A large literature spanning some 35 years investigates variation in racial socioeconomic inequality across communities. A widely accepted hypothesis in this tradition holds that this variation is determined in large part by intercommunity variation in race differences in education. Strong positive correlations between educational inequality and socioeconomic inequality have been reported in many studies and interpreted as support for the thesis. Thus measures of educational inequality are now routinely included as independent variables in community-level analyses of racial socioeconomic inequality (either as controls or as variables of intrinsic interest) and their omission is likely to be criticized by reviewers. The present article argues that this practice is flawed and likely to generate misleading results. It then suggests that disaggregated data must be used in order to take account of the impact of race differences accurately in education on racial socioeconomic inequality. Examples of analyses using disaggregated data are presented, and it is shown that their results contradict the results of a conventional analysis based on the same data. The article thus concludes that the findings of many widely cited studies that include measures of aggregate educational inequality in models of racial socioeconomic inequality must be called into question. Additionally, the article suggests that the relevance of the kinds of issues examined here is not limited to the analyses of inequality, but applies broadly to many other research literatures.

Community-Level Analyses of Racial Socioeconomic Inequality

A Cautionary Note

MARK A. FOSSETT
University of Texas at Austin

Variation in racial socioeconomic inequality across communities has been the object of empirical analysis by sociologists for over 35 years. Early studies in this tradition sought to explain variation in racial inequality in terms of community characteristics thought to affect race differences in oppor-

AUTHOR'S NOTE: *I am grateful to C. Gray Swicegood, Omer R. Galle, W. Parker Frisbie, Tom Pullum, Robert P. Althauser, two anonymous reviewers, and the editor for helpful comments on earlier versions of this article. The research*

SOCIOLOGICAL METHODS & RESEARCH, Vol. 16 No. 4, May 1988 454-491

tunities for socioeconomic attainment (e.g., community size, the relative size of the black population, industrial structure). More recent studies have extended these models by incorporating additional independent variables measuring group differences on individual characteristics such as education (e.g., the ratio of black and white median education). This development has been encouraged by the ascendance of the status attainment and human capital perspectives in stratification research. Their emphasis on the role of individual characteristics in the process of socioeconomic attainment has led investigators to hypothesize that intercommunity variation in racial economic inequality may be explained in part by variation in racial education differences. Studies attempting to test this notion by using measures of aggregate educational inequality to predict racial economic inequality have found that they invariably have sizable and significant effects. As a result, the importance of education differences for explaining community-level variation in racial inequality is widely accepted and failure to include measures of educational inequality in models of racial inequality is almost certain to draw critical reactions from reviewers.

The present article argues that the accepted practice of including measures of racial educational inequality in models of racial economic inequality to assess and/or control for the impact of education differences on economic inequality is problematic for at least three reasons. First, the practice ignores race differences in economic returns to education and their role in determining the extent to which differences in levels of education affect economic inequality. Second, it requires the implausible assumption that the correlation between economic inequality and educational inequality reflects only causal association and is completely free of any spurious component. Third, the practice increases the probability of the "partialing fallacy," a problem that results when two or more independent variables share explanatory

reported in this article was supported in part by NSF Grant SES 80-07901 and NICHD Grant HD 16837.

EDITOR'S NOTE: *This article was processed under the editorships of both George W. Bohrnstedt and J. Scott Long.*

power and most of their joint explanatory power is allocated to one independent variable on the basis of nonsubstantive, statistical reasons. As a result of these problems, estimates of the effects of educational inequality on economic inequality are likely to be upwardly biased, and estimates of the effects of other independent variables are likely to be downwardly biased. Thus studies that include measures of educational inequality in analyses of economic inequality are likely to generate misleading results and ill-founded substantive conclusions.

The sections that follow discuss these and related issues in greater detail. The next sections then identify appropriate strategies for taking account of race differences in education in models of racial economic inequality and present empirical analyses that contrast the results of the accepted or conventional practice with the results of analyses that use the more appropriate strategies. The final section discusses the implications of the central findings of the article and notes that the kind of problems identified are not specific to the analysis of racial inequality, but are instead examples of general problems that have significance for a variety of other kinds of comparative studies that attempt to take account of group differences in individual characteristics such as education while investigating group differences on another variable (e.g., fertility, residential location, marriage, and infant mortality).

MODELS OF RACIAL INEQUALITY

Conventional models of intercommunity variation in racial inequality include two kinds of independent variables with very different theoretical underpinnings. The first kind indexes aspects of the social, demographic, or economic macro structures of communities such as industrial composition and percentage black. Their relevance is suggested by *ecological* or *structural* theories that hold that black opportunities for socioeconomic attainment are affected by the macro features of the social system in which they reside. Scholars building on this theoretical tradition have produced a number of influential analyses examining

the relationship between racial inequality and community characteristics such as size, occupational structure, industrial structure, relative size of the black population, location in the South, and strength of the vote for racist political candidates, among others (e.g., Turner, 1951; Blalock, 1956, 1957, 1959; Heer, 1959; Glenn, 1963, 1964, 1966).

The second kind of independent variable common in conventional analyses of racial inequality measures black-white differences on key individual characteristics known to play a role in socioeconomic attainment. Perhaps the most important and widely used independent variable of this type is race differences in education (e.g., the index of dissimilarity, the ratio of mean or median education, and the difference in percentages completing high school). The relevance of this and similar independent variables is suggested by theories of *individual status attainment* (Blau and Duncan, 1967; Becker, 1964; Mincer, 1974) and related studies that highlight the fact that race differences in socioeconomic attainment result in part because blacks are disadvantaged with respect to characteristics such as education (e.g., Siegel, 1965; Lieberson and Fuguitt, 1967; Duncan, 1969; Daymont, 1980). As noted earlier, these variables generally are strong predictors of racial economic inequality and are included as key explanatory variables in most recent comparative studies of racial inequality (e.g., Becker, 1971; Bahr and Gibbs, 1966; Jiobu and Marshall, 1971; Roof, 1972; Hill, 1974; Masters, 1975; Spilerman and Miller, 1976; LaGory and Magnani, 1979; Elgie, 1980; Semyonov and Scott, 1983; Semyonov et al., 1984).[1]

The present article views the hypothesis that intercommunity variation in racial inequality may be caused in part by variation in race differences in characteristics such as education as an important hypothesis. It accepts that empirical analyses of racial inequality that fail to take account of the role of race differences in education are incomplete and less than fully satisfactory. The concern voiced here, however, is that the conventional practice of taking account of education differences by including measures of educational inequality in models of economic inequality is inappropriate and is likely to lead to biased results and incorrect substantive conclusions. The article now turns to discuss the reasons for this concern.

RACE DIFFERENCES IN RETURNS TO EDUCATION

The intuitively appealing practice of regressing measures of income inequality on measures of educational inequality focuses exclusively on race differences in *levels* of education. The flaw in this strategy is that education is linked with income at the individual, not the aggregate, level and the impact of education differences on income inequality thus depends on the nature of race differences in the process of socioeconomic attainment—specifically race differences in returns to education.

Conventional analyses take account of race differences in returns to education by implicitly assuming they are invariant across different communities. This assumption is implied by the interpretation of the regression coefficient for educational inequality as the average increase in economic inequality associated with a unit increase in educational inequality. Unit changes in educational inequality will have the same impact on economic inequality in different communities *only if* race differences in returns to education *do not vary* across communities.

There is no empirical or theoretical basis for the assumption that race differences in returns to education are invariant across communities. Race differences in returns to education are a key aspect of racial stratification systems and the fundamental presumption underlying comparative analyses of racial inequality is that racial stratification systems vary across communities. Thus, until otherwise shown, investigators should assume that race differences in returns to education vary substantially across communities. It follows that the conventional practice of regressing income inequality on educational inequality *cannot* yield accurate estimates of the impact of education differences on income inequality. The predictions resulting from such an analysis will be in error and the errors of prediction can be easily anticipated. The impact of education differences on income inequality will be overestimated in communities where race differences in returns to education are larger than average and underestimated in communities where differences in returns to education are smaller than average.

DECOMPOSING RACE DIFFERENCES IN INCOME

The linkages between race differences in education, race differences in returns to education, and racial income inequality can be discussed more precisely by drawing on the well-developed literature on techniques for decomposing group differences in means on a dependent variable (Kitagawa, 1955; Winsborough and Dickinson, 1971; Althauser and Wigler, 1972; Iams and Thornton, 1975). For purposes of discussion, it is convenient (but not necessary) to assume that income is a linear function of education such that

$$W = a + bX,$$

and

$$B = c + dZ,$$

where W and B represent white and black means for income, respectively, a and c represent white and black regression constants, b and d represent white and black slopes for education, and X and Z represent white and black means for education.

This example is simplistic in the sense that the income attainment process involves many additional variables and may involve nonlinear relationships and/or interactions between variables. However, the simplicity of the example does not limit its relevance. The points developed below readily generalize to situations involving multivariate and/or nonlinear functions. The extension to multivariate situations is straightforward (see Althauser and Wigler, 1972). In the case of functions that cannot be fitted using linear models, "conversion matrices" can be used to map black and white education distributions onto income attainments, and decompositions analogous to those introduced below can be expressed in matrix terminology (Althauser and Wigler, 1972).

The algebra provided below shows that the difference in mean income between whites and blacks (W-B) can be decomposed into the four terms shown in equation 1 (the derivations have been

previously provided in Winsborough and Dickinson, 1971, and Althauser and Wigler, 1972):

$$\begin{aligned} W - B &= (a + bX) - (c + dZ) \\ &= (a - c) + bX - dZ \\ &= (a - c) + (d + (b - d))(Z + (X - Z)) - dZ \\ &= (a - c) + dZ + d(X - Z) + (b - d)Z + (b - d)(X - Z) - dZ \\ &= (a - c) + d(X - Z) + (b - d)Z + (b - d)(X - Z). \end{aligned} \quad [1]$$

Equation 1 has certain important characteristics. The difference terms in the equation (i.e., (W – B), (a – c), (X – Z), and (b – d)) all involve the subtraction of a mean or parameter for blacks from the corresponding mean or parameter for whites. Additionally, the difference terms on the right-hand side of the equation are always weighted by a mean or parameter for blacks or are weighted by another difference term. These characteristics are significant because, as will be seen below, alternative decompositions change this pattern in ways that alter the substantive interpretation of the decomposition.

Equation 1 is useful in the present context because it shows that the impact of racial education differences on race differences in income depends on race differences in returns to education. To develop the point it is helpful to examine the terms that involve education.[2]

The term d(X – Z) represents a *levels of education effect.* It indicates the portion of the difference in income between blacks and whites that can be uniquely attributed to differences in education. That is to say, it registers the portion of the black income deficit that would be eliminated if black education were set equal to white education while race differences in returns to education were held constant.

The term (b – d)Z represents a *returns to education effect.* It indicates the portion of the difference in income between blacks and whites that can be uniquely attributed to differences in returns to education. That is to say, it registers the portion of the black income deficit that would be eliminated if black returns to education were set equal to those observed for whites while race

differences in education were held constant.

The term (b – d)(X – Z) represents a *joint effect* of race differences in levels of education and returns to education. It indicates the portion of the difference in mean income between blacks and whites that is associated with education but *cannot be attributed uniquely* to either the impact of differences in levels of education or the impact of differences in returns to education. Instead, it reflects the portion of the impact of differences in levels of education that is linked with differences in returns to education, or, alternatively, the portion of the impact of differences in returns to education that is linked with differences in levels of education. The "joint" aspect of this term can be seen in the fact that the portion of the black income deficit it embodies can be eliminated *if and only if* black levels of education *and* returns to education are both increased to the levels observed for whites.

The joint effect term in the decomposition presented in equation 1 is significant because it shows that a portion of the income gap between blacks and whites that is associated with race differences in education cannot be interpreted as a "pure" effect of race differences in education because its impact is contingent on the elimination of race differences in returns to education. This is important in the present context because it indicates that the reduction in inequality that would result if only black-white education differences were eliminated would be much smaller than the reduction that would be realized if education differences were eliminated after differences in returns to education were eliminated first.

The joint effect term is also important because it shows that, all else equal, the part of the black-white income gap that can be uniquely attributed to education differences will grow smaller as the severity of racial stratification becomes greater. Thus assuming that race differences in levels of education and returns to education are greater in communities where racial stratification is more severe, the actual impact on inequality of reducing the black-white education gap by a fixed amount (say one year of schooling) will be smaller in communities where the education gap is large, and larger in communities where the education gap is small. This most ironic fact is quite different from what would be

expected based on the conventional approach to taking account of the impact of differences in education on income inequality.

ALTERNATIVE DECOMPOSITIONS

There are several alternative decompositions of black-white income differences that do not include the joint effect term found in equation 1. It is useful to review these alternatives to show that they do not "eliminate" the joint effect term but instead distribute it to other terms, significantly changing the character of the decomposition and its substantive interpretation.

The decomposition shown in equation 2 reflects the approach implicit in Duncan et al. (1968) (Althauser and Wigler, 1972: 113-114). It "eliminates" the joint effect in equation 1 by combining it with the levels of education term. As a result, black-white differences in education (X – Z) are weighted by the white returns to education parameter (b) instead of the black returns to education parameter (d). The new term b(X – Z) thus indicates the portion of the black income deficit that would be eliminated if black education levels were increased to the levels observed for whites *after* black returns to education were first increased to the level observed for whites (i.e., after d is changed to b).

$$W - B = (a - c) + b(X - Z) + (b - d)Z \qquad [2]$$

The interpretation of this decomposition is substantially different from the interpretation of the decomposition given in equation 1. In equation 1 the impact of differences in levels of education is calculated given the *observed* differences in returns to education. In equation 2 it is computed under the hypothetical assumption that there are *no* differences in returns to education. Such a calculation may be useful for exploring the maximum *potential* impact of eliminating differences in education. But it does not alter the fact that the actual impact of eliminating education differences is tied to differences in returns to education.

The decomposition shown in equation 3 reflects an alternative suggested by Althauser and Wigler (1972: 108-109). It "eliminates" the joint effect term in equation 1 by combining it with the

returns to education term. As a result, black-white differences in returns to education (b – d) are weighted by the white level of education (X) instead of the black level of education (Z). The new term (b – d)X thus indicates the portion of the black income deficit that would be eliminated if black rates of return to education were raised to the observed rates for whites *after* black education is first raised to the level observed for whites (i.e., after Z is changed to X).

$$W - B = (a - c) + d(X - Z) + (b - d)X \qquad [3]$$

Again, the interpretation of the decomposition given in equation 3 is quite different from the interpretation of that given in equation 1. In equation 1 the role of differences in returns to education is calculated given the *observed* differences in levels of education. In equation 3 the impact of differences in returns to education is computed under the hypothetical assumption that there are *no* differences in levels of education. While exploring such a decomposition may be useful for answering certain questions, the fact remains that a portion of the income gap between blacks and whites is bound up in the "joint" effect of differences in returns to education and differences in levels of education.

The final alternative considered here corresponds to an approach suggested by Kitagawa (1955) and Coleman and Sorenson (1970) and examined in several treatments (Winsborough and Dickinson, 1971; Althauser and Wigler, 1972; Iams and Thornton, 1975). Shown in equation 4a, it "eliminates" the joint effect term in equation 1 by distributing half of it to the levels of education term and half to the returns to education term. Since this may not be apparent at first glance, equations 4b, 4c, and 4d are provided to show that this is in fact the case.

$$W - B = (a - c) + ((b + d)/2)(X - Z) + (b - d)\ ((X + Z)/2) \qquad [4a]$$

$$= (a - c) + (d + (b - d)/2)(X - Z) + (b - d)(Z + (X - Z)/2) \qquad [4b]$$

$$= (a - c) + d(X - Z) + 1/2(b - d)(X - Z) + (b - d)Z + 1/2(b - d)(X - Z) \qquad [4c]$$

$$= (a - c) + d(X - Z) + (b - d)Z + (b - d)(X - Z) \qquad [4d]$$

In this decomposition black-white differences in levels of education (X – Z) are weighted by the average of black and white returns to education (b + d)/2 rather than the black rate of return (d). The term ((b + d)/2)(X – Z) thus indicates the proportion of the black income deficit that would be eliminated if black education were raised to the level observed for whites *after* the black rate of return to education is first increased to a level midway between the observed rates of return for whites and blacks. Similarly, black-white differences in returns to education (b - d) are weighted by the average of the black and white education means (X + Z)/2 rather than the black level of education (Z). The term (b – d)((X – Z)/2) thus indicates the portion of the black income deficit that would be eliminated if black rates of return to education were raised to the observed rates for whites *after* black education is first raised to a level midway between the education levels observed for blacks and whites.

As in the previous examples, the interpretation of equation 4a is quite different from the interpretation of equation 1. In equation 1 the role of differences in levels of education is calculated given the *observed* differences in returns to education. In equation 4a the impact of differences in levels of education is computed under the hypothetical assumption that differences in returns to education have been reduced by one-half. Some may argue that allocating the joint effect term to the levels of education term and the returns to education term in this manner is a reasonable way to eliminate the inelegant and inconvenient joint effect term. The point stressed here is that this approach introduces an arbitrary assumption that differences in returns to education are reduced by one-half *before* the impact of differences in levels of education is evaluated, and that differences in levels of education are reduced by one-half *before* the impact of differences in returns to education is evaluated.

SUMMARY AND IMPLICATIONS OF THE REVIEW OF DECOMPOSITIONS

Decompositions are useful to the extent that they shed light on a particular question. The decomposition given in equation 1 is

well suited for clarifying the fact that the impact of differences in education on income inequality depends on the extent of differences in returns to education.[3] The alternative decompositions given in equations 2, 3, and 4a differ from the decomposition given in equation 1 because they eliminate its joint effect term by arbitrarily distributing it to other terms in the decomposition. As a result their interpretations are fundamentally different and do not negate or modify in any way the implications of the decomposition provided in equation 1.

Conventional community-level studies ignore the role of race differences in returns to education. Their design presumes that a given level of educational inequality will have the same impact on income inequality in all communities. The decomposition provided in equation 1 shows that, if differences in returns to education vary across communities, the impact on income inequality of a given level of educational inequality will also vary across communities. Since it is reasonable to assume that race differences in returns to education do vary across communities, it is reasonable to conclude that the practice of taking account of the impact of education differences on income inequality by regressing income inequality on educational inequality is subject to error.

SPURIOUSNESS IN THE CORRELATION BETWEEN INEQUALITY IN EDUCATION AND INCOME

The practice of regressing income inequality on educational inequality in conventional analyses requires that the association between educational inequality and income inequality be free of any spurious component. This assumption is implausible. Race inequality in education and income are almost certainly spuriously correlated. As a result, conventional analyses overestimate the impact of educational inequality on income inequality. This section outlines the reasons for expecting inequality in income and education to be spuriously correlated.

Unlike other independent variables commonly used in analyses of income inequality (e.g., percentage black, industrial structure), it is *not* reasonable to assume that educational inequality and income inequality do *not* share common causes. To the contrary, the stratification processes that generate race inequality in education and income are certain to be closely interrelated and few would question the assumption that their community-level determinants overlap to a very large extent. For example, it would not be controversial to assert that both are influenced by variation in the extent and intensity of racially prejudiced attitudes held by whites. Since inequality in education and income share common causes, the effect of inequality in education on inequality in income will contain a spurious component in conventional analyses unless each of three conditions is met: all common determinants of inequality in education and income must be included in the analysis, each must be measured accurately, and the relationships must be specified correctly. None of these should be taken for granted.

First, it is unlikely that all common determinants will be included in the analysis. Many community-level variables that theory suggests should influence both inequality in education and income are not included in the analyses (e.g., the extent to which blacks are organized to pursue social, political, and economic rights). Moreover, as theory is incomplete in this area, it is undoubtedly the case that unknown variables influencing both forms of inequality are omitted from the models. Thus inequality in education and income will be spuriously correlated due to their common dependence on variables not represented in the analysis.

Second, it is unlikely that all common determinants of inequality in education and income will be measured accurately. As a result, their effects on both forms of inequality will not be fully captured and the resulting errors will be correlated. For example, theory suggests that the level of race prejudice of the white population would affect both forms of inequality, but it is not measured directly in any study. Percentage black is often used as a proxy for white racial attitudes but the causal chain for this interpretation is complex (Wilcox and Roof, 1978) and few if any

would argue that it is a perfect measure. Thus inequality in education and income are likely to be spuriously correlated because the errors of prediction (resulting from measurement error in common determinants) are correlated.

Third, it is an open question whether the effects of the common determinants of inequality in education and income are well specified (i.e., nonlinearities and interactions are taken into account if they exist). If they are not, the effects of these variables will not be fully captured and the resulting errors will be correlated. Exploratory research can provide insight into the question of whether the effects of the independent variables are properly specified, but it cannot guarantee this. If the effects of common determinants are not properly specified, inequality in education and income will be spuriously correlated.

In sum, there are many ways for a spurious correlation between educational inequality and income inequality to arise. Two of these (omitted variables and faulty measurement) are virtually certain to be present. Thus it is implausible to assume that educational inequality and income inequality are not spuriously correlated. As a result, the conventional approach overestimates the impact of race differences in education on income inequality because the entire association between the two variables is interpreted as a causal effect of educational inequality on income inequality.

THE PARTIALING FALLACY

Because race inequality in education and income are so intimately related, the inclusion of educational inequality in models of income inequality would tend to create problems even if the correlation between the two were not spurious. Specifically, its inclusion is likely to cause the effects of other independent variables to be underestimated under certain circumstances due to what Gordon (1968) has termed the "partialing fallacy." This fallacy is a by-product of the general problem of multicollinearity and is likely to occur when redundant predictors that have

unequal correlations with the dependent variable are included in the same equation. In discussing this, it is useful to consider a concrete example—the relative size of the black population.

The relative size of the black population, measured as a percentage of the population of the community, has long been theorized to be a determinant of the level of race prejudice and discrimination in local communities (Hawley, 1944; Williams, 1947; Allport, 1954; Blalock, 1967). Accordingly, many investigators have hypothesized that racial inequality along a variety of dimensions of socioeconomic status will be correlated with percentage black and research has repeatedly found percentage black to be positively associated with inequality in education, occupation, and income (Blalock, 1956, 1957; Becker, 1971; Jiobu and Marshall, 1971; Wilcox and Roof, 1978).

Percentage black and educational inequality contain "redundant" information when they are included in the same equation predicting income inequality. That is, they are correlated with each other as well as with income inequality, and much of their capacity to explain income inequality is held in common. It is also the case that the two variables have unequal correlations with income inequality—educational inequality has a stronger correlation with income inequality than does percentage black.[4] These two circumstances establish the preconditions for the "partialing fallacy" wherein much, sometimes most, of the explanatory power that the weaker predictor shares with the stronger predictor is distributed to the stronger predictor. Thus the coefficient for educational inequality is enhanced at the expense of the coefficient for percentage black. The fallacy in this is that there is no theoretical basis for portioning the redundant explanatory power shared by the two variables in this manner (Gordon, 1968). Nevertheless, an unwary investigator may mistakenly conclude that percentage black does not have an effect on income inequality and that educational inequality has an exceptionally large effect.

The partialing fallacy is a by-product of the general problem of collinearity and is troublesome under any circumstances. It is particularly vexing in the present example, however, because

the explanatory power of a spurious predictor (educational inequality) is enhanced at the expense of a predictor that is thought to be a determinant of income inequality (percentage black).

APPROPRIATE STRATEGIES FOR ANALYSIS

Having argued that the conventional practice of including measures of educational inequality in community-level models of income inequality is flawed, the question arises: What is an appropriate strategy for investigating the impact of race differences in education on income inequality in a community-level analysis? The answer offered here is that, since the theoretical model underlying the analysis includes explanatory variables suggested by both individual-level theory (status attainment theory) and aggregate-level theory (ecological theory), an appropriate empirical analysis must take account of each kind of explanatory variable using data appropriate for its level of theoretical specification. That is to say, an appropriate analysis must draw on *both* individual-level data and aggregate data.

There are at least two ways individual and aggregate data can be brought to bear in a single analysis exploring variation in racial income inequality across communities. The first approach is to investigate contextual models of individual income attainment that include race, education, community characteristics, and various interactions that assess the impact of community characteristics on race differences in the process of income attainment.[5] This approach seems well suited for addressing questions concerning how ecological factors affect race differences in the *process* of income attainment.

A second approach is to draw on data that disaggregates income by race and education separately in each community and conduct a conventional community-level analysis using measures of income inequality that have been adjusted—*on a community-by-community basis*—to "remove" statistically the effect of race differences in education. This strategy seems better suited for addressing theories of variation in the outcome of racial income

stratification in different communities. The next section provides an example using this approach.

COMMUNITY-LEVEL ANALYSIS

This section serves several purposes. First, it illustrates how community-level and individual-level data can be combined to investigate racial stratification systems. Second, it presents evidence showing that the correlation between income inequality and educational inequality contains a sizable spurious component. Finally, it provides concrete evidence of how conventional analyses of racial income inequality can yield misleading results when educational inequality is included as an explanatory variable.

The analysis proceeds in two stages. The first stage replicates a conventional community-level analysis of racial income inequality. The second stage performs a reanalysis that uses disaggregated data to take account of the impact of race differences in education on racial income inequality in a more appropriate manner. The analysis presented is *not* offered as a comprehensive test of theories of community-level variation in racial income inequality. That task is beyond the scope of the present study and is left to later research.

REPLICATION OF A CONVENTIONAL ANALYSIS

The analysis begins by replicating Becker's (1971: 123-126) study of income inequality in southern metropolitan areas in 1950. The study is a good choice for replication because it is simple and uncomplicated and because it is one of the first to employ the analytic strategy that is now the norm in community-level analyses of racial income inequality.

Becker's study examined the hypothesis that racial income inequality is a function of the relative size of the black population. To test the hypothesis he first estimated the zero-order correlation between percentage black and income inequality and found a

strong, positive association that suggested support for the hypothesis.[6] Becker then raised the alternative hypothesis that income inequality is a function of black-white education differences. To assess the relationship between percentage black and income inequality while taking account of race differences in education, Becker estimated the partial correlation between percentage black and income inequality, controlling for educational inequality.[7] This proved to be much smaller than the zero-order correlation between percentage black and income inequality and he concluded that percentage black did not have a direct effect on income inequality. He also concluded that educational inequality was the major determinant of variation in income inequality and that the zero-order correlation between percentage black and income inequality was a by-product of percentage black's association with race differences in educational attainment.

Becker's analysis is replicated here using data for 84 southern SMSAs in 1970.[8] The three variables corresponding to those used in Becker's study are income inequality, educational inequality, and percentage black. Income inequality is measured by the percentage difference between the mean income for white and black males ($100 * (W_I - B_I)/W_I$) and educational inequality is measured by the percentage difference between the mean education for white and black males ($100 * (W_E - B_E)/W_E$).[9] Results similar to those reported below were also obtained using alternative measures of inequality such as the index of net difference (Lieberson, 1975), the index of dissimilarity (Duncan and Duncan, 1955), and the percentage difference in median income and education.

Two methodological changes are introduced in the replication. The first is the use of regression analysis rather than correlation analysis to explore the relationships in the data. This reflects improvements in conventional statistical methodology since Becker's analysis. The second change is the use of three control variables in the regression analysis: *manufacturing concentration* measured by the percentage of the SMSA labor force employed in manufacturing, *population growth* measured by the percentage change in population between 1960 and 1970 (using constant 1970

boundaries), and *SMSA size* measured by the natural logarithm of the SMSA population. Controls such as these are routinely used in the current literature. The findings reported below are unchanged when the control variables are omitted from the analysis. Descriptive statistics for the variables included in the replication of Becker's study are presented in Table 1.

Table 2 presents the results of two regression analyses that replicate Becker's study. The results reported here closely reproduce the results reported in his earlier analysis. Percentage black has a statistically significant, positive effect on income inequality in equation 1. But, when educational inequality is added to the analysis in equation 2, it becomes the most powerful predictor of income inequality and the effect of percentage black is reduced to near zero. Becker (1971), Jiobu and Marshall (1971), and others have interpreted similar findings as indicating that income inequality is primarily a function of race differences in education and that percentage black does not have an important direct effect on income inequality but rather influences income inequality via its impact on race differences in education.

REANALYSIS USING DISAGGREGATED DATA

This section investigates variation in income inequality using a more appropriate strategy for controlling for the impact of race differences in education. The key difference in this reanalysis is the introduction of a new measure of income inequality that is adjusted to eliminate the impact of race differences in education on racial income inequality. This alternative measure is the percentage difference between B_I, the observed mean income for black males, and W_I^*, the *hypothetical* mean income for white males that obtains when the white mean is standardized to the *observed* black education distribution by using the proportionate education distribution for blacks to compute a weighted average of the white income means for each education category.[10] The resulting income comparison thus reflects the level of inequality that would result *if* both whites and blacks had the observed black education distribution and each group maintained their *observed*

TABLE 1

Descriptive Statistics and Zero-Order Correlations for Income Inequality, Educational Inequality, and Selected Independent Variables for 84 Southern SMSAs in 1970

			Zero Order Correlations					
	$\bar{X}$	sd	(1)	(2)	(3)	(4)	(5)	(6)
(1) Educational Inequality	21.9	6.9	1.0000					
(2) Income Inequality	49.4	6.8	0.6458	1.0000				
(3) Percent Black	18.5	9.7	0.6252	0.4042	1.0000			
(4) Size (ln Population)	12.5	0.9	-0.0871	0.0208	0.0537	1.0000		
(5) Growth (1960-1970)	17.9	15.6	0.2755	0.1700	0.0567	0.3785	1.0000	
(6) Manufacturing	20.1	8.3	-0.2934	-0.2629	0.0010	0.0190	-0.2202	1.0000

NOTE: Educational Inequality = the percentage difference between white and black mean education for males in 1970; Income Inequality = the percentage difference between white and black mean income for males in 1970; Percent Black = the percentage of the SMSA population classified as black in 1970; Size = the natural logarithm of SMSA population in 1970; Growth = the percentage change in population for the SMSA between 1960 and 1970 using constant 1970 boundaries; and Manufacturing = the percentage of the SMSA labor force employed in manufacturing industries in 1970. More detailed definitions are provided in the text.

pattern of converting education into income. It is termed *net income inequality* to signify that it estimates inequality net of adjustments for race differences in education distribution.[11]

It is important to note that net income inequality is computed separately for each SMSA. Income data disaggregated by sex, race, and education were collected for each SMSA in the analysis. Then the white mean income was standardized to the black education distribution separately for each SMSA by the method of direct standardization.[12] Thus a total of 84 separate standardization analyses were performed with each one being performed using disaggregated income data specific to the SMSA in question.

Significantly, the adjustment for differences in education varies depending on the race differences in returns to education observed in each SMSA. Thus aggregate income inequality is not necessarily reduced by the same amount in cities with similar

TABLE 2
Standardized Regression Coefficients (betas) from Two Regressions of Income Inequality on Selected Independent Variables for 84 Southern SMSAs in 1970

	Equation	
Independent Variables	(1)	(2)
Percent Black	.399*	-.003
Size (ln population)	-.044	.107
Growth 1960-1970	.108	-.068
Manufacturing	-.245*	-.090
Educational Inequality	....	.650*
R Square	.241	.432
Adjusted R Square	.202	.396

NOTE: Variables are defined in the text and in Table 1.
*Probability of chance occurrence is less than 5%.

levels of educational inequality. For example, Washington, D.C., and Galveston, Texas, register identical levels of educational inequality (21.2) and similar levels of aggregate income inequality (49.4 and 52.2, respectively), but have markedly different scores on net income inequality (34.7 and 43.9, respectively). The adjustment is larger in Washington (14.7 points) than Galveston (8.3 points) because black returns to education are higher (relative to white returns) in Washington.

In contrast, the conventional approach implicitly adjusts income inequality by the same amount in both cities. Since it takes account of education differences by regressing aggregate income inequality on educational inequality, the unstandardized regression coefficient for educational inequality (.563) determines the amount by which income inequality is adjusted when race differences in education are set to zero. The level of educational inequality is identical in Washington and Galveston (21.2). Thus the implied adjustment for education differences is also identical (11.9 points) for both cities.

This highlights the difference between the two approaches. The

conventional approach ignores variation in differences in returns to education. The method used here takes them into account.

As a general rule, standardizing white income to black levels of education reduces, but does not eliminate, the average income difference between blacks and whites. The mean for net income inequality is 38.7, which compares with the mean for income inequality of 49.4. Thus eliminating education differences between blacks and whites using standardization techniques reduces income inequality by approximately 10.7 points (or about 21.7%) on average.[13]

Net income inequality is positively correlated with both income inequality ($r = .927$) and educational inequality ($r = .446$). Both of these correlations have important substantive implications. The first indicates that, while standardizing for education differences reduces the *level* of inequality, it does not significantly alter the variation in income inequality across areas. This is important because it suggests that, at the most, less than 15% of the variation in income inequality is explained by variation in race differences in education ($1.0 - 0.927^2 = .141$). This contradicts the conventional interpretation of the strong zero-order correlation between educational inequality and income inequality ($r = .646$), which suggests that more than 40% of the variation in income inequality is explained by variation in educational inequality.

The correlation between educational inequality and net income inequality is important because it provides evidence that educational inequality is spuriously correlated with income inequality. The basis for this conclusion is that, by definition, net income inequality is free of the influence of race differences in education (it reflects inequality between the predicted incomes of whites and blacks with the same level of education) and should not be correlated with educational inequality. At least, *it should not be correlated with educational inequality if the effect of educational inequality reflects only the impact of race differences in education on income inequality.* The fact that it has a powerful effect even after education differences between blacks and whites are eliminated strongly suggests that inequality in income and education are correlated for other reasons. This interpretation is all the

more compelling since, as noted earlier in the article, a spurious correlation between inequality in education and income should be expected due to their common dependence on explanatory variables that either are not included in the analysis or are included but are measured imperfectly.

Table 3 presents the results of an analysis that is conceptually analogous to the analysis reported in equation 2 of Table 2. Both analyses assess the relationship between percentage black and income inequality while taking account of education differences. In equation 2 of Table 2, education differences are taken into account in the conventional manner. In Table 3, education differences are taken into account using the more demanding method of standardizing income comparisons on education separately for each SMSA.

Significantly, the results of the analysis reported in Table 3 are at odds with the results of the conventional analysis reported in Table 2. In Table 3, percentage black has a statistically significant, direct, positive effect on net income inequality. Thus when education differences between blacks and whites are taken into account using disaggregated data, income inequality is found to be more pronounced in communities with relatively large black populations. This contradicts the conventional analysis in Table 2 that takes account of the impact of education differences at the aggregate level and indicates that percentage black has no direct effect on income inequality.

EXPLANATION OF DIFFERENCES

The disparity between the findings presented in Tables 2 and 3 compellingly demonstrates that the methodological issue of whether race differences in education are taken into account using aggregate data or individual-level data has important implications for the results obtained. Two factors are seen as accounting for the differences between the results of the two approaches.

The first is that the impact of education differences on income inequality is inextricably tied to race differences in returns to education and cannot be accurately assessed by using only aggre-

TABLE 3
Standardized Regression Coefficients (betas) from the Regression of Net Income Inequality on Selected Independent Variables for 84 Southern SMSAs in 1970

Percent Black	.344*
Size (ln population)	-.003
Growth 1960-1970	.009
Manufacturing	-.252*
R Square	.181
Adjusted R Square	.139

NOTE: Net Income Inequality = the percentage difference between the white mean income standardized to the black education distribution and the observed black mean income for males in 1970. Other variables are defined in the text and in Table 1.
*Probability of chance occurrence is less than 5%.

gate data on education differences between blacks and whites. The conventional analyses ignore this fact and assume that race differences in education of a fixed magnitude will have the same impact on income inequality in all communities. In contrast, the approach used here explicitly takes account of variation differences in returns to education. Through the technique of standardization, the education differences between blacks and whites are taken into account, and the extent to which this adjustment reduces income inequality varies depending on the *observed* group differences in converting education into income in each SMSA. Thus the method more accurately "removes" or controls for the impact of education differences on income inequality.

The second factor that accounts for the disparity between the two sets of results is the fact that inequality and education and income are spuriously correlated. As a result, the effect of educational inequality on income inequality is inflated by a spurious component and this has two undesirable consequences. It exaggerates the estimated impact of education differentials on income inequality, and it reduces the estimated effect of community characteristics that are correlated with educational inequality.

These consequences are demonstrated in the regression analy-

TABLE 4
Standardized Regression Coefficients (betas) from the Regression of Net Income Inequality on Selected Independent Variables Including Educational Inequality for 84 Southern SMSAs in 1970

Percent Black	.127
Size (ln population)	.073
Growth 1960-1970	-.086
Manufacturing	-.172
Educational Inequality	.345*
R Square	.236
Adjusted R Square	.186

NOTE: Variables are defined in the text, Table 1, and Table 3.
*Probability of chance occurrence is less than 5%.

sis presented in Table 4. This analysis is identical to the analysis reported in Table 3 with the exception that educational inequality is included as an independent variable in the analysis in Table 4. Of course, *the inclusion of this predictor is not justified on theoretical grounds* since the dependent variable has already been adjusted to remove the impact of race differences in education on racial income inequality. The only purpose in performing the analysis is to show the undesirable consequences that result when education differences are taken into account at the aggregate level.

The first consequence is that educational inequality has a strong, statistically significant, positive effect on net income inequality. There is no basis for interpreting this as a causal effect. There is, however, ample reason to interpret it as a spurious effect that exaggerates the role of education differences in determining income inequality.

The second and equally important consequence is that the effect of percentage black on net income inequality is neutralized. This is an example of the partialing fallacy described by Gordon (1968), which occurs when correlated predictors that have unequal correlations with the dependent variable are included in the same equation. The result in the present case is that the effect of percentage black is estimated to be near zero. The reason for

this is that percentage black's explanatory power (r = .342) is distributed to educational inequality, a predictor that shares most of percentage black's ability to predict net income inequality and that has a stronger zero-order correlation with net income inequality (r = .446).

FURTHER COMMENT ON THE PARTIALING FALLACY

The partialing fallacy presents a very formidable problem for conventional analyses that take account of education differences at the aggregate level. Many variables that affect income inequality are also correlated with educational inequality and their ability to predict income inequality is generally held in common with educational inequality's ability to predict income inequality. Thus, since educational inequality is likely to have a stronger correlation with income inequality, the partialing fallacy is likely to cause some (perhaps most) of the effect of each of the other independent variables to be distributed to educational inequality.

There are no simple solutions to this problem. When educational inequality is included in the model, the resulting collinearity complicates the analysis and strong guidance from theory is needed to disentangle the effects of the different independent variables. Otherwise, it is not possible to resolve the issue without drawing on additional information from outside the model (Gordon, 1968).

This provides yet another reason for using disaggregated data to take account of the impact of education differences on income inequality instead of exercising control at the aggregate level. *This approach eliminates educational inequality from the model and in so doing eliminates this source of multicollinearity and reduces the likelihood of the partialing fallacy.*

CONTEXTUAL ANALYSIS

This section of the article provides an example of a contextual analysis of income attainment that addresses the question of

whether percentage black affects racial income inequality when race differences in education are taken into account at the individual level. The example serves two purposes. First, it provides an illustration of an alternative method for taking account of education differences in an appropriate manner while investigating ecological variation in racial income inequality. Second, it provides another comparison point between the conventional community-level approach and analyses that draw on disaggregated data to control for the impact of education differences on income inequality.

Data on income, age, education, and race are taken from Table 138 of the Sixth Count Summary Files of the 1970 U.S. Census (U.S. Bureau of the Census, 1971) and consist of mean income for males aged 25-64 grouped by race, age, education, and SMSA for all 88 southern SMSAs identified in the 1970 Census. The units of analysis are a total of 5,632 income groups: 64 per SMSA reflecting 2 categories of race by 4 categories of age by 8 categories of education. It is significant to note that, since income is grouped by categories of the independent variables (age, education, race, and SMSA), regression analyses using these data yield unbiased parameter estimates of the underlying individual-level relationships (Langbein and Lichtman, 1978: 23-25).

There are at least two reasons for using these data rather than the only viable alternative—the 1970 County Group Public Use Samples (U.S. Bureau of the Census, 1972). First and most important, the data used here are available for all SMSAs and exploit the full range of variation in community characteristics. In contrast, the 1970 PUS identify only SMSAs with populations of 250,000 or more. This eliminates more than half (48) of the southern SMSAs in 1970, restricts the variation in community characteristics in the sample, and adversely affects comparability between the contextual analysis and the community-level analyses presented above.

Second, the Sixth Count Summary Files are based on a 20% sample. The County Group Public Use Sample for 1970 is a 1% sample. The difference in sample size is considerable and has important practical consequences for the efficiency of the estimates of contextual effects on black income since the black

samples for many small SMSAs are not large.

One minor disadvantage associated with using the Sixth Count data is that the variance in income is underestimated since income is grouped. As a result, measures of explained variance are inflated and tests of statistical significance for coefficients are optimistic. Fortunately, these considerations are not important in the present situation because interest is focused on the parameter estimates for the independent variables that are unbiased. Additionally, the sample size is very large so the danger of incorrectly interpreting chance effects as real effects is very slight.

It should be stressed that the purpose of this section is to compare the results of the conventional aggregate analysis with the results from a contextual analysis that takes account of education differences at the individual level. Thus the fact that some variables routinely included in models of individual income attainment (e.g., occupation, parental background, and number of siblings) are not represented in the regression is not important. Their presence or absence does not affect the methodological issue being examined.

OPERATIONALIZATION OF INDEPENDENT VARIABLES

Education is measured in terms of eight categories (see note 9) that are coded 0–7. The resulting scheme closely parallels the scaling of education used in Blau and Duncan (1967) and gives greater weight to key "credentializing" years (such as years 12 and 16) by scaling them as equal in importance to three or four noncredential years of schooling. The square of education is used in the regression because preliminary analysis indicated it was a more accurate predictor of income for both blacks and whites. Substantively, this indicates that college years of schooling have a greater impact on income than high school years, which in turn have a greater impact on income than grade school years. Additional preliminary analyses using Public Use Sample microdata indicated that the coding scheme for education used here predicted income more accurately than education measured in single

years of schooling completed. Thus the level of precision in measuring education is more than adequate for the purposes of predicting income.

Age is measured by the midpoint of the age range for the group (e.g., 30 for groups aged 25-34). The square of age is included along with age in the regressions to capture nonlinearity in the relationship between age and income. More complex codings were considered but were not used since they did not affect the results in any appreciable way.

Race is coded 1 for blacks and 0 for non-Hispanic whites.

Percentage black, manufacturing concentration, community size, and community growth between 1960 and 1970 are all measured in the same manner as in the aggregate analysis described earlier. Their values are appended to the income, education, age, and race information for each income group.

To assure that the intercept is meaningful, the contextual variables are expressed as deviations from values near their means (size is centered on 13, and growth, percentage black, and manufacturing are centered on 20). Similarly, age is expressed in terms of deviation from age 40. The intercept thus indicates the expected income for males age 40 with less than five years of education who reside in an "average" city.

MODEL SPECIFICATION

Race, education, and age are included in the income regression in additive form and interactions between race and education, and race and age are included to capture race differences in the education-income, and age-income relationships. Similarly, the community characteristics are included in the model both in additive form and in interactions with race.

Due to the fact that all independent variables other than race are involved in interactions with race, the additive effects of age, education, and the community characteristics indicate their effect on white income. The interactions of age, education, and the community characteristics with race indicate the extent to which their effect on black income deviates from their effect on white

income. Negative interactions indicate that the white-black income gap is positively associated with the variable (i.e., as the variable increases, the white-black gap widens). Positive interactions indicate the reverse.

Results of the regression analysis are presented in Table 5. Examination of the effects of education and age show that these variables operate in the familiar and expected ways. White income increases with education. It also increases at a decreasing rate with age until about age 50-55 at which point income peaks and begins to decline with further increases in age. Race has a sizable negative effect and black income is approximately 36% lower than white income at the intercept (age 40, low education, average community characteristics). The slope for education is approximately 50% lower for blacks than whites, and black income is lower than white income at all ages.

The additive effects of the community characteristics show that they have appreciable effects on white income. Net of education and age, white income is higher in large cities, higher in growing cities, and lower in cities with greater manufacturing concentration. White income is also higher in cities with larger black populations, a finding consistent with predictions that whites benefit in economic terms when blacks are present in large numbers (Glenn, 1963, 1964; Dollard, 1957; Frisbie and Niedert, 1977).

The interactions between race and the community characteristics indicate that the community characteristics have different effects for blacks than whites and thus have implications for income differences between whites and blacks. In the interests of brevity, only the effect of percentage black is discussed in detail since its effects are most directly relevant.

Percentage black has a negative interaction with race that implies an overall negative effect on black income. Thus, in light of percentage black's positive effect on white income, it obviously has a positive effect on racial inequality. *Significantly, percentage black's implied effect on inequality in the contextual analysis is the same as its effect on net income inequality reported in Table 3.* Thus both of the approaches using disaggregated data to take account of education differences when investigating racial inequal-

TABLE 5
Unstandardized Regression Coefficients (betas) from the Regression of Income in 1969 on Selected Individual and Community Characteristics

Age	164.460
Age Squared	-7.857
Education Squared	231.901
Race	-2365.989
Race * Age	-131.336
Race * Age Squared	4.874
Race * Education	-120.057
Percent Black	17.282
Size	617.881
Growth 1960-1970	10.376
Manufacturing	-6.233
Race * Percent Black	-23.159
Race * Size	-174.916
Race * Growth 1960-1970	-12.766
Race * Manufacturing	9.208
Constant	6388.740

NOTE: Age = age in years; Education = a scale based on completed schooling; Race = a dummy variable for race with 0 = white and 1 = black. Contextual variables are defined as indicated in the notes for Table 1. To aid in the interpretation of the constant and race effects, Age, Percent Black, Size, Growth, and Manufacturing are expressed as deviations from values near their respective means.

ity generate similar substantive findings. Neither approach yields results consistent with the findings of the conventional analysis reported in Table 2, which takes account of education differences at the aggregate level.

DISCUSSION

This article has shown that the conventional strategy of controlling for educational inequality at the aggregate level is flawed.

It yields upwardly biased estimates of the effect of education differences on income inequality and biased estimates of the effects of other independent variables under certain circumstances. This was shown in the analysis sections of the article where the findings of Becker's (1971) conventional investigation of racial income inequality were first replicated and then contradicted by two separate reanalyses that used disaggregated data income data to take account of the impact of education differences on income inequality in a more appropriate manner.

The reversal of Becker's findings is important because the methodology used in Becker's analysis is found in many influential and widely cited early studies (e.g., Bahr and Gibbs, 1966; Jiobu and Marshall, 1971; Roof, 1972; Hill, 1974) and is now the norm in more recent studies (e.g., Masters, 1975; Spilerman and Miller, 1976; LaGory and Magnani, 1979; Elgie, 1980; Semyonov and Scott, 1983; Semyonov et al., 1984). One implication of this article, then, is that the empirical findings of a great number of community-level studies of racial socioeconomic inequality must be called into question. This includes many influential and widely cited studies and the vast majority of recent studies. It is important to stress, however, that it is only the results of the empirical analyses of these studies that are called into question. Their theoretical contributions remain significant and will no doubt continue to guide research in this area.

A second implication of this article is that more demanding analyses will be required to test more fully the theoretical models advanced in this literature. Theories of variation in racial inequality across communities now draw on both theories of individual socioeconomic attainment (which suggest the relevance of characteristics such as education) and theories of ecological or structural determinants of relative socioeconomic opportunities open to blacks (which suggest the relevance of factors such as percentage black). Thus it is necessary to draw on both individual- and community-level data to provide satisfactory tests of the hypothesized causal model. The theories cannot be evaluated using aggregate data alone.

The practical difficulties involved in testing theories of community-level variation in racial income inequality are greatly

increased. The data requirements are severe and the data analysis is more complex. Unfortunately, less troublesome alternatives are not viable. Theories of racial inequality have become more sophisticated by invoking both individual-level and community-level explanatory variables. Consequently, more complex strategies of analysis are required to evaluate these theories.

To date, few studies have combined individual and aggregate data to test these more complex theories of community-level variation in racial inequality. Thus many of the key hypotheses advanced in this literature have yet to be adequately tested. The challenge for future research in this area is to adopt more rigorous methodological approaches to investigating these issues and assess community-level models of racial inequality in an appropriate manner.

The final point to be made is that the problems identified in this article are not limited to studies investigating racial economic inequality. To the contrary, the problems are likely to occur in all efforts to estimate models of aggregate-level dependent variables that include aggregate-level independent variables whose theoretical relevance derives from micro-level relationships. A brief list of illustrative examples includes (a) analyses predicting race differences in infant mortality with race differences in income, education, and residential segregation (Jiobu, 1972), (b) analyses predicting male-female differences in occupational attainment with female education (Abrahamson and Sigelman, 1987; Rogers and Goudy, 1981), (c) analyses predicting race differences in occupation with race differences in age and sex distribution (LaGory and Magnani, 1979), (d) analyses predicting residential segregation between ethnic groups based on group differences in education, occupation, and income (e.g., Jiobu and Marshall, 1971; Roof and Van Valey, 1972; Marshall and Jiobu, 1975; Roof et al., 1976; Hwang and Murdock, 1983; Hwang et al., 1985), and (e) analyses predicting race differences in fertility with race differences in age, marital status, education, female labor force participation, occupation, and income (Marshall and Sinnot, 1971). The common denominator in these examples is that the predictor variable is suggested by the existence or expectation of a relation-

ship between the dependent variable and the independent variable at the individual level (i.e., mortality with income, education, and residence; occupation with education; occupation with age and sex; residential location with education, occupation, and income; and fertility with age, marital status, education, labor force participation, occupation, and income).

The expectation that the group differences measured by these independent variables are important factors affecting the dependent variable may be completely justified. The problem is that the aggregate correlations examined may not accurately capture the relationships. These correlations are likely to be misleading if the individual-level relationship underlying the aggregate relationship varies across areas in some significant way, or if the aggregate correlations contain spurious components. Generally, the possibility that the individual-level relationship varies across areas cannot be ruled out without examining disaggregated data. And, perhaps more important, the probability of spurious associations seems to be relatively high in these situations because it is not unusual for the independent variable and the dependent variable to be related to common causes. In light of the potential distortions that may be introduced in a conventional aggregate analysis in these situations, researchers are well advised to draw on disaggregated data to examine more appropriately the nature of the relationship.

NOTES

1. It is important to emphasize that the logic of taking account of education differences at the aggregate level is grounded in theories of *individual* socioeconomic attainment. Aggregate controls might be valid if their use was dictated by ecological theory, but this is not the case in the studies cited in this article. Instead, controls for education differences are dictated by assumptions (often explicit) about the role of education in the process of individual socioeconomic attainment.

2. The first term of the decomposition, (a - c), is of little direct interest here because it does not involve education in any way. It reflects an additive race effect. The portion of the income difference between blacks and whites associated with this term is unaffected by changes in differences in levels of education or changes in differences in returns to education.

3. The discussion in this section has focused on group differences in mean income but it applies with equal force to all measures of racial income inequality, including ratios of mean or median income, the index of dissimilarity, the index of net difference, and comparisons of proportions above or below selected income levels. Group differences in mean income have been discussed here for convenience and algebraic simplicity alone. Measures that are not easily manipulated algebraically (e.g., the index of dissimilarity) can be decomposed by performing multiple-standardization analyses and then performing appropriate calculations to compute the value of each component. For example, to decompose the index of dissimilarity on income (D) first compute D between the observed income distributions. Then standardize the black education distribution to the white education distribution and compute the education standardized index of dissimilarity on income (D_E*). Next standardize the black returns to education to the white returns to education and compute the resulting returns standardized index of dissimilarity on income (D_R*). The components can then be computed as follows. The unique impact of differences in education (I_E) will be given by $D - D_E$*. The unique impact of differences in returns to education (I_R) will be given by $D - D_R$*. The joint effect of differences in education and differences in returns to education will be given by $D - (I_E + I_R)$.

4. For example, analysis presented later in the article shows that the correlation between educational inequality and income inequality is .646 while the correlation between percentage black and income inequality is .404. In Jiobu and Marshall (1971) these correlations are .850 and .640, respectively, and similar patterns of correlations are reported in many other studies.

5. Examples of these interactions would include two-way interactions between race and community characteristics that would assess the impact of race on the effects of ecological factors on socioeconomic attainments. Parcel (1979) provides an example of a contextual approach that investigates the effects of ecological and individual factors on black and white earnings attainments.

6. Becker used percentage nonwhite to measure percentage black and a ratio of black and white median income to measure income inequality.

7. Becker measured racial education differences by the ratio of black and white median years of schooling completed.

8. Data on income by level of education and race for men aged 18 and above are taken from Tabulation 138 of the Sixth Count Summary Tapes of the 1970 census (U.S. Bureau of the Census, 1971). The South is defined in terms of the standard census regions. SMSAs are excluded from the analysis when the income data for black males aged 18 and above for the SMSA were based on a sample of fewer than 100 individuals.

9. In the Sixth Count data education is reported using eight categories based on years of schooling completed: 0-4 years, 5-7 years, 8 years, 9-11 years, 12 years, 13-15 years, 16 years, and 17 or more years.

10. White mean income is standardized to the black education distribution to maximize the efficiency of the hypothetical mean. Standardized means have greater standard errors than observed means. However, the increase in the standard error of the standardized white mean is much smaller than the increase in the standard error of the standardized black mean because the sample sizes for blacks are much smaller, especially at higher levels of education.

11. One reviewer suggested that net income inequality is a crude measure because it controls for race differences in *quantity* of education but does not take account of race differences in *quality* or *kind* of education. This is true, but it does not affect the central

point of the present article. The existing literature has adopted a certain strategy for controlling for race differences in *quantity* of education. This article discusses whether this method is satisfactory or flawed. The question of whether it is useful to control for race differences in quantity of education without also controlling for quality or kind of education is a separate and distinct issue.

12. The white income distribution is standardized to the black education distribution by the method of direct standardization outlined in Kitagawa (1955), Althauser and Wigler (1972), and standard texts on demographic techniques such as Shryock and Siegel (1976). The analysis was also replicated using the procedure of regression standardization. The results are identical using either standardization technique. The results for direct standardization are presented here because this method can be used to standardize a wider variety of measures of inequality including the index of dissimilarity, the index of net difference, the ratio of median income, and other measures.

13. Note 11 is also relevant here.

REFERENCES

ABRAHAMSON, M. and L. SIGELMAN (1987) "Occupational sex segregation." Amer. Soc. Rev. 52 (October): 588-597.

ALLPORT, G. W. (1954) The Nature of Prejudice. Boston: Beacon.

ALTHAUSER, R. P. and M. WIGLER (1972) "Standardization and component analysis." Soc. Methods & Research 1: 97-135.

BAHR, H. M. and J. P. GIBBS (1966) "Racial differentiation in American metropolitan areas." Social Forces 45: 521-532.

BECKER, G. S. (1964) Human Capital. New York: National Bureau of Economic Research.

BECKER, G. S. (1971) The Economics of Discrimination. Chicago: University of Chicago Press. (First Edition, 1957.)

BLALOCK, H. M., Jr. (1956) "Economic discrimination and Negro increase." Amer. Soc. Rev. 21: 584-588.

BLALOCK, H. M., Jr. (1957) "Percent nonwhite and discrimination in the South." Amer. Soc. Rev. 22: 677-682.

BLALOCK, H. M., Jr. (1959) "Urbanization and discrimination in the South." Social Problems 7: 146-152.

BLALOCK, H. M., Jr. (1967) Toward a Theory of Minority Group Relations. New York: John Wiley.

BLAU, P. M. and O. D. DUNCAN (1967) The American Occupational Structure. New York: John Wiley.

COLEMAN, J. S. and A. B. SORENSON (1970) "Occupational status changes for blacks and non-blacks during the first ten years of occupational experience." Johns Hopkins University Center for Social Organization Report 123.

DAYMONT, T. N. (1980) "Racial equity or racial equality." Demography 17 (November): 379-393.

DOLLARD, J. (1957) Caste and Class in a Southern Town. New York: Doubleday. (First Edition, 1937, Yale University Press.)

DUNCAN, O. D. (1969) "Inheritance of poverty or inheritance of race?" pp. 85-110 in D. P. Moynihan (ed.) On Understanding Poverty. New York: Basic Books.

DUNCAN, O. D. and B. DUNCAN (1955) "A methodological analysis of segregation indexes." Amer. Soc. Rev. 20 (April): 210-217.

DUNCAN, O. D., D. L. FEATHERMAN, and B. DUNCAN (1968) Socioeconomic Background and Occupational Achievement: Extensions of a Basic Model. Washington, DC: Department of Health, Education, and Welfare.

ELGIE, R. A. (1980) "Industrialization and racial inequality within the American South, 1950-1970." Social Sci. Q. 61: 458-472.

FRISBIE, W. P. and L. J. NIEDERT (1977) "Inequality and the relative size of minority populations: a comparative analysis." Amer. J. of Sociology 82 (March): 1007-1030.

GLENN, N. D. (1963) "Occupational benefits to whites from the subordination of Negroes." Amer. Soc. Rev. 28: 443-448.

GLENN, N. D. (1964) "The relative size of the Negro population and Negro occupational status." Social Forces 43: 42-49.

GLENN, N. D. (1966) "White gains from Negro subordination." Social Problems 14: 159-178.

GORDON, R. A. (1968) "Issues in multiple regression." Amer. J. of Sociology 73 (March): 592-616.

HAWLEY, A. H. (1944) "Dispersion versus segregation: apropos of a solution of race problems." Papers of the Michigan Academy of Science, Arts, and Letter 30: 667-674.

HEER, D. M. (1959) "The sentiment of white supremacy: An ecological study." Amer. J. of Sociology 64 (May): 592-598.

HILL, R. C. (1974) "Unionization and racial income inequality in the metropolis." Amer. Soc. Rev. 39: 507-522.

HWANG, S. and S. H. MURDOCK (1983) "Segregation in nonmetropolitan and metropolitan Texas in 1980." Rural Sociology 48 (Winter): 607-623.

HWANG, S., S. H. MURDOCK, B. PARPIA, and R. R. HAMM (1985) "The effects of race and socioeconomic status on residential segregation in Texas, 1970-1980." Social Forces 63 (March): 732-747.

IAMS, H. M. and A. THORNTON (1975) "Decomposition of differences: a cautionary note." Soc. Methods & Research 3: 341-352.

JIOBU, R. and H. H. MARSHALL, Jr. (1971) "Urban structure and the differentiation between blacks and whites." Amer. Soc. Rev. 36: 638-649.

KITAGAWA, E. M. (1955) "Components of a difference between rates." J. of the Amer. Stat. Assn. 50 (December): 1168-1194.

LAGORY, M. and R. J. MAGNANI (1979) "Structural correlates of black-white occupational differentiation: will U.S. regional differences remain." Social Problems 27: 157-169.

LANGBEIN, L. I. and A. J. LICHTMAN (1978) Ecological Inference. Beverly Hills, CA: Sage.

LIEBERSON, S. (1975) "Rank-sum comparisons between groups," in D. Heise (ed.) Sociological Methodology 1976. San Francisco: Jossey-Bass.

LIEBERSON, S. and G. V. FUGUITT (1967) "Negro-white occupational differences in the absence of discrimination." Amer. J. of Sociology 73 (September): 188-200.

MASTERS, S. H. (1975) Black-White Income Differentials: Empirical Studies and Policy Implications. New York: Academic Press.

MARSHALL, H. H., Jr. and R. JIOBU (1975) "Residential segregation in the United States cities: a causal analysis." Social Forces 53 (March): 449-460.

MARSHALL, H. H., Jr. and J. SINNOT (1971) "Urban structure and the black/white fertility differential: examination of the 'assimilationist' model." Social Sci. Q. 52 (December): 588-601.

MINCER, J. S. (1974) Schooling, Experience, and Earnings. New York: National Bureau of Economic Research.

PARCEL, T. L. (1979) "Race, regional labor markets, and earnings." Amer. Soc. Rev. 44 (April): 262-279.

ROGERS, D. L. and W. GOUDY (1981) "Community structure and occupational segregation, 1960 and 1970." Rural Sociology 46 (Summer): 263-281.

ROOF, W. C. (1972) "Residential segregation of blacks and racial inequality in southern cities: toward a causal model." Social Problems 19: 393-407.

ROOF, W. C. and T. L. VAN VALEY (1972) "Residential segregation and social differentiation in American urban areas." Social Forces 51 (September): 87-91.

ROOF, W. C., T. L. VAN VALEY, and D. SPAIN (1976) "Residential segregation in southern cities: 1970." Social Forces 55 (September): 59-71.

SEMYONOV, M., D. R. HOYT, and R. I. SCOTT (1984) "Place, race, and differential occupational opportunities." Demography 21 (May): 259-270.

SEMYONOV, M. and R. I. SCOTT (1983) "Percent black, community characteristics and race-linked occupational differentiation in the rural South." Rural Sociology 48: 240-252.

SHRYOCK, H. S. and J. S. SIEGEL (1976) The Methods and Materials of Demography. New York: Academic Press.

SIEGEL, P. M. (1965) "On the cost of being a Negro." Soc. Inquiry 35: 41-57.

SPILERMAN, S. and R. E. MILLER (1976) Community and Industry Determinants of the Occupational Status of Black Males. Discussion paper #330. Madison: University of Wisconsin, Madison, Institute for Poverty Research.

TURNER, R. H. (1951) "The relative position of the Negro male in the labor force of large American cities." Amer. Soc. Rev. 16: 524-529.

U.S. Bureau of the Census (1971) Technical Documentation of the Sixth Count Population Summary Tapes. Washington, DC: Government Printing Office.

U.S. Bureau of the Census (1972) Public Use Samples of Basic Records from the 1970 Census: Description and Technical Documentation. Washington, DC: Government Printing Office.

WILCOX, J. and W. C. ROOF (1978) "Percent black and black-white status inequality: southern versus nonsouthern patterns." Social Sci. Q. 59: 421-434.

WILLIAMS, R. M., Jr. (1947) "The reduction of intergroup tensions," chap. 3 in Social Science Research Bulletin #57.

WINSBOROUGH, H. H. and P. DICKINSON (1971) "Components of Negro-white income differences," pp. 6-8 in the Social Statistics Section of the Proceedings of the American Statistical Association.

Mark A. Fossett is Research Scientist and Director of Data Services at the Population Research Center, the University of Texas at Austin. He is currently investigating racial and sexual inequality in metropolitan and nonmetropolitan areas, measures of inequality, and the effects of the sex ratio on various social and demographic variables.