workers</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Social Security</strong></td>
<td></td>
</tr>
<tr>
<td>Old Age and Survivors Insurance</td>
<td>1.2%</td>
</tr>
<tr>
<td>Disability Insurance</td>
<td>1.2</td>
</tr>
<tr>
<td>Unemployment Insurance</td>
<td>0.3</td>
</tr>
<tr>
<td>Workers' Compensation and Black Lung</td>
<td>0.7</td>
</tr>
<tr>
<td>Railroad Retirement</td>
<td>*</td>
</tr>
<tr>
<td>Veterans' Disability Compensation</td>
<td>0.4</td>
</tr>
<tr>
<td>Medicare</td>
<td>*</td>
</tr>
<tr>
<td><strong>Public Assistance</strong></td>
<td></td>
</tr>
<tr>
<td>AFDC</td>
<td>0.6</td>
</tr>
<tr>
<td>Supplemental Security Income and Veterans' Pensions</td>
<td>0.1</td>
</tr>
<tr>
<td>Food Stamps and Housing Assistance</td>
<td>0.3</td>
</tr>
<tr>
<td>Medicaid</td>
<td>*</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>4.8</td>
</tr>
</tbody>
</table>

*Under .05%.

Old Age and Survivors Insurance

OASI may induce workers to work less during their prime years, in view of the expected stream of net benefits when they retire; that very availability of benefits, of course, has led to a substantial withdrawal from the labor force among men over 65. The earnings test, which (at least in 1981) reduces benefits for recipients under 72 by 50 percent of earnings in excess of $5000, also might discourage work effort in retirement years. This loss of labor may be offset by the tendency of younger men to work more to offset the earnings test that Social Security will impose upon them when they are older.

Of the dozen empirical studies analyzed by Danziger, Haveman, and Plotnick, all agreed that labor supply was reduced and retirement decisions increased by Social Se-
curity, but no two came up with wholly comparable estimates. Different data sets, different and in some cases faulty methodologies, and a different selection of variables help explain this consequence. Only one study, for instance, considered the joint labor supply decisions of husband and wife, and none explored the advantages of panel data for this kind of research. Despite these weaknesses, it is clear from Table 1 that Social Security is one of the two biggest sources of the estimated reduction in labor supply that can be traced to government transfer programs, accounting for about one-quarter of the total.

Disability Insurance

Another large contributor to the decline in labor supply is DI. Relaxed eligibility determination and higher benefit levels have countered the effects of DI's stringent definition of disablement to expand outlays rapidly. The probability that men aged from about 45-50 to 60 will participate in the labor force has been found to fall significantly as the ratio of potential DI benefits to the wage increases. The dynamics of this process involve more than the provision of benefits to "clearly disabled" individuals, but the size of the labor supply effect cannot be altogether reliably established without better measures of "true disability," expected labor market income, and transfers from other sources.

Unemployment Insurance

Compared to Social Security and DI, this third nationally available social insurance program has a small effect upon labor supply. True, the incidence of unemployment will increase if workers who desire to quit can arrange to collect UI, or if workers more often enter temporary or seasonal work because UI benefits will be available. And the duration of unemployment is also likely to be affected when lost wages are replaced by income from UI. But to offset these effects, some people may enter the work force to qualify for future benefits, or work more hours to raise the benefits to which they are entitled.

The most robust effects of UI that have been found have to do with the duration of unemployment. The best available estimates, the authors believe, establish a one-week loss of employment hours for each ten-percentile-point increase in the replacement rate for lost wages, and for each ten-week extra extension of benefits.

Aid to Families with Dependent Children

The three programs just considered, including the two that have by far the largest estimated impact on labor supply, are all social insurance programs, in some sense "earned" by recipients through earlier contributions and current inability to work. AFDC, in contrast, is a welfare program that has been the focus of particular controversy. In some states, AFDC, augmented by Food Stamps and Medicaid, may provide a larger net income for women with several children than full-time work at the minimum wage—a work disincentive which is compounded by the program's high benefit reduction rate in relation to earned income. The labor supply of women who are heads of households is well known to be much more elastic than that of prime-aged men; it is a rational decision on the part of some mothers to reduce labor supply when confronted with such strong work constraints. Best estimates are that AFDC reduces the labor market effort of the average female family head by roughly 180 hours a year, for a total reduction in labor supplied by all workers of about 0.6 percent.

Estimates of overall reduction

The rest of the figures in Table 1 are based upon evidence from the programs discussed above with some other considerations, and are thus more speculative. The decline in work effort, as a percentage of total hours supplied to the economy, is placed at 4.8 percent.

What is to be concluded from these estimates? First, they are based upon research studies with rather disparate results; hence they must be interpreted cautiously. Second, they exclude any labor supply effects arising from the taxes used to finance these transfers—taxes that may well induce changes in the labor supplied by those not receiving transfers. Third, they are based on a comparison with a situation in which no public transfers exist, but other factors remain the same. This condition does not make allowance for private employer or household transfers that, in the absence of public transfers, might well induce some of the same drop in labor supply. For this reason the figures in the table may overstate the net impact of the transfer system. Bearing in mind that transfer recipients come primarily from the ranks of the less skilled or unskilled, whose wages when they work are low, the total loss in earnings to the economy is probably in the range of 3.5-4.0 percent.

Private saving

Because private saving is, in part, undertaken to smooth out income flows across one's lifetime, and because the Social Security retirement program provides the largest amount of transfers, most analysis has focused upon the effect of OASI upon private saving. There are several mechanisms by which this effect may exert itself, and their influence is not all in the same direction. First, pay-as-you-go Social Security benefits may substitute for, and hence decrease, private saving. The benefits, however, tend to induce early retirement, which may increase the need for savings. Second, because the program shifts income from children (taxpayers) to parents (beneficiaries), parents may increase their savings to maintain a target level of bequests and to offset the taxes their children pay. Finally, if it is true that the more money people
have, the greater the fraction they save during their lives, the equalizing effects of Social Security within generations may reduce private savings. The total effect on private savings is the sum of these effects.

The size of Social Security's impact upon saving has been subject to much recent controversy. In a 1974 time-series study, Martin Feldstein argued that Social Security could have reduced private savings and investment by about 38 percent, with a corresponding drop in GNP of about 15 percent. These results have been heavily criticized, both for the assumptions and for faulty estimating procedures. Reestimations correcting for errors have found small or insignificant effects. Part of the general difficulty with the subject, the authors comment, arises from the weaknesses of time-series models for deriving the estimates of the savings effects of Social Security. But microdata and other cross-sectional studies have also failed to confirm that Social Security has any significant negative influence on private savings, although it probably depresses the level of saving somewhat.

It is clearly premature to draw firm conclusions for Social Security's effects on savings. There are serious uncertainties about the appropriate model to use, and about the appropriate measure of Social Security wealth. Effects on bequests have received little consideration, and there are data problems in distinguishing real savings, as opposed to financial flows. Nor can it be expected that other transfer programs would have an influence even as large as Social Security's, for none is on a pay-as-you-go basis or is tightly tied to the life cycle. In sum, and given the wide variation in scholarly estimates, the authors conclude that the true effect of Social Security on saving must lie "somewhere in the range 0-20%." They are inclined to accept the lower end of the range as most probable.

Redistributive effects

The redistributive studies discussed by the authors utilize three summary indicators to measure the redistributive impact of transfers: the incidence of income poverty, the share of aggregate income going to the poorest fifth of households, and the Gini coefficient. These studies employ a substantially different methodology from those analyzing labor supply or saving responses. The latter rely on the application of multiple regression techniques to cross-sectional or time-series data to estimate behavioral responses to policy-induced price, income, or wealth changes. Redistributive studies, in contrast, rely on more straightforward calculations from aggregate data or microdata, and the range of estimated effects is substantially narrower than the labor supply and saving effects.

Measuring redistributive effects accurately nevertheless presents formidable problems. First, there are the deficiencies and discrepancies in the major data sets available for this work—the Current Population Survey microdata tapes and the Panel Study of Income Dynamics. Second is the issue of what economists call the "counterfactual"—what would the income distribution be if there were no transfer programs?

Most studies duck this last issue, measuring redistribution as the simple difference between a household's income after transfers and its income before any transfers are taken into account. But to the extent that transfer-induced declines in labor supply are concentrated among these with low incomes, this approach underestimates the market (pretransfer) incomes of those at the bottom of the income distribution.

Furthermore, since transfers encourage some low-income households to split into small units—older people or young people choosing to live alone instead of with relatives, for instance—there will be more such low-income households than there would have been without transfers. These factors, among others, can readily lead researchers to the conclusion that "measured" market income is more unequally distributed than "true" market income. These problems are compounded when efforts are made to measure trends over time.

Nevertheless, the studies that are examined are in basic agreement over effects, despite substantial variation—indeed, controversy—over the income unit, income definition, income accounting period, valuation of in-kind benefits, and other aspects. Best estimates indicate that cash and in-kind transfers reduced the percentage of persons living in households with incomes below the poverty line by 53 percent in 1968 and by 78 percent in 1980. They have increased the income share of the poorest fifth in recent years by about 4.6 to 5.9 percentage points, and have reduced the Gini coefficient by about 19 percent. Table 2 summarizes the authors' general conclusions about the multiple effects of income transfers.

<table>
<thead>
<tr>
<th>Labor Supply, Savings, and Redistributive Effects of Income Transfer Programs: A Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effect of Major Income Transfer Programs</td>
</tr>
<tr>
<td>Labor supply</td>
</tr>
<tr>
<td>Private savings</td>
</tr>
<tr>
<td>Income poverty</td>
</tr>
<tr>
<td>Income inequality (Gini coefficient)</td>
</tr>
</tbody>
</table>
A critical fact, only now emerging clearly, is that the impact of government transfers on poverty, though very great, has not increased as quickly as the costs of the programs. Thus it is fair to ask whether an expansion of existing programs along present lines would produce returns in the form of diminished poverty that are commensurate with their cost. In the final section of the paper the authors conclude that if current benefits were expanded, they would not produce sizable additional reductions in poverty, as officially measured, or in inequality—most of the additional payments would go to recipients who have already been raised above the poverty line by transfers. Meanwhile, labor supply and savings costs would not diminish: the aggregate loss in earnings and savings per dollar of additional transfers may well increase.

From their review Danziger, Haveman, and Plotnick venture two suggestions:

- Reform, not continued proportional expansion of current programs, is the route to take. Reforms can be designed to reduce disincentives for work and saving without sacrificing the distributional effects that have been achieved. (For some proposals, see "Work and Welfare: New Directions for Reform," *Focus*, 4:1, 1979.)

- Reductions in or elimination of current programs will undoubtedly increase income poverty, and will achieve only small increases in work effort and savings.

At bottom, the authors emphasize, "the research findings are too varied, too uncertain, themselves too colored with judgment to serve as more than a rough guide to policy choices. Perhaps future methodological developments and improvements in data . . . and estimation techniques . . . can decrease the domain over which value judgments now reign."
The dynamics of poverty

by Elizabeth Evanson

For those who experience it, is poverty a lasting or a temporary condition? What types of individuals and households are more likely to endure either "permanent" or "transitory" poverty, and what programs are more likely to be effective in reducing the two types of need? The answers to these questions have important bearing on the direction of government policy: short-term measures such as income transfers are adequate to relieve temporary distress; long-term measures including structural changes in the labor market as well as investment in education, training, and special services are needed to address persistent poverty.

Efforts to classify poverty as persistent or transitory, and to identify the demographic groups susceptible to either variety, have been clouded by the use of different data sources and varying definitions of poverty, persistent, and transitory (see Table 1). As will be seen below, changing emphases in research (and policy) over the last twenty years have sometimes affected the conclusions that are drawn about the phenomenon of poverty. Recently, two Institute-affiliated researchers working independently of each other have suggested that poverty is more persistent than research of the last decade has implied.

Peter Gottschalk, economist at Bowdoin College and a researcher associated with the Institute, examines earnings mobility among married couples. He focuses particularly on the amount of mobility that is due to random fluctuations in earnings and not steady movement up or down earnings ladders. A surprisingly large share, up to two-thirds, of alterations in earnings appears to represent transitory variations; but when those variations are eliminated, the proportion of permanently low earners among the poor is still larger than previously perceived. Lee Rainwater, sociologist at Harvard and member of the Institute's National Advisory Committee, examines labor earnings plus cash transfers and other income. Counting the poor together with the "near poor"—those somewhat above the poverty line—he finds persistent poverty of considerable magnitude, estimated to be as much as 16 percent of all preretirement families.

These two scholars employ distinct definitions and separate data bases in their analyses, yet their conclusions share common features. This commonality is the more striking in view of the increasing salience, over the 1970s, of the view that poverty in the United States was, in large part, a transitory phenomenon. But that view itself re-

Table 1

<table>
<thead>
<tr>
<th>Author</th>
<th>Poverty Line</th>
<th>Persistently Poor</th>
<th>Transitorily Poor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Morgan (1974)</td>
<td>Relative (income-to-needs ratio; lowest fifth of population)</td>
<td>Income poor all 5 years: 9% of families</td>
<td>Income poor at least 1 of 5 years: 35% of families</td>
</tr>
<tr>
<td>Levy (1977)</td>
<td>Absolute (official definition)</td>
<td>Income poor at least 5 of 7 years: 43% of 1967 poverty population (5% of all individuals)</td>
<td>Income poor 2 of 7 years: 30% of 1967 poverty population (11% of all individuals)</td>
</tr>
<tr>
<td>Coe (1978)</td>
<td>Absolute (official definition)</td>
<td>Income poor all 9 years: 1% of all individuals</td>
<td>Income poor at least 1 of 9 years: 25% of all individuals</td>
</tr>
<tr>
<td>Gottschalk (1980)</td>
<td>Absolute (nontransitory earnings less than 125% min. wage for full-time work)</td>
<td>Earnings poor all 6 survey years: 40% of middle-aged couples who were poor in any 1 year</td>
<td>Earnings poor at least 1 of 6 survey years: 12% of middle-aged couples</td>
</tr>
<tr>
<td>Rainwater (1980)</td>
<td>Relative (income-to-needs ratio; poor = 50% of median, near poor = 50-70%)</td>
<td>Income poor all 10 years: 5% of all individuals (poor or near poor, 11.6%); poor on average all 10 years, 9.4%; (poor or near poor on average, 16.5%)</td>
<td>Income poor or near poor 1-6 out of 10 years: 19% of all individuals</td>
</tr>
<tr>
<td>Hill (1981)</td>
<td>Absolute (official definition)</td>
<td>Income poor all 10 years: 7% of all individuals</td>
<td>Income poor at least 1 of 10 years: 24% of all individuals</td>
</tr>
</tbody>
</table>

Source: Data source is the Panel Study of Income Dynamics for all authors except Gottschalk, who used the National Longitudinal Surveys of Labor Force Participation.

...pects a rejection of a still earlier concept. The importance of Gottschalk's and Rainwater's findings emerges clearly when set against earlier conceptions of the dynamics of poverty.
Contrasting viewpoints: two decades

Henry Aaron and Frank Levy have eloquently described one attitude toward the poor that shaped government policy in the mid-1960s, when efforts to combat poverty and—first things first—to define and measure it were launched. The main feature of that attitude was that poverty was a fixed condition; the operative phrase, borrowed from anthropology and applied with something less than rigor, was “culture of poverty.” Whatever its conceptual drawbacks, the intent of the term was to convey the idea that the poor were trapped in a condition of economic immobility, most of them destined to remain poor from year to year, or even, without the intervention of outside forces such as the government, from generation to generation.

During this period the Social Security Administration developed its measure of poverty, the threshold still officially used. This was the estimated minimum income (cash income including government cash transfers) that would permit families, whatever their size and structure, to purchase the Economy Food Plan of the U.S. Department of Agriculture. This measure merged with the concept of an immobile portion of the population with results described by Levy as a kind of “queue” theory:

The poverty population was thought of as forming a long queue in which the length of the line might change but people always kept their places. Thus, under suitable economic conditions, some individuals could be removed from poverty and the poverty population would be reduced accordingly. But if the number of poor people remained constant, the same individuals would remain in poverty (p. 6).

These early efforts to comprehend the nature of poverty lacked the time-series sources that now exist. Only with the development of such longitudinal data sets as the University of Michigan’s Panel Study of Income Dynamics (PSID) did it become possible to gauge turnover in the poverty population with any degree of accuracy. The replacement of cross-sectional studies with longitudinal analysis permitted examination of more than just turnover: it allowed researchers to look for segments within the poverty population that were more likely to remain poor than others. Then policy could, in theory at least, be aimed at particular groups with particular problems.

In 1974, James Morgan was the first to make use of the PSID for this purpose. He analyzed the initial five years (1967-1972) in terms of income poverty, which he defined, like the official measure, by cash income and transfers as related to a family’s needs. Unlike the official fixed line, however, Morgan’s threshold was a relative one: the poor were those who in any particular year were in the lowest fifth of the population as judged by the income/needs standard. Morgan and his colleagues found that 35 percent of the families in their representative national sample were poor in at least one of the five years. Morgan next asked “Who climbs out?”—who crosses the line that divides poor from nonpoor? His answer was that 11 percent of those who started below the line (set at twice the needs standard) in 1967 had moved above it by 1972, while 9 percent who started above the line dropped below it. In other words, the total percentage of poor changed somewhat, but different individuals made up the new percentage.

When Morgan looked for the particular characteristics of those who succeeded in climbing out, he found that in families with the same head over those years, education and ability were the primary sources of escape from poverty; being younger also helped. Getting married was a way out of poverty, and getting divorced was a way into it. Blacks incurred greater risk of being persistently poor than did whites, and needed more years of schooling than whites to avoid poverty.

In 1977 Levy analyzed seven years of data from the Panel Study sample to determine what had happened to those Americans who in the starting year, 1967, were deemed poor by the official standard. Of those 22.3 million people, 30 percent were only temporarily poor—or as he put it, “poor by mistake”—that year, because they were poor in only one other year through 1973. But 43 percent of those poor in 1967 were “permanently” poor, meaning poor in at least five years out of the seven (note that his measure is not poor every year). The remaining 27 percent “cycled in and out,” spending about equal amounts of time under and over the threshold (p. 13).

Levy wanted to test the hypothesis that poverty status was passed on from one generation to the next, so he compared the 1973 status of young people who in 1967 (then aged 16 to 25) were poor and who subsequently left their families and formed their own households. About 90 percent of the new white households, and 80 percent of the nonwhite ones, had moved above the poverty line in 1973; some were considerably above. Levy concluded that those figures “suggest strongly that poverty status is not something passed mechanically from one generation to the next” (p. 30). He also examined the characteristics of the “permanently” poor and found that if he sorted out poor people by the characteristics of the household head in 1967, a little more than one-third were in households headed by an aged or disabled person and about the same proportion were in households headed by men (nonaged and nondisabled); over one-quarter were in households headed by women.

Levy stressed that the proportion that remained poor was substantial, but by the late 1970s it was the transience rather than the persistence that had caught the interest of researchers and policymakers alike, and which ultimately came to dominate the picture. This view was reinforced
by Richard Coe's 1978 study of the same data source, which found that poverty was less persistent (the "persistently poor" being the 1 percent who were poor every one of the nine years examined) but more pervasive (one-quarter of the nation was poor at least one of the nine years) than annual figures alone would indicate.

Another study from the 1970s that showed mobility rather than stability was conducted by Bradley Schiller. His, like Gottschalk's, was a study of labor market earnings, and he noted a considerable degree of movement by individuals. His focus was not on the poor alone but on almost all groups. Using the Longitudinal Employer Employee Data (LEED) file of the Social Security Administration, Schiller wanted to learn how much movement there was by prime-aged male workers from one segment of the earnings distribution to another over the period 1957-1971. His data excluded those who earned very little or no income (the bottom 10 percent) and were therefore not covered by social security; also excluded by definition were workers in the public sector outside the social security system. Within those limits, he found a great deal of mobility: 70 percent changed segments; the average move spanned one-fifth of the total distance from top to bottom of the distribution. Movement was greater in the middle portions of the distribution, least at the extremes. At the lower end, one-third of the workers remained immobile, indicating that many of the poor stayed poor. Schiller compared black with white earnings mobility and pointed out that the poorer blacks experienced less mobility, but blacks at the upper levels experienced more: "What this means is that black workers have an easier time precariously clinging to the higher earnings positions" (p. 935).

**Views from the 1980s**

Gottschalk's subject, like Schiller's, is earnings mobility, but he asks a different question about different people in a different time: If we look at low earners in the years 1966 to 1975 (which included a recession), to what extent can we determine that yearly changes in their earnings represent random fluctuations rather than steady movement up or down earnings scales? Or, as he puts it, "Are people with low earnings in one year experiencing a transitory drop in earnings, or do they have permanently low earnings?" A particularly important aspect of his research deals with a specific segment of the population—the working poor. For if employed people cannot earn their way out of poverty, labor market strategies to improve earnings capacity may be called for.

The population whose earnings Gottschalk analyzes consists of a sample of nearly 1500 middle-aged couples, 30 to 44 years old in 1966, who are in one of the National Longitudinal Surveys of Labor Force Participation (NLS), directed by Herbert S. Parnes. These couples represent workers during their years of highest earnings. Gottschalk focuses on earnings both of the husband and of the couple, the latter included to capture joint household decisions about allocation of time to market work. In contrast with Schiller, he is interested in mobility among those in the lower part of the earnings distribution—whose incomes fall at or below a threshold equivalent to 125 percent of the 1975 minimum wage for full-time work. (He also did some calculations using the official poverty line.)

To separate permanent from transitory changes in earnings, he fitted a time trend to each person's earnings history. The changes in earnings over the six observation periods in the eleven-year span were then separated into transitory changes, meaning movement around the trend, and nontransitory changes, or movement along the trend. The results, Gottschalk believes, suggest that about two-thirds of the mobility among not only low earners but all across the earnings distribution is due to random fluctuation.

Eliminating those fluctuations, he found that the proportion of permanently low earners was high. Among those couples with low earnings in any one year, 40 percent had low earnings in all six survey years, and 67 percent of those with one low-earnings year were also below the threshold in more than half the years. Actual earnings, which included random variation, showed much less persistence: By that measure only 14 percent of low-earning couples in any one year were under the threshold all six years.

Who are those most likely to be among the permanently low earners? In Gottschalk's sample, nonwhites, the elderly, and residents of rural areas had a significantly higher than average probability of having persistently low nontransitory earnings. His findings document a situation concisely expressed several years earlier by Henry Aaron:

Random events—the business cycle, plant closings, family problems, and, one suspects, interpersonal difficulties on the job—play an important role in the dynamics by which families sink into poverty or rise from it. The problem of poverty is in fact a continuum of problems, ranging from those of households who cannot ever earn as much as the officially designated thresholds, through other families who sometimes earn more but never much more than official thresholds, to a fraction that experiences poverty for a relatively brief time and then emerges from it (pp. 36-37).

The labor market circumstances for many of the poor are clearly adverse, and policy considerations might well profitably address the economic structures that constrain low earners.

(continued on p. 19)
The formal evaluation of social welfare programs has imposed stringent demands upon research methodologies in economics and other social sciences. Often the validity of the performance criteria must be reassessed. We now know, for instance, that simply counting numbers served or dollars spent provides insufficient information. The Institute was closely associated with the development of target effectiveness as a rule-of-thumb measure. More recently, Institute researchers, among others, have promoted another criterion: the level of program participation. In the article that follows, Jennifer Warlick, Research Associate at the Institute, assesses the advantages and pitfalls of this measure.

The decade of the seventies witnessed an intensive evaluation of the nation's public assistance programs, which resulted in two proposals for comprehensive welfare reform and a series of proposals for lesser reforms. The performance criteria used to evaluate the various programs included concepts that had become familiar during the previous decade: horizontal and vertical equity (equal treatment of equals, equivalent treatment for nonequals); target efficiency; adequacy; clarity; simplicity; and the presence of incentives for work, family stability, savings, and sharing of income and wealth within families. The seventies marked the development of a new criterion: the level of program participation.

Government officials, program administrators, and welfare rights advocates know well that the number of people who receive benefits from government programs is less, sometimes dramatically so, than the number of those eligible to participate. The failure to reach all those for whom programs are intended is frequently seen as a flaw in program design or administration. Thus the participation rate has become an indicator of the success of a program.

Historical review

The issue of participation first arose when the Aid to Families with Dependent Children (AFDC) caseload grew explosively, more than doubling across the nation between 1967 and 1971. In a 1973 study, Barbara Boland determined that 55 percent of the increase in the basic family (BF) portion—by far the major portion of AFDC—was due to increasing participation among eligibles, and only 45 percent could be attributed to growth in the number of poor eligible families. The estimation of aggregate participation, measured as the ratio of participating to eligible households among all types of AFDC families, increased from 58 to 84 percent in the years 1967 to 1970, while participation by the female-headed families in the caseload climbed from 63 to 91 percent. Identifying increasing participation as a major cause of growth of caseloads and costs helped to defuse escalating concern that the number of poor was increasing during a period of national prosperity. The discussion of growth also underscored a proposition set forth earlier by Frances Piven and Richard Cloward: program caseloads may be controlled not only by manipulating the size of the eligible population but also by discouraging or encouraging participation.

As the seventies progressed, participation rates of other programs were studied. Maurice MacDonald, an Institute researcher, discovered that nationwide participation in the Food Stamp Program averaged 38 percent in 1975. Accusing program administrators of negligence in some cases and blatant resistance to awarding eligible persons the benefits to which they were entitled in others, welfare rights advocates brought suit against states with low participation rates. The courts ruled in favor of potential participants, ordering that outreach programs be created to inform eligible households about the availability of food stamps and to ensure that the stamps were applied for and received with dispatch.
In 1972 Congress created the Supplemental Security Income (SSI) program, the first federal guaranteed annual cash income program in our nation's welfare history. Because eligibility for the program is limited to the aged, blind, and disabled, the Social Security Administration was given administrative responsibility for the program. Alerted to the problem of low participation rates in other programs, SSA initiated a program to inform the eligible population that benefits would be available beginning January 1, 1974. Despite these outreach efforts, which were nationwide in scope and cost $26 million by 1976, participation during the first two years of operation was disappointingly low—estimated for the aged at 50 percent, for the blind and disabled at 67 percent.

Attention has recently turned again to AFDC. Because Boland's work exerted a major and enduring effect on thinking regarding participation, the Urban Institute was awarded a government contract to review her study. Analysis revealed that the way in which Boland derived the numerator in her measure of participation caused her to overestimate the true participation rate by as much as 15 percentage points. According to the Urban Institute study, there have been two distinct periods of growth in AFDC-BF during the decade 1967-1977: In the first five years, participation rose by 104 percent (from 45 to 92 percent); in the second five years the rate rose marginally, by only 3 percent. Although these estimates differ substantially from Boland’s, which were 13 and 15 percentage points higher than the Urban Institute’s for 1967 and 1970, they do not contradict her basic conclusion that participation increases were responsible for much of the caseload growth between 1967 and 1970. In contrast to Boland’s results, they indicate that participation in AFDC-BF reached saturation in 1973 rather than 1970. The Urban Institute study suggests, but does not prove conclusively, that the rapid growth between 1967 and 1973 can be explained by the combination of active outreach efforts, a favorable social and legal climate for welfare expansion, and increased total benefits due to the rapidly growing availability of food stamps and medical care to AFDC participants.

The Urban Institute study also analyzes growth in the unemployed parent (UP) portion of the AFDC program, where participation rates rose by 167 percent between 1967 and 1977 (from 27 to 72 percent). In contrast to the basic family program, UP has not shown a pattern of smooth growth, although the general movement has been upward. Variations appear to be best explained by fluctuations in the level of unemployment.

**Measuring participation**

The measure that was first developed, and is the one most commonly used, to gauge level of program participation is the ratio of the number of actual “filing units” to the estimated total units eligible to receive program benefits. A filing unit is defined as the relevant residence group (i.e., household) filing a single application. The definition varies by program: for food stamps it is all persons living under a single roof and sharing cooking facilities, regardless of blood relationship; for SSI it is a single individual or married couple, which means that there may be multiple filing units within a single household. To calculate the ratio—called the “caseload” or “population participation rate”—it is necessary to count the number of participating filing units and to identify all eligible filing units.

An accurate count of participants can usually be obtained from microdata surveys—surveys of individual households—which question respondents about receipt of welfare. These counts can be verified by checking against administrative records. Because eligible nonparticipants do not identify themselves, counting the eligible population is much more difficult. Most programs have multiple eligibility criteria, all of which must be satisfied. Researchers approach this problem by matching the characteristics of single filing units with the eligibility criteria within a microsimulation model of eligibility (see “The Modern Miracle of Microsimulation Modeling,” *Focus*, 4:2, 1980). Even so, there is no available benchmark equivalent to administrative records against which to verify estimates so obtained.

The population participation rate indicates what proportion of the targeted population actually receives benefits, but does not differentiate this population by level of need. It therefore ignores differences in economic circumstances among members of the eligible population. In programs such as AFDC and SSI, benefits vary inversely with the amount of nonwelfare income of the filing unit. If the level of participation is low, it is difficult to determine a program's target efficiency on the basis of the population participation rate alone. Is the program reaching those with the greatest or least need? It follows that this rate may shed very little light on the question of the degree to which program costs will rise as participation increases.

To help answer these questions, some researchers have recently begun to calculate a second measure of participation: the ratio of program benefits disbursed to the hypothetical total which would have been distributed had all eligibles participated. As in the case of the population participation rate, the numerator—benefits actually disbursed—of this “expenditure participation rate” is more easily obtained than is the denominator—the hypothetical total. The numerator may be obtained directly from program administrative records, but the denominator requires estimation of the sum of benefits available to non-participating eligible filing units. Once again, employment of a microsimulation model is required.
Unlike the population participation rate, the expenditure participation rate, by indicating the percentage of benefits actually disbursed, provides a measure of the degree to which the economic needs of the targeted population are met. Yet it too is an inadequate predictor of how costs will change with increasing participation. Moreover, the cost implications of rising participation in one program may extend beyond that program if increased participation in the initial program leads to higher enrollment in other programs. For example, SSI beneficiaries are automatically enrolled in Medicaid in 28 states, regardless of the amount of their SSI benefits. It is therefore possible for an individual receiving minimal SSI benefits (e.g., one dollar a month) to receive medical services financed by Medicaid valued in the thousands of dollars. It follows that the increase in Medicaid expenditures resulting from increased participation in SSI could dwarf the corresponding change in SSI outlays. While persons interested in SSI's ability to meet the income needs of the eligible population might focus on the expenditure participation rate, Medicaid officials concerned with caseloads and expenditures would find the potential changes in the SSI population participation rate more relevant. In general, it is safest to consider both rates in conjunction.

Estimates of the population and expenditure participation rates for AFDC and SSI are presented in Figure 1. Comparing these rates within a single program, one is struck by their similarity. In no case is their difference greater than 5 percentage points. Differences across programs are more marked. The population participation rate in SSI is only half that in AFDC-BF and two-thirds that in AFDC-UP. In SSI, the expenditure rate is greater than the population rate, suggesting that the neediest of the eligible participate with greater frequency than those with less need (a point illustrated in Figure 2). This hypothesis is not borne out by the AFDC programs, in which the population rate exceeds the expenditure rate.

Users of participation rates should also be aware that aggregate or programwide participation rates such as those cited above may mask substantial variation across states and/or across different categories of recipients. The national population participation rate for filing units in the AFDC-BF program during 1976 is estimated at 83 percent, but state participation rates range from a high of 95 percent in the District of Columbia to a low of 56 percent in Arizona. The national SSI participation rate during 1975 was 47 percent, but the state rates range from 20 percent (Nebraska) to 77 percent (Louisiana).7

There is also substantial variation in SSI participation rates when eligible filing units are categorized by demographic characteristics. One of every two aged eligible individuals with an eighth-grade education or less participates, whereas only 1 of 5 of those with a college education does so. Rural residents are 40 percent more likely to participate than nonrural residents, and southerners are 60 percent more likely than nonsoutherners. Figure 2 illustrates another kind of variation: participation in relation to levels of benefit entitlement (the amount of benefit for which a recipient is eligible).

Can the rates be accurately measured?

Precise measurement of participation rates is a goal that has not been reached. The need to identify nonparticipating eligibles leads to reliance upon large microdata bases. Several features of these data can lead to imprecise estimates:

<table>
<thead>
<tr>
<th>Program</th>
<th>1976</th>
<th>1977</th>
</tr>
</thead>
<tbody>
<tr>
<td>AFDC-BF</td>
<td>94%</td>
<td>89%</td>
</tr>
<tr>
<td>AFDC-UP</td>
<td>72%</td>
<td>69%</td>
</tr>
<tr>
<td>SSI</td>
<td>47%</td>
<td>52%</td>
</tr>
</tbody>
</table>

Figure 1. Population and Expenditure Participation Rates for Three Programs

Population rate = number of people participating as percentage of estimated total eligible
Expenditure rate = amount spent on benefits as percentage of hypothetical total expenditure if all eligibles participated

Source: AFDC rates from Michel (1980); SSI rates estimated by the author from Survey of Income and Education. Figure by University of Wisconsin Cartographic Laboratory.
- Underreporting of nonpublic assistance income
- Reporting of income and other data for a time period different from program accounting periods
- Absence of detailed information regarding respondents' assets and other economic characteristics relevant to eligibility determination
- Insufficient sample size to support estimates for individual states

The first and third features are more likely to lead to overestimates of the size of the eligible population and hence to artificially low participation rates. The second often produces the opposite result: For example, persons earning their annual income in the first six months of a calendar year and experiencing unemployment for the remaining months may be classified as ineligible on the basis of their annual income when they were actually eligible for benefits from a program with a shorter accounting period. It follows that the size of the eligible population will be underestimated and the measured participation rate will be higher than its true value.

A paradoxical situation results when persons who report receipt of welfare benefits are classified as ineligible by the microsimulation model of eligibility. Such persons are referred to as ineligible participants. Should they be included in the calculation of participation rates? The answer may depend on a researcher's belief about the true eligibility status of such persons. It is known that a significant number of people apply for and receive welfare benefits fraudulently. The researcher may believe that a majority of ineligible participants are fraudulent recipients, and may exclude them from the calculation. If, however, the classification of ineligible participants results from an imprecise microsimulation model, and such participants are in fact eligible, then they should be included in the calculation. Researchers hope to find that ineligible participants are indeed ineligible, because the opposite conclusion raises the possibility that a significant proportion of the nonparticipating eligible population is also being misclassified as ineligible.
The absence of detailed data regarding the nature and amount of a filing unit's asset holdings may also significantly affect estimated participation rates. Because the value of assets is included in determining eligibility for most public assistance programs, the problem cannot be ignored. It is common practice to impute assets to filing units according to demographic characteristics, using the known asset value of persons with similar characteristics, or else on the basis of reported nonemployment income. Unfortunately, estimated participation rates are quite sensitive to the chosen imputation method.

To illustrate this sensitivity, the author calculated the population participation rate for SSI in 1974 using three different imputation methods. The first assumed that interest, rents, and dividends actually reported represented a 6 percent return on the total stock of assets. The second and third methods employed two-step and one-step regression procedures respectively to assign asset values to filing units. The estimated participation rates produced under these three methods varied dramatically, from a low of 42 percent when no asset screen was employed to 71 percent when assets were predicted with the simplest of the regression procedures. Calculations also show that variation in SSI participation rates attributable to including and excluding ineligible participants is substantial, ranging from 6 to 20 percentage points depending upon the method of asset imputation.

The Urban Institute study cited above approached the problem of ineligible participants in a different way: Rather than simply including or excluding them, it adjusted estimated participation rates on the basis of data regarding AFDC case and payment error rates regularly collected by the U.S. Department of Health and Human Services as part of its Quality Control Program. The adjustment lowered the estimated population and expenditure participation rates by averages of 7 and 11 percentage points respectively over the period 1973 to 1977. The upward trend in both rates in that period was not changed, however.

What policy role for participation rates?

The previous discussion has established that reliance on a single participation standard, such as the population participation rate, may be misleading in many policy contexts, that national participation rates mask significant variation in participation by geographic residence and recipient characteristics; and that measured participation rates are highly sensitive to the methodologies used in their estimation. What then is the appropriate policy role of participation rates? How much confidence should be placed in them as measures of program performance? Do they enlighten policy discussions or perhaps misguide them?

In the absence of improved data sources, regular use of participation rates as a measure of administrative performance appears unwarranted. Despite their statistical inadequacies, however, participation rates have to good purpose focused attention on program accessibility and on obstacles to participation, leading to efforts to increase public knowledge of program availability, to simplify complex application forms, to reduce waiting time between application and benefit receipt, and to eliminate demeaning treatment of actual and potential recipients. Moreover, comparative studies of variations in participation rates by demographic groups are useful in identifying the different effects of outreach efforts according to the circumstances of filing units, thus helping to target these efforts more efficiently. To the extent that estimation biases are randomly distributed by state, comparative studies of state participation rates can be quite useful. Exact cardinal rankings are not necessary to determine that some states perform relatively better in this respect than others. And the practices of states with high participation rates may be successfully applied to those with low participation rates. Similar lessons may be learned from examination of variations in participation rates for a particular program through time. Thus, despite the fact that absolute participation rates measured at a single time are suspect as a measure of single program performance, relative studies of participation have enhanced and can continue to enlighten the policymaking process.


7Lynn Ware, "AFDC Basic Eligibles and Program Participation Rates" (memo to David Arnando, Family Assistance Studies, SSA), Public Assistance Data Analysis Laboratory, Social Welfare Research Institute, Boston College, 1980. Estimates of SSI benefits are the author's.
choice when entering the Catholic schools, the results obtained by Coleman and his colleagues would be more persuasive if they were shown to hold even among students of the same religion. But no such difference associated with religion can easily explain the relative success of Hispanics in the Catholic schools. Still, the larger effect of parental educational differences among students in the non-Catholic private schools, where there is a greater heterogeneity of background associated with diversity among schools, argues for the role of selection in explaining the apparent success of private schools with minority and lower socioeconomic students. The possibility remains that curricula and pedagogy in the Catholic schools are actually less stratified and more inclusive in their objectives and direction than they are in other schools, but this possibility should be held in abeyance until Coleman or another researcher demonstrates that the effects of family background are actually smaller within individual Catholic schools than within individual public schools. Coleman’s method of analysis, which compares students to the average for all students in each school sector rather than to the average for a student’s own school, leaves open the possibility that some atypical schools, whose achievement levels exceed the levels expected on the basis of the schools’ socioeconomic or racial composition, account for the findings. If this is the case, the results imply no unusual efficacy in general by Catholic schools for reducing the link between social background and achievement.

Do private schools increase segregation?
One argument against using public monies to subsidize attendance at private schools is that private schools directly or indirectly restrict minority enrollment, and that the movement of whites into nonpublic schools would aggravate racial segregation. Coleman, Hoffer, and Kilgore reject this argument by attempting to show that segregation within the private school sector is substantially lower than that within the public sector. If attendance at private schools was assisted, they suggest, minority enrollments would rise more than white enrollments, for a net reduction in segregation nationwide. I consider the report’s analyses and interpretations on this point so flawed as to be seriously misleading.

To begin with, Coleman and his coauthors treat the United States as if it were a single school district. This means that the “segregation” of whites in Salt Lake City from blacks in Philadelphia counts as much as segregation of whites from blacks within Washington, D.C., or of Montgomery County whites from District of Columbia blacks. The relative scarcity of nonpublic schools in certain of the overwhelmingly white Mountain and Pacific states contributes to the overall differences between levels of segregation in the public and nonpublic sectors. When narrower units are considered, public schools appear considerably less segregated than the report’s figures imply, and when levels of segregation in parochial and public school systems have been compared within the same large cities, the parochial systems have been found in some cities to be the more segregated.

To measure segregation, the authors employ an index Coleman introduced in earlier work. The index expresses the difference between the total percentage black in a population of students and the percentage of black schoolmates characteristic of the average white student in this population as a proportion of the total percentage black. Coleman claims that this measure is “standardized” and that it reflects only the distributions of students to schools by race. This claim is mistaken. Even when the patterns of distribution of students to schools by race is identical in two settings, the Coleman index can differ if the relative proportions of blacks and whites differ between the two settings. The differences between the proportion of students in private schools who are black and the proportion in the public schools who are black in fact explains half the apparent difference in the degree of segregation in the private and public school sectors. Measures of segregation which are not sensitive to this complication show much greater comparability across private and public schools.

Even more disturbing than Coleman’s approach to the measurement of segregation is his assumption that even if there were an appreciable shift of students to the private schools, the internal distribution of students by race among schools would remain unchanged. This is hardly credible. For the country as a whole, it would take the addition of just two all-black high schools of 2,500 students each to the private sector to raise from 13 to 28 the percentage of black students in private schools which are 80 to 100 percent black, a figure 5 percent higher than in the public schools.

The depiction of the private school sector as internally less segregated than the public, and the extrapolation of that picture into the future is far less warranted than the authors of “Public and Private Schools” would suggest.

Who would benefit most from tuition subsidies?
One of the unanswered questions concerning tuition tax credits and educational vouchers is what their effect would be on private school enrollments. Critics fear that disproportionate numbers of middle-class whites would take advantage of the opportunity to leave the public schools. Coleman and his colleagues decline to speculate on the enrollment consequences of current proposals because they are uncertain of the price and supply responses these proposals would prompt. Nevertheless, they attempt to allay concern by determining the enrollment responses which would be produced by an additional $1000 of income for every family with secondary school children.

Coleman, Hoffer, and Kilgore relate the proportion of students attending private schools to student-reported pa-
rental income. They find that private school enrollments increase most rapidly as income increases among Hispanics, and more rapidly for whites with low incomes than for similar blacks, but more rapidly for blacks than for whites among those with high incomes. On the basis of these calculations, the report predicts that the preponderance of white or high-income students among shifters would be smaller than their preponderance among current private school students. From this the authors conclude that such students would not enjoy differential benefits from governmental assistance for private school attendance. The empirical findings are not at all convincing and the interpretive logic is deceptive.

Secondary school students' reports of their parents' incomes are not wholly accurate. Accurate measurement of parental income would most likely increase the estimate of the extent to which private school enrollment rises with additional family income. Taking into account other factors which influence choice of schools would also increase the apparent impact of income increases, and probably more so among whites than blacks. For example, wealthier communities have better public schools and are likely to have lower attendance at private schools than would otherwise be expected. Communities with more minority students are likely to have lower family incomes but greater attendance at private schools than would otherwise be expected, at least among whites. If blacks are limited in their residential choices by discrimination, black neighborhoods will be more economically heterogeneous than would otherwise be the case and family income and local public school characteristics will be more highly associated among whites than blacks. Holding constant public school characteristics is therefore likely to raise the effect of income on private school choice more for whites than for blacks. Finally, close inspection of the relationship between income and private school choice in these data shows that income gains tend to produce larger increases in private school enrollment among families whose incomes are already high than among families with initially low incomes. Taken together, all these factors suggest that white middle-class families would be the group most likely to avail themselves of financial subsidies to attend nonpublic schools.

Even if this were not the case and the report's projections were credible, the authors' logic for assessing who would benefit most from government tuition assistance is specious. By their accounting, if the proportion of white middle-class students among shifters were smaller than the proportion of white middle-class students now in private schools, we are to conclude that minority and low-income students are the chief beneficiaries of the policy! What, of course, should be of interest is whose chances of entering the private schools are raised most and, combining new entrants with current private school students, whose chances of utilizing the assistance are greatest. By these criteria, I have little doubt that the benefits of tax-supported assistance for private school tuition would disproportionately benefit middle- and high-income whites rather than black or low-income families, and would very likely aggravate racial and socioeconomic segregation.

**Conclusion**

In sum, Coleman, Hoffer, and Kilgore's conclusions regarding the unusual effectiveness of nonpublic schools for enhancing scholastic achievement are unconvincing. The representativeness of the nonpublic schools and the appropriateness of the tests are in doubt. Not all strategies available to reduce biases due to initial selectivity were attempted, and more appropriate measures of achievement growth and alternative points of comparison suggest appreciably smaller benefits from attendance at private schools than those reported in "Public and Private Schools." Nevertheless, the argument that more demanding and more orderly schools would show educational benefits appeals to common sense, and the hints in the Coleman report that this may be true should be pursued.

The risks in crediting the report's conclusions about the effect of private schools on segregation are far more serious than entertaining its conclusions about achievement. Coleman and his coauthors have exaggerated the differences in internal segregation between public and private schools, they have most likely underestimated the segregation impact tuition subsidies would have, and they omit the very simple fact that one important reason segregation appears modest in the private sector is that so few minority students are within its embrace. To support tuition subsidies in the expectation, based on this report's "finding," that segregation would not be increased would be myopic.

In a *New York Times* Op-Ed column, June 20, 1981, Coleman asked "Should social research directly address divisive issues of social policy?" I believe that the answer is yes, but that we should do so with James Q. Wilson's recent admonishment in mind: "There is little wrong with intellectuals taking part, along with everyone else, in the process by which issues are defined, assumptions altered, and language supplied. But some of them—university scholars—are supposed to participate under a special obligation—namely, to make clear what they know as opposed to what they wish."
ground, found a modest effect of curriculum placement on eleventh grade achievement (Sequential Tests of Educational Progress) and on twelfth grade Math Preliminary Scholastic Aptitude Tests. They found no significant effect on the Verbal PSAT scores. Alexander et al. do not analyze data available from 19 other schools because of the absence of information on race. They nowhere establish that the data they do utilize is representative, so skepticism of their results is warranted.


7The use of summer learning as a control for student differences is a relatively recent innovation in research on school effects. See especially Barbara Heyns, Summer Learning and the Effects of Schooling (New York: Academic Press, 1978).


9Calculated on the basis of Table 6.2.5 in the report, assuming 90 percent of all students are in public schools, 6.7 percent in Catholic schools, and 3.3 percent in other private schools.

10Table 6.2.5 shows that, adopting Coleman's corrections for missing dropout data, Catholic school seniors get 2 more items correct than public school seniors. I adjusted this for background differences between Catholic school and public school seniors, basing my adjustment on Coleman's reported results of biases due to background calculated without corrections for dropping out.

11The authors do not report standard deviations for the subtests which they analyzed. They report them only for the full tests. To approximate standard deviations for the subtests, I assumed the coefficient of variation (i.e., the ratio of the mean to the standard deviation) was the same for each subtest as it was for the corresponding full test.

12See, for example, Jencks and Brown.


16David James, unpublished reanalyses of the High School and Beyond data. James calculated a segregation index of 0.498 for public schools and 0.285 for private schools. But when he assumed the private schools had the same proportion of black students as the public schools did, the segregation index for private schools rose to 0.391 even though the calculation assumed no change in how students were distributed to schools according to race. In contrast to the large differences between the Coleman segregation indices for private and public schools, James found the Gini index (G) and the index of dissimilarity (D) to be more comparable: G for the public schools was 0.703 compared with 0.627 for the private schools, and D was 0.871 for public schools and 0.812 for private schools. See Zoloth for descriptions of these measures.

17Glen Cain suggests this possibility in an unpublished memorandum to Arthur Goldberger, Madison, Wisconsin, July 1981.

18See Coleman, Kelly, and Moore for evidence of "white flight."

19See Goldberger.

20Hispanics might realize a proportionate benefit because their enrollment in nonpublic schools is close to their percentage among all high school students, and Coleman's results predict a higher responsiveness among Hispanics to income increases than among white Anglos. The higher dropout rate among Hispanics would reduce the benefit families actually received.

families from poverty; the couple's other income kept another 20 percent above the line; contributions by others in the household moved a small proportion, 3.5 percent, across the threshold. The comparable figures for sources moving families above near poverty are 43 percent, 12 percent, 27 percent, and 4 percent. These figures, minus income from assets and transfers, are those that Gottschalk examined; they underscore again the need for anti-poverty strategies directed at earnings. The considerable role for "other income" illustrates both the effectiveness of and the dependence upon government transfers.

In the preretirement group—also Gottschalk's population of interest—Rainwater finds that 11 percent of the total population never rose above the near-poverty level in ten years, and if temporary additions to family income from children or others are disallowed, "it would not be at all unreasonable to expect a persistently poor group of pre-retirement families on the order of 16 percent." Anti-poverty policy in the United States has not focused upon persistent poverty, Rainwater points out. In Britain and Sweden, by contrast, general employment and wage solidarity policies have been joined to the social insurance programs that are directed toward alleviating temporary poverty.

Another recent analysis of Panel Study data, by Martha Hill, economist at the University of Michigan, reasserts the significance of transient income poverty over the decade 1969-1978. In any single year, about 8 percent of the individuals were poor by the official definition; over the ten years, 0.7 percent were poor every year and 24 percent were poor in at least one year. But 40 percent of those poor at any time during the decade were poor only one year, in comparison to the 3 percent of the ever poor who were in poverty all years; the remaining 57 percent fell between the two extremes. Yet Hill too finds that persistent poverty afflicted a substantial number. Over each of two five-year periods (used for analysis to detect structural differences in the decade), 11 to 12 percent who were poor in any one year were poor all five years.

The characteristics that Hill finds strongly associated with permanent poverty include less than full-time employment: 75 percent of the persistently poor had household heads working less than 500 hours per year, yet 63 percent of the temporarily poor (poor one or two years out of ten) were in full-time worker households. (Both sets of figures reinforce Gottschalk's conclusion concerning the inability of many workers to earn their way out of poverty.) Blacks and female-headed households formed much larger percentages, 60 percent in each case, of the persistently poor than of the temporarily poor. Furthermore, "as poverty became more persistent, there were also substantial shifts toward larger proportions with disabled heads, female heads 65 or older, and unmarried female heads with children" (p. 111). Hill therefore suggests that programs to reduce persistent poverty should be directed toward households headed by blacks, women, and those working less than full time, whereas programs to relieve temporary poverty should be aimed more toward full-time workers.

Longitudinal studies have thus opened new possibilities for understanding the nature of temporary and enduring poverty and for determining which demographic groups are at greater risk of one or the other. That understanding can give a more effective direction to policies and programs intended to alleviate distress and make unproductive members of society more productive. All of the researchers cited above stress that their efforts are only first steps toward such understanding. More studies tracking the experience of households and individuals over time are needed to enlarge our comprehension of the complex forces that direct groups into conditions of short-run or long-run hardship.

Selected papers


3Levy used a scale factor, applied to the panel's annual sampling weights, to obtain aggregates corresponding to the national population.
As rising marital instability and out-of-wedlock childbearing put an increasing number of children at risk of living in a single-parent family, the question of financial responsibility for children takes on new urgency. One child in two born today will spend some part of childhood in a single-parent family before age 18. Only 60 percent of absent fathers pay anything at all toward the support of their children. And some 40 percent of families headed by women have incomes below the poverty line. Add haphazard legal and administrative approaches to assigning responsibility for children—approaches that have created gross inequities among families—and the gravity of the situation becomes apparent.

Recognition of this has led researchers at the Institute to design and undertake a major research endeavor that is empirically analyzing the extent and nature of child support arrangements nationwide and, in collaboration with the Wisconsin Department of Health and Social Services, is designing a reformed child support system for the state of Wisconsin (see Focus, 4:1, 1979, for a description). As part of that endeavor, a workshop with participants from universities and from state and federal government agencies was held in Madison on April 22 and 23, 1981. The papers presented at this workshop addressed the larger social, legal, economic, and psychological issues raised by the existing child support system and offered proposals to reform it.

The workshop was organized by Judith Cassetty, Assistant Professor of Social Work, University of Texas at Austin. A list of papers presented, with primary authors, follows.

**Introductory session**


*The Role of the State in Child Support Reform Efforts*, Bernard Stumbras, Administrator, Wisconsin Division of Economic Assistance.

**Session 1: The legal system**

*Child Support Enforcement: Legislative Tasks for the 1980s*, Harry Krause, Professor of Law, University of Illinois. Critiques by H. Robert Hahlo, Professor of Law, University of Toronto; Isabel Marcus, Professor of Law, Government, and Public Affairs, University of Texas at Austin.

**Session 2: Patterns of support**

*Child Support: Who pays what to whom?* Maurice MacDonald, Associate Professor of Family Resources and Consumer Sciences and IRP Research Affiliate, University of Wisconsin-Madison; Annemette Sørensen, IRP Research Affiliate. Critiques by Walter D. Johnson, Director, Family Support Project, Sangamon State University, Springfield, Illinois; Wendy Wolf, National Commission for Employment Policy; Alastair Bissett-Johnson, Professor of Law, Dalhousie University, Halifax, Nova Scotia.

**Session 3: The parental duty to support**

*Fundamental Issues of Child Support*, Martin Levy, Professor of Law, University of Louisville Law School; Jessica Kingsley, J.D., University of Louisville Law School. Critiques by Edward M. Young, Center for Health Services Research, School of Medicine, University of Southern California; John Sampson, Professor of Law, University of Texas at Austin; Sheila Kamerman, Professor of Social Policy and Planning, Columbia University.
Session 4: Financial elements of the parental support issue

Developing Normative Standards for Child Support and Alimony Payments, Isabelle Sawhill, Economist, Urban Institute. Critiques by Barbara Bergmann, Professor of Economics, University of Maryland; Carol Bruch, Professor of Law, University of California, Davis.

Session 5: Nonfinancial issues related to parental support

Bread and Roses: Nonfinancial Issues Related to Fathers' Economic Support of Their Children Following Divorce, Judith Wallerstein, Professor of Psychology, University of California, Berkeley; Dorothy Huntington, Ph.D. Critiques by Martha Cox, Institute for Child and Family Policy, University of North Carolina, Chapel Hill; John Santrock, Professor of Psychology, University of Texas at Dallas.

Session 6: Options for reform

Current Proposals for Reforming Laws and Mechanisms for Collecting Adjudicated Support, Sanford Katz, Professor of Law, Boston College.


Child Support in the 21st Century, David L. Chambers, Professor of Law, University of Michigan.

Recent Institute Publications


Class Structure and Income Determination by Erik O. Wright. 1979. $21.00.


Food, Stamps, and Income Maintenance by Maurice MacDonald. 1977. $16.50 (paper $8.00).


Income, Employment, and Urban Residential Location by Larry Orr. 1975. $12.00.

Integrating Income Maintenance Programs edited by Irene Lurie. 1975. $23.50.


Protecting the Social Service Client: Legal and Structural Controls on Official Discretion by Joel F. Handler. 1979. $13.00 (paper $6.00).

Order Form for Institute BOOKS

Send to: Academic Press, Order Department
111 5th Avenue
New York, N.Y. 10003

Customers ordering from Academic Press, New York—send payment with order and save postage and $.50 handling fee.

Name: ____________________________________________
Address: Number and Street ____________________________________
City_________________ State_________________ Zip________

BOOK TITLE(S) 1.________________________________________
2.________________________________________
3.________________________________________

Payment Enclosed ___ Bankamericard no. _______________________
American Express no. _______________________
Diners Club no. _______________________

Signature: __________________________________________
(all orders subject to credit approval)

Prices subject to change without notice. For book orders add applicable sales tax.

Order Form for Institute DISCUSSION PAPERS AND REPRINTS (free of charge)

Send to: Institute for Research on Poverty
3412 Social Science Building
University of Wisconsin
Madison, Wisconsin 53706

Name: __________________________________________
Address: Number and Street ____________________________________
City_________________ State_________________ Zip________

DISCUSSION PAPER nos. ___________________
REPRINT AND SPECIAL REPORT nos. ___________

Note: DO NOT ORDER DISCUSSION PAPERS OR REPRINTS IF YOU RECEIVE EITHER CATEGORY ON A REGULAR BASIS

Order Form for FOCUS NEWSLETTER (free of charge)

Send to: Institute for Research on Poverty
3412 Social Science Building
University of Wisconsin
Madison, Wisconsin 53706

Name: __________________________________________
Address: Number and Street ____________________________________
City_________________ State_________________ Zip________