An Analysis Of Income Concentration In Small Areas#

LARRY D. SCHROEDER*

Studies relating socio-economic attributes of the population of an area to income inequality of that population are fairly numerous.[1] The current paper is another such effort but with a slightly different focus. Rather than observe populations living within regions, states or even counties, we here concentrate upon persons living within a single major city—Atlanta, Georgia. Income inequality as observed among families living with Census tracts is used as the unit of analysis.

While the use of Census tract data has both positive and negative aspects, as outlined in Section I, its major strength is that it allows one to study the income distributional attributes of predominantly white and predominantly black populations all located within the same area labor market. The empirical results reveal major divergencies in the relationships between socio-economic characteristics and income inequality for these two racial groups.

The analytical basis of the study is outlined in Section II. Section III elaborates on the testable hypotheses including a description of the variables used in the empirical portion of the paper while the fourth section reports on the empirical results. The results suggest a primary divergence in income inequality correlates for the racial subgroups.

I. SMALL AREA INCOME ANALYSIS

In the study by Farbman (1973) there is a review of several arguments in favor of and opposed to the use of small area (e.g., county-level data) in empirical analyses of income distribution. Although these arguments are reasonable, we would suggest that there are two reasons why further disaggregation to census tracts within a single metropolitan area is desirable. Most important is the fact that by concentrating on populations within a single metropolitan area, we standardize for geographical differences in labor market demand. As Farbman (1973: 334) suggests, when using counties as the geographic base, workers living in one county may be working in another county, thus the distribution of income may be ". . . only loosely related to the economic characteristics of that county." By using a single metropolitan area, the influence of such geographically-related demand characteristics is minimized. The

*Associate Professor of Public Administration, Syracuse University.

observed variations in inequality and socio-economic characteristics therefore more adequately reflect micro-level differences in labor market behavior.

A second argument for emphasis on census tract level data is the opportunity such data present in assessing racial differences in the effects of explanatory variables on income concentration. One consistent finding in each of the empirical analysis cited above is the positive relation between percent of the population which is non-white and income concentration. However, for these state or county-based studies, few observations were possible from within areas in which most, if not all, of the population is non-white. The latter possibility is, of course, quite likely within a single city, especially in the South. Given the opportunity afforded by tract data we replicate our analysis for predominantly white and predominantly non-white tracts as well as for all of the tracts taken simultaneously. By comparing results for the two subgroups additional insight is gained regarding labor market behavior and income generation.

It can be argued that any findings based on areas as small as census tracts (with an average population of approximately 4,000) are simply results of residential choice. To the extent that persons tend to live in tracts with others of similar socio-economic backgrounds, income concentration will be lower than for alternative, larger geographic areas. On the other hand, homogeneity of the socio-economic characteristics also means that these data provide a more direct measure of the hypothesized relationships.[2] We now outline the general analytical framework used to generate these hypotheses.

II. ANALYTICAL FRAMEWORK

There are several theoretical foundations for an analysis of income inequality, including stochastic, human capital and "ability" models.[3] While it is not the intent here to develop any new approach, we will briefly suggest how the human capital approach leads to particular hypotheses spelled out in more detail in the following section. In addition we suggest how, when using the tract-level data at hand, additional hypotheses concerning how certain socio-economic characteristics might further influence family income inequality.

Human capital models of income distribution concentrate almost totally upon the determinants of earnings. They illustrate how such factors as education, on-the-job training and other activities such as health improvements, mobility or labor market information constitute investments which, given a rate of return, influence individual-level earnings.

Chiswick and Mincer (1972) and Chiswick (1974) note how the dispersion of earnings as generated under the traditional human capital model is expected to be a function of the levels and variances of schooling and age (as a proxy for experience) as well as the variance in weeks

worked. Additionally, this model indicates that correlations among these variables also contribute to earnings inequality.

Although the levels and variability of human capital variables are posited to affect income inequality, we also recognize the importance of housing location choice in the determination of observed inequality. The above cited model is based on individual income maximizing behavior and an apparently random grouping of individuals. Within small areas such as Census tracts such randomness is unlikely to occur if persons choose to live in areas where others with similar socio-economic characteristics live. Thus, one is more likely to observe less concentration in the distribution of income than in a more heterogeneous group. The construction of several different variables designed to reflect the heterogeneity of the population are described in the following section.

III. MEASURED VARIABLES AND PRINCIPLE HYPOTHESES

Since the exact form of the hypotheses depends upon the definition of the dependent variable, we first elaborate on the particular measure of income inequality used in the analysis. We then turn to a discussion of the independent variables and their hypothesized relationship with the measure of inequality.

The measure of income inequality used here is the Gini coefficient of concentration.[4] This variable measures the area between the Lorenz curve and the diagonal line of perfect equality of income distribution as a proportion of the total area under the diagonal line. The coefficient is thus positively related to the level of income concentration or income inequality within a population.

To estimate Gini coefficients for each census tract we utilize the distribution of families by income class as published in the U.S. Bureau of the Census, Census of Population and Housing: *1970 Census Tracts* (1972a). However, since this data source does not contain income class means, we assume that the income class mean, for income classes below the median income class, is equal to the class mean for the United States as a whole and as published in the *Current Population Report* (1972b). For income classes lying at and above the median income level for the tract, we assume that the upper tail of the income distribution follows a Pareto distribution. We then use least-squares regression to estimate income class means in the above-median income classes.[5]

Five different variables are used to reflect the different levels and dispersions of human capital characteristics of the population within a Census tract as specified in the human capital model. The first variable, median education (found directly in the Census tract data) is hypothesized to be positively related to the level of income concentration in the tract. The dispersion of the distribution of education in the tract is also hypothesized to be directly related to the level of income inequality. The dispersion measure used here is the variance of edu-

cation computed for each tract from the grouped data published by the Census Bureau (1972a) with midpoints of the education groups used as class marks.

The third and fourth human capital variables are the mean and variance of the age distribution within a tract for all persons older than 16. Since these data are not provided directly by the Census, they were computed from the grouped data shown in the Census Bureau (1972a), once again using midpoints of the groups as class marks. Assuming that both age variables act as proxies for experience in the labor force, we expect that both will positively affect the level of income inequality within the tract.

For the final human capital variable we use the tract unemployment rate for males at the time of the 1970 Census. Chiswick and Mincer (1972) as well as Chiswick (1974) note the weakness of using this variable as a proxy for the variance in weeks worked—the theoretically proper variable as drawn from the basic model. However, since a summary measure of weeks worked is not available directly from the published tract-level data Census, it was deemed preferable to include unemployment, with the expectation that higher unemployment rates for males would be positively associated with income inequality.[6]

In the first estimation below of the determinants of small area inequality only these five variables are utilized. As noted above, for small-area data residential choice may also be an important inequality determinant. We use two different proxies for residential choice. The first of these is based upon the percent of the population which has not moved to a new address within the past five years. We hypothesize that this variable is non-linearly related to income inequality. This non-linearity is based on the premise that persons do choose to live among others with similar socio-economic backgrounds. Thus, an area which is *currently* experiencing much change (i.e., may have approximately half of its population who have moved in during the past several years) will exhibit greater income inequality than tracts which either have experienced little turnover in the past or have experienced a great deal of turnover with most of its population new in-migrants. This expected non-linearity leads to the use of two variables of population stability—the percent of the population in the tract living at the same address as five years previous to the census and the square of this variable. The parabolic relationship hypothesized would lead to the expectation that the percent of nonmovers would be positively related to income inequality and the squared amount would be negatively related to inequality.

While the observed mobility of the population may provide some indication of residential choice, occupational characteristics may also affect small area income inequality.[7] One occupational characteristic of the labor force which is associated with the "riskiness" of earnings stability and therefore likely to be positively associated with greater

income inequality[8] is the proportion of persons in self-employed occupations. We thus hypothesize a positive relationship between this variable and the tract-level Gini coefficient.[9]

A second purpose of this paper is to explore possible racial differences in determinants of income inequality. As noted in the following section, the segregated housing patterns of Atlanta allow us to consider predominantly white and black-subpopulations operating in the same labor market. Chiswick (p. 116-118) notes how racial differences in the levels of human capital investments, in rates of return to such investments or in secular changes in younger cohort earning profiles could lead to flatter cross-sectional income profiles for blacks than for whites, and thus, *cet. par.*, lesser income inequality for blacks. Segmentation of the data and replicated analysis for the racial sub-groups allows us to explore the racial differences in income inequality for persons within the same spatial labor market.

IV. EMPIRICAL RESULTS

As indicated above, the data used in this analysis were all obtained from published *1972 Census of Population and Housing* data for census tracts in Atlanta, Georgia. Atlanta provides an interesting geographical location for these tests since it contains a substantial black population with certain areas containing fairly high-income blacks. Also, the segregated pattern of housing lends itself to a segmentation of the data into white and black areas.[10] (There are few other minority groups that populate Atlanta.) The method of segmentation is to divide the tract-level data according to racial (black/white) majority populations. In this way the sample is divided into two subsamples of approximately equal sizes (59 majority white tracts and 53 majority black tracts.)[11]

Information regarding the underlying income data is exhibited in Table 1 where summary statistics of the tract-level Gini coefficients and median family income for the tracts is presented for the entire sample of 112 tracts as well as the two subsamples.

TABLE 1

Summary of Tract Level Income Concentration and Median Income

Variable	Sample	Mean	Standard Deviation	Minimum	Maximum
Gini	All tracts	.357	.068	.217	.557
	White tracts	.345	.074	.217	.557
	Black tracts	.370	.059	.244	.531
Median Family Income	All tracts	$ 8,732	$5,001	$1,951	$29,793
	White tracts	11,049	5,599	1,951	29,793
	Black tracts	6,152	2,354	2,273	14,329

The income data in Table 1 indicate a general lower income for pre-dominantly black areas relative to predominantly white areas, although there is less deviation in the black median income levels. For pur-poses of comparison, median income for all black families in Atlanta in 1969 was $6,451 compared with $8,399 for the city as a whole.

Table 2 presents the regression results for three different formula-tions of the income inequality relations. (Table 3 contains definitions of the terms used in the table along with means of the variables.) Re-gression 1 contains the five basic human capital variables and proves to be quite interesting. While the five variables together are signifi-cant for both the pooled and racially segmented subgroups, the level of schooling in the tract is found to have opposite effects for the two races (which explains its insignificance in the pooled regression). After accounting for other schooling, employment and age characteristics, white tracts exhibit the hypothesized positive relation between educa-tion and income inequality. On the other hand, in predominantly black tracts greater median education is negatively related to income in-equality. This is likely due to the flatter income profiles found for blacks (an extremely likely phenomenon in Southern Cities such as Atlanta) as hypothesized by Chiswick.

As hypothesized, higher tract unemployment rates are found to be associated with greater income inequality, with the size of this effect greater in black areas with their greater incidence of unemployment. While the variability of schooling is significant at the .10 level for the pooled data, neither of the segmented groups exhibit significant effects of the variance of education. The variance in age is significant as hypo-thesized for the pooled and white subgroup but insignificant for blacks. Average age, which does not show much variability across tracts, is not significant for either the segmented or pooled data. For cross sectional data such as these, the explanatory power of the regressions is quite high, especially for the predominantly black areas.

None of the five variables used in equation 1 corrects for residential choice effects. In equation 2 and 3 the mobility and self-employment variables are used to attempt to remedy this deficiency. While none of these variables are significant at the .10 level (although all but the mobility variables for black tracts are of the hypothesized signs), their inclusion does have some effect upon the human capital variable coef-ficients. Of primary interest is the lack of significance in the education variable in white tracts. The insignificance in the median education and variance of age variables is likely due to collinearity among all these variables for, while each is insignificant, when all are included in the regression the explanatory power of the white results increase from .29 to .37. For the white tracts there is a positive correlation of .18 be-tween non-movers and education suggesting the role played by "estab-lished" (in general upper income) white neighborhoods. This is to be compared with the -.27 correlation between education and non-movers

TABLE 2

Results of Tract-Level Gini Coefficients Regressed on Socio-Economic Characteristics[a]
(Standard Errors in Parentheses)

Regression	Group	Constant	Median School	Male Un-emp.	Var. School	Mean Age	Var. Age	% Self Employ	% No Move	(% No Move)²	R²	F
1	All	.120	.314 (.337)	.661 (.218)**	.448 (.239)*	.019 (.189)	.038 (.014)**				.24	6.90**
1	White	—.024	1.416 (.490)**	.671 (.260)**	.706 (.455)	—.131 (.255)	.052 (.020)**				.29	4.36**
1	Black	.261	—1.226 (.658)*	1.004 (.358)**	.476 (.357)	.206 (.250)	.012 (.021)				.43	7.14**
2	All	.134	.032 (.422)	.632 (.220)**	.501 (.246)**	—.004 (.229)	.033 (.015)**	.266 (.233)	.164 (.235)	—.229 (.272)	.26	4.54**
2	White	—.033	1.003 (.661)	.581 (.271)**	.620 (.465)	.069 (.331)	.031 (.022)	.395 (.344)	.337 (.399)	—.599 (.442)	.37	3.72**
2	Black	.287	—1.330 (.726)*	1.011 (.383)**	.514 (.391)	.156 (.320)	.013 (.024)	.160 (.348)	—.038 (.289)	.038 (.341)	.43	4.23**
3	All	.171	.069 (.413)	.657 (.218)**	.458 (.240)*	—.073 (.209)	.038 (.014)**	.236 (.230)			.25	5.92**
3	White	—.009	1.353 (.662)**	.672 (.272)**	.684 (.483)	—.152 (.294)	.052 (.020)**	.047 (.329)			.29	3.56**
3	Black	.276	—1.296 (.680)*	.995 (.362)**	.495 (.362)	.157 (.272)	.014 (.022)	.157 (.330)			.43	5.89**

[a]Definitions and means for variables given in Table 3.
*Denotes significant at less than 0.10 level
**Denotes significant at less than 0.05

TABLE 3

Definitions and Means of Regression Variables

Variable	Definition	All	Tract Means White	Black
Gini	Gini Coefficient	.357	.345	.370
Median School	(Median Schooling For Persons 25 years and older)/100	.108	.117	.098
Male Un-emp.	Percent males 16 years and older unemployed	.038	.035	.042
Var. School[a]	(Variance of School for Persons 25 and older)/100	.124	.107	.142
Mean Age[a]	(Mean age of persons 16 and older)/100	.413	.424	.400
Var. Age[a]	(Variance of age for persons 16 and older)/100	3.018	3.080	2.949
% Self Employ	Number of self employed persons in a tract as a proportion of all employed persons in the tract	.045	.061	.028
% No Move	Proportions of persons 5 and and older living in same house as in 1965	.445	.426	.467
(% No Move)^2	Square of % No Move	.217	.200	.236

[a]Computed from class data using mid-points as class marks

are likely to be less able economically to move, possibly due to the dual constraints of job locations and public transportation.

The self-employment variable had essentially no effect upon the results, possibly because of little variability in this variable across tracts. The unemployment variable, on the other hand, was consistently significant and extremely stable across the various formulations.

V. SUMMARY AND CONCLUSION

Based on previous work regarding the human capital approach to income distribution, we have hypothesized several variables as important determinants of income concentration in small areas. Because of the importance of residential choice when considering smaller geograph-

ical areas, we have also posited how the mobility and occupational characteristics of the population may affect income inequality. These hypotheses were then tested using 1970 Census tract data for Atlanta.

The results of the empirical tests were quite interesting, particularly the finding that while higher education levels are positively associated with income inequality in predominantly white areas, education and inequality were negatively related in black tracts. At the same time unemployment had a significant and positive effect upon income inequality within all tracts. The variability of age and education demonstrated some positive effect on inequality, but not consistently so. Likewise the mobility and occupation proxy variables used here did not significantly influence that tract-level Gini coefficient.

Two conclusions that can be reached from this analysis are not especially limited to income distribution studies. One is that small area data can provide useful insight into labor and earnings behavior of persons confined to a single labor market. The other is that theories purported to be universal may break down when different subgroups, such as blacks, and whites, are analyzed separately.

FOOTNOTES

[1]See, for example, Al-Samarrie and Miller (1967), Aigner and Heins (1967a), Farbman (1973) or Sale (1974).

[2]Another "weakness" is that by concentrating on a single metropolitan area we cannot measure income distributional impacts of the economic characteristics of the area. This also implies that it is not possible to test the proposition presented by Kuznets (1955) regarding the effect of economic development on income equality.

[3]See Mincer (1970) or Blinder (1974, pp. 3-16) for review of distribution theories.

[4]We recognize that the Gini coefficient is certainty not a perfect measure of income inequality. For an example of the differences in income equality measures and their effect upon the ordinal ranking of inequality results see Aigner and Heins (1967b).

[5]See Aigner and Heins (1967b) for details on this estimation of means.

[6]For additional discussion of the effects of unemployment upon income concentration see Reder (1969, pp. 238-240).

[7]See Wolff (1976) for a discussion of earnings inequality within narrowly defined occupations.

[8]This was originally emphasized by Friedman (1953) who hypothesized that the degree of risk attached to an occupation will be positively related to the dispersion of incomes to persons in that occupation.

[9]As a measure of occupational heterogeneity we also constructed two- and three-occupations "concentration ratios" for tracts. The results for other variables in the regressions were not substantially affected while the concentration variables did not generate any additional information.

[10]That segregated housing patterns hold within the City of Atlanta is suggested by the fact that 86 of the 112 census tracts used in this analysis contain either greater than 90% or less than 10% black persons.

[11]Five tracts were excluded from the analysis either because of extremely small populations due to bisection by the city boundary or because they had greater than 15% of the male population residing in institutions.

REFERENCES

Aigner, D. J. and A. J. Heins (1967a), "On the Determinants of Income Equality," *American Economic Review,* (March) 57: 175-184.

Aigner, D. J. and A. J. Heins (1967b) "A Social Welfare View of the Measurement of Income Inequality," *Review of Income and Wealth,* pp. 12-25.

Al-Samarrie, A and H. P. Miller (1967) "State Differentials in Income Concentration,"

American Economic Review, (March) 57: 59-72.

Becker, G. S. (1967) *Human Capital and the Personal Distribution of Income; An Analytical Approach.* W. S. Woytinsky Lecture No. 1, (Ann Arbor: University of Michigan).

Blinder, A. S. (1974) *Toward an Economic Theory of Income Distribution* (Cambridge; MIT Press).

Chiswick, B. R. and J. Mincer (1972) "Time-Series Changes in the Personal Income Inequality in the United States from 1939, with Projections to 1985," *Journal of Political Economy* (Supplement, May-June), pp. 534-566.

Chiswick, B. R. (1974), *Income Inequality* (New York: NBER).

Farbman, M. (1973) "Income Concentration in the United States," *Review of Economics and Statistics,* 55: 333-340.

Friedman, M. (1953) "Choice, Chance and the Personal Distribution of Income," *Journal of Political Economy,* (August), 59: 277-290.

Kuznets, I. B. (1955) "Economic Growth and Income Inequality," *American Economic Review,* (March), 45: 1-28.

Mincer, J. (1970) "The Distribution of Labor Incomes; A Survey with Special Reference to the Human Capital Approach," *Journal of Economic Literature,* 8: 1-26.

Reder, M.W. (1969) "A Partial Survey of the Theory of Income Size Distribution," *Six Papers on the Size Distribution of Wealth and Income,* L. Soltow (ed.), (New York: NBER).

Sale, T. S., III (1974) "Interstate Analysis of the Size Distribution of Family Income, 1950-1970," *Southern Economic Journal,* (January) 40: 434-441.

U.S. Census (1972a) *Census of Population and Housings 1970 Census Tracts,* (Washington: Government Printing Office, 1972).

U.S. Census Bureau (1972b) *Current Population Reports,* Series P-60, No. 75, "Income in 1969 of Families and Persons in the United States," (Washington: Government Printing Office).

Wolff, E. N. (1976) "Occupational Earnings Behavior and the Inequality of Earnings by Sex and Race in the United States," *Review of Income and Wealth,* (June), pp. 151-166.