The Perils of Provocative Statistics
JAMES P. SCANLAN

Female-headed families comprise a larger proportion of the poor than they did the preceding year; the observer concludes that female-headed families are worse off. Blacks consistently make up the overwhelming majority of freshmen disqualified from participation in intercollegiate athletics by Proposition 48, the academic requirement introduced by the National Collegiate Athletic Association to determine which freshman athletes will be eligible; the observer concludes that the NCAA's policy has a severely disproportionate impact on blacks. The ratio of the black infant-mortality rate to the white infant-mortality rate increases; the observer concludes that the health of black infants has declined.

Seemingly sensible conclusions all, and ones the reader can probably remember at least hearing something about in recent years. It would be only slight overstatement, however, to say that the conclusions are altogether wrong. But since many people (and in particular many social commentators) are poorly grounded in rudimentary mathematics, the conventional understanding of the statistics that bear on these and other important issues is often 180 degrees off the mark.

To see how the conventional understanding is often in error, imagine a country with a population of eight families. Six families, or 75 percent, are poor. Two of those are female-headed families, so in this society 33 percent of poor families are female-headed. Ten years later, poverty is down to 25 percent, or two families. But one female-headed family–one less than before–is still poor. Commentators point to the familiar fact that female-headed families, who once composed 33 percent of the country's poor families, now make up a shocking 50 percent. Although this scenario might appear to be far-fetched, the use of what seems bad news to hide what may really be good news has become common in discussions of social policy.

Consider the poverty debate. Which figure is more meaningful: the proportion of female-headed family members who are poor or the proportion of the poor who are in female-headed families? As the simplified account above illustrates, the first statistic tells us more about the economic plight of female-headed families; but ever since people began speaking of "the feminization of poverty" in the late 1970s, attention has focused on the perceived increase in the latter proportion.

In 1980 the National Advisory Council on Economic Opportunity made a frequently quoted pronouncement: "All other things being equal, if the proportion of the poor who are in female-headed families were to increase at the same rate as it did from 1967 to 1977, they would comprise 100 percent of the poverty population by about the year 2000!" In the ensuing years, the theme was pursued throughout the media. For example, a 1983 column in the *Washington Post* that questioned the Reagan administration's concern for women noted:

> Women have become much more economically vulnerable in the past 30 yeas than is generally understood…. In 1959 only 14.8 percent of [whites] below the poverty line were [in families] headed by women; in 1980, more than a quarter of them were. The figures for black families are even more staggering: 24.4 percent of [blacks] below the poverty line in 1959 were [in families] headed by women, but 58.6 [percent] of them were by 1980.

The preoccupation with these provocative data, however, has led people to ignore certain critical features of the feminization of poverty. Most significant is the fact that, as a rule, the feminization of poverty varies inversely with the amount of poverty. That is, when poverty is decreasing, female-headed families—being those most susceptible to poverty—will comprise a larger proportion of the poor, even as the poverty rate for such families is also declining.

Thus a major reason for the dramatic increase in the feminization of poverty—which actually occurred between 1959 and the mid-1970s—was an unprecedented *reduction* in poverty that included major reductions in the poverty of female-headed families. Between 1959 and 1974, for example, the poverty rate for whites in female-headed families dropped from 40 percent to 28 percent as the overall white poverty rate declined from 18 percent to 9 percent. Though far less poverty-prone than in 1959, members of female-headed had almost doubled their representation among the white poor (from 15 percent to 27 percent), while their representation among the overall white population had grown by only about one-fourth.

Is the feminization of poverty bad?

Decreases in poverty do tend further to "feminize" the poverty that remains. But this does not mean that female-headed families do not share fairly in the reduction of poverty, as some commentators have asserted and as certain features of the data might suggest. Rather, it is in the nature of "normal" or bell-shaped distributions for a group that is poorer on average to comprise a larger proportion of each poorer segment of the population.

In 1979, for example, members of female-headed families comprised 28 percent of the population that fell below one-and -one-fourth times the poverty line, 32 percent of the population that fell below the poverty line itself, and 35 percent of the population that fell below three-quarters of the poverty line. If everyone who had been above three-quarters of the poverty line had been raised above the line, then the proportion of the poor made up of female-headed family members would have risen.

Conversely, an across-the-board increase in poverty would have caused a drop in the proportion of the poor made up of female-headed families. In fact, after a striking and uninterrupted increase between 1959 and 1974, the proportion of the poor who were in female-headed families declined with the economic stagnation of the late 1970s and the substantial rise in poverty that ensued in the early 1980s—even though such families were becoming a larger proportion of the total population. While decreases in poverty since 1983 have somewhat increased the feminization of poverty for both whites and blacks, in 1989 the proportion of the white poor who were in female-headed—27.5 percent—was still only marginally higher than the 27.2 percent of 1974; for blacks, the 59.4 percent of 1989 remained significantly lower than the 1978 high of 61.8 percent. This history has been ignored in most commentary, however, which has presented differences between 1959 and the mid-1980s as if they reflected a continuing trend.

The underlying phenomenon manifests itself in other mathematical relationships that may also misleadingly suggest a change in the relative well being of two groups with different income distributions, such as differences between percentage reductions in poverty rates and changes in ratios of poverty rates. But whenever there is an overall decrease in poverty, the poorer group will tend to have a smaller percentage decrease in its poverty rate than other groups, and the ratio of the poverty rate of the poorer group to that of wealthier groups will increase. In the hypothetical reduction in poverty described above, the poverty rate in female-headed families would fall by 26 percent (from 34.4 percent to 25.2 percent), while the poverty rate for all other people would fall by 36 percent (from 9.6 percent to 6.1 percent); and the ratio of the poverty rate of people in female-headed families to that of other people—which was 3.6 to 1—would increase to 4.1 to 1.

Anyone inclined to believe that these figures actually mean that such families are relatively worse off should merely consider the effect of the change on people who are not poor. The hypothetical drop in the poverty rate for people in female-headed families—from 34.4 percent to 25.2 percent—means that the proportion of people in female-headed families who are not poor would rise from 65.6 percent to 74.8 percent, representing an increase of 14 percent. That increase would be several times greater than the 4 percent increase (90.4 percent to 93.9 percent) in the proportion of other people who are not poor. But it would make no sense to say that female-headed families were moving more rapidly out of poverty, just as it would not make sense to say that female-headed families were disproportionately increasing their representation among the shrinking group that remains poor.

In its 1980 report, the National Advisory Council purported to illustrate the "deepening inequality between men and women" by noting that "in 1967, a woman heading a family was about 3.8 times more likely to be poor than a man heading one, [but by] 1977, after more than a decade of anti-discrimination efforts, she was about 5.7 times more likely to be poor." The direction of that change, however, reflected a benign truth—namely, that the economic circumstances of male and female family heads improved measurable during this period.

The misunderstanding of this matter has also led to unwarranted judgments that the relative status of female-headed families has improved. Census data for 1979, for example, indicate that if poverty had increased such that all people previously below one-and-one-fourth times the poverty line were now in poverty, the poverty rate for female-headed families would have increased by only 23 percent (from 34.4 percent to 42.3 percent), while the poverty rate for other people would have increased 43 percent (from 9.6 percent to 13.7 percent). Thus in 1986 the Democratic staff of the Joint Economic Committee would thoughtlessly conclude that increases in poverty after 1979 had not harmed female-headed families as much as other families.

To be sure, at various times real changes in the relative well-being of female-headed families, which are not simply functions of overall increases or decreases in poverty, do occur. For example, an obvious reason to expect an increasing disparity between the well being of two-parent families and that of female-headed families is the dramatic rise in the employment of married women during the 1970s. In the wake of this rise, comparisons between female-headed families and two-parent families have increasingly become comparisons of one- and two-earner families.

Of course, the feminization of poverty is also much influenced by the increase in the proportion of the population that is in female-headed families (although contrary to much commentary, this increase should not simply be equated with the worsening relative economic status of such families). The increase has had effects that

have varied along lines of race and over time; it accounted for just a third of the initial uninterrupted feminization of white poverty from 1959 to 1974, but about three-fifths of the black increase. Continuing changes in family composition have counteracted somewhat the "defeminizing" tendency of increases in overall white poverty since 1974. Thus these changes have become the predominating factor in the post-1959 increase in the feminization of both black and white poverty.

None of the numerous commentaries that claim real changes in the relative well-being of female-headed families, however, indicates a complete understanding of the underlying functional relationships, much less carries out the complex analysis required to distinguish real changes from apparent ones.

The racial impact of Proposition 48

The misfocus upon the feminization of poverty is by no means the only case in which misunderstandings of properties of normal distributions have led to confusion. Consider the controversy over the NCAA's freshman-eligibility standards. In recent years it has been noted in various places that blacks have accounted for around 90 percent of the entering college freshmen barred from participation in intercollegiate athletics for failing to satisfy the academic requirements of the NCAA's Proposition 48. One year, in fact, the *Atlanta Constitution* reported that at major schools in the South, blacks made up almost 95 percent of the disqualified athletes. Such figures were given considerable attention during debates over the NCAA's January 1989 adoption of Proposition 42, which would have denied scholarships to freshmen failing to meet Proposition 48 standards. The figures were universally taken to mean that the NCAA's policy had a tremendous racial impact, and this belief figured prominently in the NCAA's January 1990 decision to modify Proposition 42 to permit athletes who were ineligible for athletic scholarships to receive need-based scholarships.

Yet the popular understanding was exactly wrong. Since on average whites usually have performed better academically than blacks, academic requirements will tend to affect blacks more than whites. Thus the fact that blacks comprise a disproportionate share of the athletes disqualified by Proposition 48—even allowing for the high black representation among college athletes—does indicate that the NCAA's academic criteria, like other academic criteria, have had some racially disparate effect. But few people seem to realize that the overwhelming black representation among people failing to meet an academic requirement is usually a sign that the disparate effect is *small*, not large.

If whites outperform blacks on average, and if the distributions of grades and test scores for each group are roughly normal, then we would expect the proportion that blacks comprise of people at each performance level to rise as the level drops. Say, for instance, that the average score on a test is 80 for whites and 70 for blacks; since the center of the black distribution is lower than the center of the white distribution, blacks will make up a larger proportion of the people who score below 70 than they do of those scoring below 80, a larger proportion of those scoring below 60 than of those scoring below 70, and so on. Consequently, when the cutoff score is set very high, black representation among the people who fail will be only slightly higher than black representation among those taking the test. When the cutoff score is set very low, however, and only people at the left-most tail of the overall distribution are disqualified, the proportion that blacks make up of those disqualified will be much greater.

The standard way of measuring the racial impact of a test or other selection criterion is to compare the proportion of blacks who pass with the proportion of whites who pass. The higher the cutoff score, the greater will be the disparity between the proportions of whites and blacks who pass the test. On the other hand, the lower the cutoff score, the closer the black pass rate will be to the white pass rate, even as blacks come to comprise a higher proportion of people who do not pass. This is why judges sometimes require that cutoff scores be lowered on particular tests that harm blacks unfairly.

Suppose, for example, that 800 whites and 200 blacks take an exam and that the white mean score is one standard deviation above the black mean score. (The term "standard deviation" refers to the departure of scores from the mean; when a distribution is normal, roughly 95 percent of the tested population will fall within two standard deviations of the mean.) Assuming that both distributions are normal, setting the cutoff at the white mean score will yield a white pass rate of 50 percent and a black pass rate of 16 percent—or 32 percent of the white rate; 30 percent (168 of 568) of those who fail and 7 percent (32 of 432) of those who pass will be black. But suppose the cutoff score is lowered to two standard deviations below the white mean. With such a cutoff, the white pass rate will be 98 percent and the black pass rate will be 84 percent; blacks will comprise 67 percent (32 of 48) of those who fail and 18 percent (168 of 952) of those who pass.[1]

---

[1] Because the size of the disparity will vary depending on whether one focuses on pass rates or failure rates (or the presence or absence of a condition), statisticians have sometimes presented disparities in terms of "odds ratios." For each group the odds of passing a test is the groups pass rate divided by its failure rate. In the situation described in the text, when the cutoff is set at the white mean, the white odds of passing would be 1 (the 50-percent white pass rate divided by the 50-percent white failure rate); the black odds would be .19 (the 16-percent black pass rate divided by the 84-percent black failure rate). The ratio of white to black odds of passing would therefore be 5.3 (1/.19), which is the same as the ratio of black to white odds of failing. While the use of odds ratios keeps the size of a disparity from turning on whether one focuses on pass rates or failure rates, the size of the odds

The state of Georgia's teacher-competency testing illustrates the same tendencies. Teachers who had repeatedly failed the teacher's exam were terminated after the ninth administration of the test in August 1987. By that time, close to 95 percent of black teachers had passed, while the white pass rate was above 99 percent. The ratio of the black to white pass rates—around 96 percent—would thus not come close to violating the "four-fifths rule" that is usually used by the federal government to trigger inquiry into whether a test actually measures the ability to perform a job. Yet the statistic that received the greatest attention was that blacks made up 75 percent (244 of 327) of the small group of people who had not yet passed the exam.

In the case of Proposition 48 as well as in that of teacher-competency testing, the emphasis on the high proportion of blacks among the people who are disqualified poses a perplexing dilemma for the policymaker. An obvious and easy means of reducing that proportion is to raise the standards. Doing so, however, will increase the programs' racial impacts as they are normally measured and will increase the absolute numbers of black athletes and teachers who will be disqualified. It is doubtful that either end is actually desired by those most concerned about the potential for such programs to limit black opportunity.

Black and white infant-mortality rates

A number of recent works on social policy have stressed the fact that in 1983 the ratio of the black infant-mortality rate to the white infant-mortality rate reached 1.98, its highest point in the forty years for which such data had been kept. The increase in this ratio has generally been taken to mean that the relative health of black Americans has materially worsened and that government support for health care has declined. Thus one pair of authors called it "a shocking indicator of the combined effects of economic recession and service cutbacks."

It is no such thing. As with the feminization of poverty, the ratio of the black to the white infant mortality rate will almost invariably increase when there is a general decline in infant mortality. That is, the more we reduce the danger of infant mortality, the more infant mortality will be restricted to those most susceptible to it, and each progressively more susceptible group will be even more disproportionately black.

In 1983, when the ratio of black to white infant mortality reached its highest point ever, the infant-mortality rates for both whites and blacks reached their lowest points ever.

---

ratio, like simple ratios of pass or fail rates, is a function of the location of the cutoff and the prevalence of the condition. In the situation described in the text, for example, when the cutoff is set at the 90th percentile for whites the odds ratio will be 11.1; it will drop to 4.9 when the cutoff is set at the 70th percentile of white scores; and it will rise to 9.3 when the cutoff is reduced to the point at which 98 percent of whites pass.

This fact, however, was ignored in the readiness to believe that the increase in the black-white ratio meant that something bad had happened. Similarly ignored were the facts that both the black and white infant-mortality rates had declined every year for the preceding two decades, and that in nine of the past ten years those declines had been accompanied by an increase in the ratio of black to white infant mortality.

This is not to deny the significance of the substantial, and seemingly persistent, disparity between black and white infant mortality. But again, if we want to understand that disparity and whether it is increasing or diminishing, we must examine the data with a greater appreciation of basic mathematics than has generally been exhibited in the commentary on these issues.

These observations about racial differences in infant mortality of course, apply as well to efforts to compare black and white mortality rates from other causes. In recent years there have been many such efforts, with commentators typically regarding an increase in the ratio of black to white mortality rates as indicating a decline in the relative health of American blacks, in general as well as with respect to particular diseases. Examination of the underlying data for some of these diseases reveals real cause for concern about the relative health of blacks; in some cases, black rates have actually risen while white rates have declined. With other diseases, however, the increase in ratios is simply what one would expect to accompany a general reduction of mortality from the disease, and it is where a disease has been almost entirely eliminated that the black-white ratio often will be the highest. When society sets about ordering its health-care priorities—and particularly when it orders them in hopes of reducing racial disparities in mortality from various diseases—a thoughtless emphasis upon increasing black-white mortality ratios could lead to grave error.

Other racial disparities

The mathematical principle described above applies to numerous other situations in which commonly cited data are misinterpreted. Thus when poverty—including black poverty—declines, blacks usually will comprise a higher proportion of the poor. There also will be an increasing disparity between black and white poverty rates, which can occur even when blacks have actually improved their relative condition. By contrast, when poverty increases, the disparity typically will decrease. In fact, one of the most dramatic changes in the ratio of black to white poverty rates occurred between 1978 and 1983, when the ratio dropped from 3.5 to 3.0 as both rates jumped substantially. The decline in the ratio obviously did not mean that black Americans were better off. Nevertheless, a number of recent works have attempted to appraise black progress in terms of changing disparities between black and white poverty rates, without showing any understanding of how

these disparities are functions of the overall amount of poverty.

Similarly, whenever unemployment—including black unemployment—decreases, one would expect to find increasing disparities between black and white unemployment rates. An October 1988 study by the Center on Budget Priorities qualified the good news that black unemployment had reached its lowest point in the decade, by noting that the ratio of black to white unemployment had reached its highest point in the decade. But the conjunction of these two occurrences is to be expected, and to learn whether disparities between the susceptibility of blacks to unemployment and that of whites are actually increasing will require a deeper inquiry.

It has recently been observed that for all the attention to out-of-wedlock births among blacks, the rates of such births are actually increasing more among whites. Yet whenever a condition tends to increase among all groups, it tends to increase at greater rates among the group that is least susceptible to it.

The black two-parent family is far closer in median income to the white two-parent family than the black female-headed family is to the white female-headed family. In 1987, for example, the median black income was 77 percent of the median white income for two-parent families, and only 57 percent for female-headed families. Yet the disparity in poverty rates is much greater between black and white two-parent families—where in 1987 the black rate, 12.3 percent, was 2.4 times the white rate of 5.2 percent—than between black and white female-headed families—where the black rate (51.8 percent) was only 1.9 times the white rate (26.7 percent). This is an almost inevitable consequence of the fact that poverty is far less prevalent among two-parent families than among female-headed families.

Similarly, the disparity between the median incomes of two-parent families and female-headed families is much greater for blacks than for whites. But because whites are generally less likely to be poor than blacks, the disparity between the poverty rates for two-parent families and those for female-headed families is much greater among whites than among blacks.

Finally, when minorities who are hired by an employer or admitted to an educational institution are found on average to be poorer performers than whites, it is sometimes perceived as evidence of affirmative action. The employer or institution is believed to have lowered its standards for minorities. But if the minorities are on average less qualified among the applicants, they will be less qualified among those hired or admitted when the selection procedures are fair—and usually even when the procedures are biased against them—simply because they will be

disproportionately represented among those who were hired or selected from among the less qualified applicants.[2]

On the other hand, that blacks tend on average to be superior performers in professional sports—for example, they have substantially higher batting averages at every position in major-league baseball—has been interpreted by a number of commentators as a sign that blacks are discriminated against even in professional sports. Yet if blacks are on average better athletes among people seeking to become professional athletes (as may be suggested by their high representation in professional sports), on average they will be better athletes among people who become professional athletes. So blacks may be discriminated against in professional athletics, but the data cited to prove it are just as consistent with fair treatment.

The same features at work in the situations described above also frequently lead various studies to find evidence of discriminatory treatment that may not really exist; imprecision in the model is often to blame for this. Consider, for example, studies designed to identify discrimination against blacks in lending practices. Even studies that attempt to adjust for socioeconomic status by grouping the people studied into several levels will often fail to adjust adequately, because within each level blacks will be disproportionately represented in the lower reaches. To the extent that socioeconomic status accounts for some part of observed differences in treatment, the failure to account fully for socioeconomic status will lead to a false perception of discriminatory treatment. And even studies that adjust for precise dollar amounts of income would need to take into account the fact that among whites and blacks earning a particular amount, whites will be disproportionately represented among people whose incomes are growing at a faster rate.

Similarly, groups that on average have more experience and education will also tend to have superior experience and education. Consequently, even efforts to study differences between groups that adjust for differences in amounts of education or experience will tend to find discriminatory treatment, so long as the adjustment ignores this qualitative difference.

This is not to say that such studies are invariably flawed. One must determine whether an inadequately specified variable is important enough to make a difference; one cannot do so, of course, without being aware of the bias imputed by a lack of specificity.

Nor do I mean to deny the immense value of statistics. They are indispensable in a complex society. At the same time, commentators often rely on data that they do not fully understand. Unfortunately, there are no invariably reliable rules of thumb for distinguishing meaningful statistics from

---

[2] For an illustration of this point, see Table IV in my article "Illusions of Job Segregation," *The Public Interest*, no. 92 (Fall 1988), p. 60.

misleading ones. Knowledge of the principles explored above is useful; and understanding the basic point that disparities tend to increase as conditions improve is essential. But this hardly prepares one for the range of similarly flawed uses of statistics that abound even in scholarly journals. Perhaps the most valuable thing to cultivate is the simple recognition that even well intentioned commentators may be misled by the superficial or the provocative. Their conclusions may not only be imprecise––sometimes they will simply be wrong.

[Further developments of the points made in the above article may be found on the Measuring Health Disparities page of jpscanlan.com and nuance of the described statistical tendency are discussed on the Scanlan's Rule page of the same site. The more significant of the published articles on the tendency include:

Divining difference. *Chance* 1994;7(4):38-9,48;

Race and mortality. *Society* 2000;37(2):19-35 (reprinted in *Current* 2000 (Feb));

Can we actually measure health disparities? *Chance* 2006:19(2):47-51.]