THE SIZE DISTRIBUTION OF FAMILY INCOME IN U.S. SMSAs, 1959*

BY MICHAEL FARBMAN

University of Glasgow

An attempt is made in this paper to identify and quantify the relative influence of several economic, social, and demographic factors on variations in the size distribution of family incomes in 208 Standard Metropolitan Statistical Areas (SMSAs) in the United States in 1959. Using a simple ordinary least squares model with Gini's concentration ratio (R) as the proxy for family income inequality, the estimating equations explain up to 89 percent of the SMSA-to-SMSA variation. The "best" explanatory variables are those having to do with size of nonwhite population, occupational structure, and median years of education. City size and region—which are represented by dummy variables—are also revealed as playing an important role, both on their own and in conjunction with other of the independent variables.

I. INTRODUCTION

The essential problem of economics is how to increase economic welfare. This problem has two separate but interrelated dimensions: (1) how to increase total output from a given stock of productive resources; and (2) how to distribute this output among the members of the community so that it will be of maximum benefit to them. This paper is concerned principally with the second aspect, otherwise referred to as the problem of distribution.

Each individual or consuming unit in a society will not necessarily have equal command over goods and services. The ability to consume is directly related to the payments received by individuals in return for the use of productive resources owned by them, supplemented by public and private transfers. In large part the extent of these payments will be a function of the mix, size, and quality (real or perceived) of these resources. From the fact that these resources are not equally distributed among all members of society, we may infer that absolute equality in the size distribution of command over goods and services is not likely either.

Certainly this phenomenon has been widely observed, at all times and in all places, often reflecting and correlating with unduly high incidence of measured poverty, even in countries with the over-all productive capacity to provide all their citizens with an "adequate" standard of living. The welfare implications of the preceding statement are obvious; however, leaving to the welfare economist the task of determining what degree of equality constitutes maximum welfare, this paper will deal with the related but no less important problem of identifying the factors which appear to be closely associated with spatial variation in the inequality of incomes. Of course, the concept of welfare in general—rather than the identification of an optimum—will necessarily underlie this study in its entirety.

*This paper was drafted for presentation at the Urban Economics Conference held July 1973 at the University of Keele, England, sponsored by the Centre for Environmental Studies of London. It is based on the author's Ph.D. dissertation.

II. Income Inequality in the United States

The United States during the 1960's saw its share of discontent stemming in large part from the persistence of poverty in the midst of unprecedented plenty. Not since the Great Depression were the needs of the poor and underprivileged the object of so much official policy and attention, not to mention propaganda. Estimates of the size of the sector receiving inadequate incomes ranged from fifteen to forty percent of the U.S. population, depending on who was doing the estimating and what criteria of adequacy were being used.

While continuing to recognise that poverty and economic under-development in rural areas constituted a significant portion of the over-all problem, economists were increasingly forced to focus their attention on the dualistic nature of opportunity and the consequent unequal income distribution which had become institutionalised in most of the larger U.S. cities and which had been associated with massive civil disorders in several of the largest. This is not to say that there was anything especially perverse about the distribution of income in urban areas in general—in fact, as Table 1 shows, if anything, urban incomes are considerably more equally distributed (in aggregate) than are non-urban incomes in the U.S.[1] But viewed on a city-by-city basis, the picture is entirely different: As one would expect there is an enormous amount of inter-city variation about the mean of virtually any single-statistic income inequality measure one can contrive.[2] From what underlying sources does this variance arise? Is there perhaps more than a casual association between those factors which can be identified as

TABLE 1

Distribution of Family Income in 1959 for the United States, Urbanized Areas, and Non-urbanized Areas

Total Money Income, 1959	Percentage Distribution of Families: Whole U.S.	Urbanized Areas	Non-Urbanized Areas
Under $1,000	5.6	3.8	9.9
$1,000–1,999	7.5	5.6	11.9
$2,000–2,999	8.3	7.0	11.7
$3,000–3,999	9.5	8.5	11.9
$4,000–4,999	11.0	10.5	12.1
$5,000–5,999	12.3	12.7	11.5
$6,000–6,999	10.7	11.5	8.7
$7,000–9,999	20.1	22.6	13.8
$10,000–14,999	10.5	12.3	6.0
$15,000 and over	4.6	5.5	2.4
Gini's *R*	0.374	0.356	0.409

Source: U.S. Bureau of the Census, *U.S. Census of Population*: 1960, Vol. 1, *Characteristics of the Population*, Part 1, U.S. Summary, 1964, Table 224.

[1]This refers to *money* incomes. Naturally, this assertion is subject to revision if the comparison is made in real terms.

[2]Viz. for our sample of 208 SMSAs in 1960, the mean of Gini's concentration ratio for family incomes is 0.359, ranging between 0.297 and 0.473, and having a standard deviation equal to 0.034.

"explaining" the inter-city variation in income inequality and those identified[3] as responsible for the urban unrest typical of the 1960's?

Apart from the fact that establishing this link *per se* is not meant to be the main thrust of this paper, any conclusion in this respect would almost certainly be dependent upon the prior identification, as mentioned above, of those factors which appear to be associated with variations in income inequality among urban areas. It is this more basic empirical relationship that we shall attempt to establish: Given the special nature of large cities in the U.S. to-day, what are the relevant economic, social and demographic factors which cause inter-city variation in the shape of the income distribution, and by proxy, welfare?

III. A Spatial Theory of Income Inequality

In recent years, as interest in the size distribution of income has slowly gathered momentum within the economics profession, an increasing number of researchers have begun turning their attention to theories and models for explaining spatial variations in inequality.[4] In 1967, articles by Al-Samarrie and Miller, and Aigner and Heins were published in the *American Economic Review* on the subject of inter-state (U.S.) variations in the concentration of family incomes.[5] Small area models have also been formulated, but in none of these has the major thrust been toward a comprehensive empirical analysis of variation in income size distributions in urban areas.[6]

Before explicitly specifying the explanatory model it will be necessary to dispose of two conceptual problems having to do with the observation unit. The first of these refers to the size of the spatial area. While several definitions of urban area could have been employed, the Standard Metropolitan Statistical Area (SMSA) was selected on account of the functional nature of its definition,[7] and also because a complete data set of relevant statistics was known to be available for estimating purposes.

The second conceptual problem has to do with the income recipient units among which inequality in each SMSA was to be measured. Here the data constraint immediately narrowed the choice to just two: individuals and families. Either could have been used, but certainly not inter-changeably since equal distribution among individuals has very different welfare implications from

[3]Cf. the *Report of the Presidential Advisory Commission on Civil Disorders.*

[4]In fact the whole sub-field of income distribution has become a growth industry within the profession, as witness the proliferation of major books on this subject in recent years by H. Lydall, M. Bronfenbrenner, J. Pen, *inter alia,* along with a host of less ambitious volumes and journal articles too numerous to mention.

[5]But see my paper in the *Review of Economics and Statistics* (1973) which is critical of the information loss attributable to aggregation which is implicit in the use of large analytical units such as states.

[6]A good example of this is Mattila and Thompson's (1968, pp. 63–80) macro-model of urban economic development, in which one small corner of this admittedly epic model is devoted to a rough identification of the factors contributing to urban income inequality, although the statistical measure of inequality is fairly crude and the estimates based only on a sample of SMSAs. See also Burns and Frech (1970), Frech and Burns (1971), and Newhouse (1971) for more of the same.

[7]For an explanation of the criteria for defining SMSAs, see U.S. Bureau of the Budget, Office of Statistical Standards, *Standard Metropolitan Statistical Areas,* (Washington: G.P.O., 1967). On the subject of the appropriateness of SMSAs for analytical purposes, see Murray (1969, 1970).

equality among families.[8] The choice, from the point of view of the objectives of this study, need not be an easy one. On the one hand, the individual income recipient unit is more useful than a family unit in reflecting differences in contribution to the production process out of which income originates. Income distribution, however, only assumes full meaning when viewed from the consumption side, and in this respect it may be argued that families are generally the units which consume. If we are willing to concede, however, that decisions to contribute to the production process may also largely come under the control of families, then the preponderance of theoretical logic is persuasive of the appropriateness of using families as the income recipient unit. At least that is our choice in this paper.

Many theories of income distribution exist, some sophisticated and highly abstract, others tending toward the intuitive. Some, such as the institutional models, de-emphasize the traditional economic aspects, while others fit comfortably into contemporary macro-economic theory. For the purposes of this study we have chosen a simple econometric model whose purpose, in accordance with Reder's dictum, is to describe "... the direction and, where possible, the extent to which changes in certain structural parameters alter the size distribution of income or some component thereof."[9] Most persons familiar with this field would probably agree that the more important of these parameters are those reflecting technology, resource utilisation, level of development, individual ability, monopoly power and taste. An empirical model of income distribution, then, should relate some measure of the distribution to the above and other relevant structural parameters of an economy.

Considering this in terms that will lead us to explanatory variables which are empirically testable, we may begin by linking the distribution of income to the distribution of factor ownership among persons, and factor scarcity in the market economy, as modified by transfers. Earnings are the rewards from participation in the productive process and constitute the major portion of personal income. This forms what Reder would call the "initial" distribution,[10] and will be followed in turn by a "secondary redistribution" which takes into account the effects of voluntary and involuntary transfers. The former (the "initial" distribution) is determined by the price of factor services multiplied by the quantities sold; but for a variety of reasons this price will not be uniform, in as much as individuals do not compete in identical markets for earnings. An individual's talents, acquired training and skills, and experience, all differentiate the services he has to offer, while imperfections in factor markets and discriminatory barriers may further influence his ability to realise his potential earnings. Thus the final distribution—the one we wish to explain—will be a function of more than just the fundamental distribution; it will necessarily be altered by the effects of inborn differences in gifts and maintenance of acquired advantages through environment, the effects of cyclical changes in economic activity, the effects of public policy, the effects of socio-geographic factors such as degree of urbanisation and

[8] It is perfectly conceivable that income could be relatively equally distributed among individuals while unequally among families.

[9] See M. W. Reder (1969, p. 205).

[10] *Ibid.*, p. 207.

geographic location, the effects of demographic factors, and the effects of discrimination.

IV. Factors Relating to Income Inequality

In the preceding section the broad factors that may explain interpersonal and spatial differences in income concentration were introduced. The purpose of this section is to specify quantifiable explanatory variables which derive therefrom, and to formulate hypotheses concerning the association of these variables with the area measure of income inequality. Only those presumably independent factors which are deemed important and which lend themselves to statistical testing will be considered.

1. *Personal Characteristics of Income Recipients*

A. *Racial discrimination* (X_1). An individual's colour must be evaluated as a characteristic associated with differences in income, not because of any established difference in talent between white and nonwhite persons, but because this attribute reflects the effect of present and past imperfection and discrimination in labour markets. For the purposes of this study, imperfections in the labour market exist when there is discrimination against the employment of a specific segment of the labour force and/or when the price for a particular type of labour is not uniform in all uses. Principally this treatment is applicable in an urban context to nonwhites and, to a lesser extent, females.[11] A relatively poor occupational mix for some group may indicate *job* discrimination, while low earning rates, standardised for occupation, may indicate dual wage scales.[12] It is the latter phenomenon which concerns us here.

Certainly there is no lack of evidence that racial discrimination has been a very real phenomenon in U.S. labour markets, normally having a most telling effect in connection with the earnings component of income. This earnings effect has been well documented at all levels and from all points of view, including studies on wage determination, human capital and rates of return to investment, occupational and labour market segmentation, and so on.[13] Of particular interest is a recent study using 1960 Census data, in which Chiswick (1973) was able to substantiate what he calls his "employee discrimination hypothesis." As opposed to the traditional view of employer-based discrimination leading to dualistic conditions on the demand side, this construct stresses conditions of supply of labour when employment is integrated. That is, owing to the fact that white workers may require a wage premium to get them to work alongside nonwhites of equivalent training and experience, a component of inequality within skill levels will result, which, up to some point, will be larger, the larger the proportion of

[11]A variable reflecting discrimination as applied to females *per se* was not included for theoretical reasons: In the first place it would be collinear with an activity rate variable which is to be included. Secondly, fluctuations in the internal composition of this group are too large and therefore unreliable to sample at a point in time. Also, Al-Samarrie and Miller (1967, p. 67) included such a variable in their study but found it not to have significance.

[12]This has been demonstrated empirically by J. Gwartney (1970).

[13]For a good up-to-date coverage of many of these studies, see O. Ashenfelter and A. Rees (1973).

nonwhites. We shall use the proportion of nonwhites residing in each SMSA as a proxy for this much-indentified dualistic effect, although without particular reference to its being either supply- or demand-determined. Specifically, our hypothesis is that there should be a direct relationship between this percentage and our measure of income inequality.

B. *Occupational affiliation* (X_2). The varying occupational structure from SMSA to SMSA would appear to be a significant source of disparity in the respective concentration ratios of SMSAs. There are generally wide variations in the levels of incomes received by families, depending on the occupation of the head. The annual incomes of families headed by professional, managerial and technical workers varied (in 1959) anywhere over a range of from approximately \$5,500 on up to six digit levels. In like manner, incomes to families headed by low-skilled workers may range from hundreds of dollars to, in many instances, over \$10,000, depending in part on the industry in which the income recipient works, union influence, availability of overtime and other pertinent factors. Such large percentage variations in the intra-group incomes of these two groups manifest themselves as high concentration ratios. On the other hand, families of workers in middle-level occupations generally have incomes which do not deviate from the median in percentage terms as much as do the two polar groups. Therefore, as a group they show the lowest concentration ratio. Stated as a rule, income is least concentrated in middle-level occupations and becomes progressively more concentrated as one moves to occupations with extremely low or high median income, particularly the former.

It follows, then, that the lower the inequality of income within an occupational grouping and the larger the representation of that occupation within the SMSA, the smaller is the inequality of total income received by residents of that SMSA, other things being equal. "Craftsmen, foremen and kindred workers", "Clerical and kindred workers", and "Operatives and kindred workers" are the three occupational classes in the U.S. labour force in 1960 which show the lowest concentration ratios,[14] and which normally constitute a relatively large percent of the employment within most SMSAs. By this logic it seems reasonable to hypothesize that there is a negative association between the concentration ratio of income received by families in each SMSA and the percent of the labour force employed in the above-mentioned middle-level occupations.

C. *Education* (X_3). Training in one form or another is the usual means by which a person qualifies for a more responsible position and thereby enhances his prospects of earning a higher level of income. This intuitive idea is supported theoretically by the human capital school (Mincer, Becker, *et al.*) who view future earnings as a flow of returns to investment in education and training, appropriately discounted. Soltow (1960) has found that incomes tend to be more equally distributed among more highly educated or skilled groups of people. He attributes this phenomenon principally to variation in the numbers of persons with complete high school educations, the educational group having the lowest internal income dispersion. It should follow then, that the concentration ratio of family

[14]Cf. H. P. Miller (1963, Table 7).

income will vary inversely with the median school years completed by persons aged twenty-five years and older in the SMSA.

2. *Source of Income*

A. *Property and wages and salaries* (X_4). There are major differences in the degree of inequality in the distribution of the various income sources, with wages and salaries most equally distributed and property income[15] the least equally distributed.[16] These two shares may have an impact on the over-all income distribution through two distinct dimensions: the degree of inequality within each of the two groups, and the relative size of each. (*N.B.*: Because of transfers they need not constitute 100 percent of income received.) If we make the not unreasonable assumption as suggested throughout the literature that property incomes are always more unequally distributed than wages and salaries, then the total variance attributable to source of income will be a simple function of the latter dimension, i.e., the relative sizes of the shares of income arising from each source. In order to combine into a single variable the independent and opposite effects of the size of the shares of wages and salaries, and property income respectively, the proxy statistic used to measure this effect will be the ratio of the percentage of income from wages and salaries to the percentage of income from property sources. The over-all relationship as defined here should be negative: the larger the fraction, the less unequal will income distribution be, other things being equal.

B. *Transfers.* Although preliminary testing of several transfer income proxies revealed little or no systematic relationship with income inequality—and for this reason they will not be explicitly employed in this model—the general relevance of this source of income to the problem at hand is anything but insignificant.

There are several *a priori* reasons why the size of transfer income would be expected to affect the shape of the income distribution. The most obvious is that almost by definition, persons or families who receive transfer incomes are synonymous with persons or families enumerated in the low tail of the distribution. Also, given the urban focus of this study, one would have expected an over-sampling of these groups since so many transfer income recipients are residents of urban areas.

Transfers would also be important to a study such as this one since the unit we are dealing with is the family, and one of the principal effects of transfers is to enable more families to exist as separately domiciled independent units than would occur in their absence. If we accept that a high level of economic development (to be discussed below) fosters greater equality, then it seems important to acknowledge that it also encourages a statistical tendency toward higher measured inequality through its increased levels of transfer programmes and the subsequent establishment of low income independent families.

[15]By property income we mean interest, dividends (other than included in business income), and net income from rents.

[16]See D. S. Projector, G. S. Weiss, and E. T. Thoresen (1969, p. 111).

Notwithstanding this *a priori* logic, as mentioned above, no controlled relationship between transfers and income inequality looked like being demonstrated. The reasons for this can only be speculated upon in this paper, but a good guess is that since transfer incomes in a local area are a function not only of federal sources, but state, county and local as well, the variance in this statistic was so inflated that no meaningful covariance relationships could be revealed.

3. *Labour Force Participation*

A. *Activity Rates* (X_5). Although several alternative means of measuring activity rates are available and theoretically justifiable, regardless of which is chosen one of its principal attributes will be that it serves a dual purpose in accounting for both unemployed and out-of-labour-force persons although it does not separate out the differential contribution of each component to the variance of the dependent variable.[17] A high employment to population ratio implies more multiple-earner families and/or fewer families with no income-earners hence higher family incomes and, *ceteris paribus*, few families falling at the lower end of the income distribution biasing the inequality statistic upwards. In relating this factor to the level of family inequality in each SMSA it is hypothesized that inequality will be negatively associated with the ratio of employment to population in the SMSA.

4. *Economic Development and Industry*

A. *Median family income* (X_6). Expecially since Kuznets' studies relating economic development to income inequality (1955, 1963), many researchers have demonstrated the tendency for personal incomes within a region to be more equally distributed the more maturely developed is the region. For instance, Kravis (1960) has found that there is a discernible, though not perfect, tendency for underdevelopment, low incomes and inequality to go hand in hand, and for development, high income and relative equality to be associated with one another. This is essentially the relationship demonstrated by Williamson (1965) for countries, states and regions, and by Aigner and Heins (1967) for states.

In addition, it has been observed that real family incomes have been moving secularly upward during the last few decades, responsive for the most part to generally full employment, an expanding economy and large productivity increases. A few subgroups, such as the aged, uneducated and low-skilled are conspicuous exceptions to this trend; however, the bulk of the remaining earners fall well within its definition. Concurrent to this general upward movement there has been a tendency towards clustering at middle income levels as real and institutional ceilings in incomes are approached (Fitzwilliams, 1964). Accordingly, it would be expected that as income rises and as a central tendency begins to appear, due to the lessening dispersion caused by marginal incomes—heretofore the strongest influence toward high concentration—the statistic reflecting concentration should decrease.

[17]This comprehensiveness is, of course, a mixed blessing, since we are then confronted with the question of whether activity rates are not *too* aggregated to be meaningful in this type of analysis.

Thus, in keeping with most of the research in this field on larger spatial units—where some measure of family income is used as a development proxy, and where the hypothesized relationship between income inequality and economic development is negative—we shall expect a similar inverse relationship for SMSAs.

B. *Industry* (X_7). Responding to the same forces as the previous variable, it would be expected that economic development leads, through capital formation, to an industrial structure with higher productivity commensurate with the level of development. Thus, as a result, there may be inter-area variations in value productivity according to degree of development, and more importantly, there will certainly be intra-area productivity variation among industries out of which variation in the amounts paid to industrial employees will arise.

The opulence of an area may be related generally to the predominance of manufacturing, especially if there are large monopoly and export components involved. As Thompson suggests (1966, p. 94), where these conditions obtain, the monopoly aspects will have the effect of transferring income from nationwide consumers to local producers, and given strong egalitarian unionism, ultimately into the hands of workers. In addition there may be a further wage roll-out effect through the multiplier boosting non-export earnings elsewhere in the local economy. All this should contribute to a relatively more equal distribution of income in the area.

Manufacturing industries with high capitalisation and selling in markets in which they exercise some control over prices will normally have high productivity—not necessarily physical productivity, but almost certainly *value* productivity. According to our theory, SMSAs possessing industries of this type should have more equal distributions of income. The translation of high productivity into income size equality will be approximated by the wage level for production workers in manufacturing which is obtained from the 1958 Census of Manufactures.[18] Hence, it is our hypothesis that there will be an inverse relationship between average hourly wage for manufacturing production workers and the inequality of family incomes in SMSAs.

5. *Demographic and Geographic Effects*

A. *Population density* (X_8). Kuznets (1955) has shown that urbanisation, industrialisation and a shift away from agriculture invariably accompany growth and reduced inequality in developed countries. The income distribution of all families in an SMSA can be viewed simply as a combination of the income distributions of the urban and the suburban populations. It is generally assumed that the median family income of urban residents is lower than that of surburban residents and that the concentration coefficient will indicate more income inequality in urban areas than suburban.[19]

[18]The author is grateful to an anonymous referee for his suggestions for improving the explanatory power of this variable.

[19]For a detailed treatment of residence patterns and incomes in urban areas, see Homer Hoyt (1966).

Thus it would appear that the larger the proportion of families in an SMSA living in what might be called its urban core, the greater will be income inequality, other things being equal. Of course it is known that quite often the central city is the home to not only large numbers of poor, but to substantial numbers of wealthy people as well. On the average this wealthy group constitutes less than 2 percent of the central districts in which they live (Hoyt, p. 13), and for this reason they may be found in areas with low median incomes, although their cohabitation in the same area as the poor will yield higher income concentration ratios than if the area were uniformly poor.

Population density—the number of residents per unit of area—should serve as a proxy for how closely an SMSA follows the above pattern. That is, greater population density implies a more highly urban than suburban character within the area. This being the case, it is hypothesized that the more highly "urban" is an SMSA as measured by population density, the higher will be inequality, other things being equal.

B. *Size of SMSA* (X_{9-11}). On both the the supply and demand sides, factors can be identified which would be expected to vary systematically with city size. In the case of supply there are several reasons why we might expect larger urban areas to show greater interpersonal income variation: First, large cities tend to be the major recipients of displaced farm workers and other low skilled workers. At the other extreme, large cities are also gathering points for the most gifted and ambitious persons, and furthermore amplify the earning abilities of residents by providing institutions offering the most advanced technical and professional training. Thus with resident human resources spanning such a wide spectrum the assertion that income inequality increases with city size can be built solely from the supply side.

The demand side is composed essentially of the complement to the above. The large city is more likely to be the home office of the large corporation complete with highly skilled researchers, executives, legal and other employees, many of whom command relatively high compensation. In sharp contrast, the smaller city is often biased toward mass production and the concomitant narrow range of wages and skills which are associated with mass production industries and industrial unions. In addition, while the jobs on production lines will often call for no greater skills than those in the service sector, the jobs typically pay better and the pay rates have a greater modal tendency.

A host of other arguments in support of increasing inequality commensurate with larger city size can also be made, *inter alia*, in terms of price structure and the weight of property incomes in larger cities. (Arguments in opposition could also be formulated, though they are less persuasive and will not be developed here.) The measure of city size to be employed will be total population. The 208 SMSAs[20] were categorised into four size groups and analysed as dummy

[20]Although 212 SMSAs are defined for 1960, only 208 are used for estimation in this study due to the absence of complete data sets for four of them.

variables,[21] with the hypothesis as above of a positive correlation between size and inequality.

C. *Region of Country* (X_{12-19}). While neo-classical theory would lead us to expect that income differentials among regions would eventually be equalised, the empirical evidence indicates otherwise.[22] Scully (1969, p. 758) has suggested three possible explanations for the failure to achieve inter-regional equality:

(1) Barriers to the free flow of resources among regions, which implies differing factor proportions.
(2) The non-homogeneity of the labour force, i.e., inter-regional variations in the quality of the labour force.
(3) Institutional factors, such as regional differences in bargaining strength and labour market discrimination against minority segments of the labour force, may impede wage uniformity.

While many of these sources of regional inequality are likely to be accounted for by other variables in the model, we would posit that there is a residual variation in income inequality from region to region which cannot be attributed to other variables and which will be described here as "pure" regional effect. For testing purposes each of the 208 SMSAs was categorised according to which of the standard regions it belonged in, and assigned a dummy variable for covariance analysis. Since we are not interested in demonstrating a systematic ordered variation between the dependent variable and region, the hypothesis in this case will merely imply a test for a statistically significant difference in the average level of the dependent variable between any two of the nine regions.

The Dependent Variable

The task of defining the "appropriate" measure of income inequality is seldom an easy one. The ultimate choice will always to a certain extent be arbitrary, since no established analytical framework necessarily implicitly prescribes or defines a unique concept of inequality which must be employed in a

[21]Although city size may of course be measured on a continuum (e.g., number of persons residing in the SMSA), we are using dummy variables expressly for the purpose of pointing up the discrete non-continuous and non-monotonic nature of the size effect. Proxying this effect by using the actual numbers of residents as an independent variable in least squares estimation (such as we shall employ) may present a very real danger of mis-specifying the underlying relationship. For example, given the spectrum over which SMSA population varied in 1960 (from 50,000 to over 10 million persons) it must be recognized that among cities of less than a million persons, 500,000 people more or less may represent or give rise to more meaningful differences in economic character than say a similar variation among cities numbering more than a million inhabitants. In the absence of an appropriate transformation (for which we have no *a priori* specification) a continuous measure of city size would implicitly value these equally, and this is intuitively a false concept. Dummy variables, on the other hand, will—at a small cost in degrees of freedom but in a sound statistical fashion—not only capture the size effect, but into the bargain will allow it to be seen and interpreted in the simple terms of a group-wise intercept shift.

[22]For detailed analyses of changes in regional income distribution, see B. F. Haley (1968) and U.S. Department of Commerce, O.B.E. (1953).

given situation.[23] We have chosen Gini's concentration ratio (R) of family incomes partly because it is based on the exceedingly useful graphical concept of the Lorenz curve, and also because it possesses certain attractive theoretical advantages such as its "mean difference" property and its independence of scale.

As with all such measures, Gini's R of course is known to suffer from certain defects, especially on the welfare side.[24] However, to paraphrase Sen (1973, p. 31), of the so-called "positive" measures, Gini's R, the coefficient of variation, and the standard deviation of the logarithms are the only single-statistic measures that can pass the statistical properties tests well enough to even be allowed to be tested for welfare implications. On grounds of the concavity of the group welfare function, Sen concludes (p. 34) that the other objections notwithstanding, the standard deviation of the logarithms of income is almost certainly inferior to Gini's R. As for Gini's R vis-a-vis the coefficient of variation, at least for the SMSA data in this study, the rank order coefficient of correlation between these two measures on various valuing assumptions constantly exceeded 0.90. For all these reasons, as well as because of its familiarity and frequency of use in similar studies, we selected Gini's R as the summary measure of family income inequality in this study.

The Data

Because of the detail in which it reports incomes and its relatively low errors, most of the data used to test this model are from the 1960 census of population. The only exceptions pertain to variables X_4 (source of income) which was compiled from unpublished Bureau of Economic Analysis data, and X_7 (average hourly wage for production workers in manufacturing), which originated in the *Census of Manufactures, 1963*, referring to the year 1958.

V. Results of Statistical Testing

In the previous sections the theoretical framework for this study was constructed and the hypotheses to be tested were outlined and explained. It is the purpose of this section to test these hypotheses and to estimate the effect of each of the independent variables on Gini's concentration ratio. In the interest of

[23]As is well known to most readers, a wide choice of inequality measures is available. Since so much literature already exists which deals specifically with the relative advantages and shortcomings of each, we do not feel compelled to duplicate the various arguments here yet again. Instead, the reader is advised to refer to Yntema (1933), Bowman (1945), Kravis (1962), Weisskoff (1970), Stark (1972), and Sen (1973), all of whom present useful general summaries of the properties of the more important measures.

[24]We refer to the ambiguity caused by crossed Lorenz curves. Given certain assumptions concerning the additivity of individuals' or families' welfare it can be shown that some social welfare function can be defined which ranks intersecting Lorenz curves differently from Gini's R. This is generally conceded to be the principal limitation of Gini's R, and it cannot be resolved in any way except by assumption. For an elaboration of this see Atkinson (1970) and Newberry (1970).

simplicity an ordinary least squares estimating technique will be employed.[25] The estimated functions are assumed to be linear and additive, and there are no interaction terms. Naturally, the usual cautions are also invoked against ascribing causality *per se* to the measured relationship where in fact it is wished to imply mere association.

It should be noted from the beginning—since the accuracy of the coefficients in the estimated regression equations depends on ascertaining or reasonably presuming this fact—that judging by the evidence of the correlation matrix in Table 2, the level of intercorrelation among the continuous variables used in this study does not appear to be intolerably large. Of course the correlation matrix alone is not the last word in this matter, and furthermore what constitutes a "large" correlation coefficient must remain to a certain extent arbitrary. But accepting the correlation matrix as a convenient proxy, the real danger of non-independence among the exogenous variables—multicollinearity—appears with one exception to be of limited likelihood. This exception is the coefficient (0.627) between variables X_7 and X_6, manufacturing wage and family income respectively. While the estimated coefficients of these two variables in the estimating equations may be somewhat ambiguous, and since neither is so highly correlated with any of the other variables, as long as proper care is exercised in interpreting the coefficients of X_7 and X_6 this minor effect can be localised and considered as relatively inconsequential.

[25]It has been pointed out that technically, since our observations consist of group means and not a complete enumeration of all families or persons on which they are based, the estimating data are themselves subject to errors which vary inversely with the size of the group whose characteristics are being summarised (i.e., the SMSA population). In short, there is the possibility of heteroscedasticity, for which the standard corrective is to use generalised least squares (or weighted regression) instead of ordinary least squares. This involves multiplying through each observation by the variate which is liable to be causing the non-constant variance of the error term, which in this study would be SMSA population (cf. Kmenta, pp. 322–29).

It is suggested in the instant case, however, that provided *all* the groups qualify in a statistical sense as "large" groups, then the properties of the means of all of these groups will not differ significantly depending on the actual numbers. In this sense all SMSAs are certainly "large" groups and consequently it may be asserted that the error variances of statistics relating to them—apart from all being infinitesimally small, since they are all divided through by large *n*'s—are more or less invariant with size and therefore do not present a real danger of heteroscedasticity. It might be noted in passing (as if to confirm this) that as far as the parameter estimates are concerned, the data in this study are insensitive to the use of weights anyway. For example,

1. OLS

$$Y = 0.6734 + 0.1135X_1 - 0.3334X_2 - 0.0083X_3 - 0.0020X_4$$
$$- 0.000009X_6 - 0.0023X_7 + 0.0000026X_8$$

2. WLS (Weight = $\sqrt{Population}$)

$$Y = 0.6410 + 0.1170X_1 - 0.3102X_2 - 0.0058X_3 - 0.0021X_4$$
$$- 0.000009X_6 - 0.0024X_7 + 0.0000029X_8$$

Two further reasons arise in this study for not using weighted least squares: First, the putative weight variable —SMSA population—is already slated for inclusion in the estimating equation, and its simultaneous use as a weight could cause larger econometric problems. Second, weighting may artificially influence the outcome of significance tests and impose extra qualifications on interpretations of the results.

TABLE 2

MATRIX OF SIMPLE CORRELATIONS FOR GINI'S CONCENTRATION RATIO AND 8 QUANTITATIVE EXOGENOUS VARIABLES, FOR 208 SMSAS, 1960

	Y	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8
Y	1.000								
X_1	0.627	1.000							
X_2	−0.680	−0.330	1.000						
X_3	−0.184	−0.116	−0.380	1.000					
X_4	−0.110	0.149	0.291	−0.271	1.000				
X_5	0.265	0.044	−0.197	−0.287	0.102	1.000			
X_6	−0.663	−0.371	0.290	0.504	−0.064	−0.444	1.000		
X_7	−0.540	−0.260	0.181	0.323	0.067	0.173	0.627	1.000	
X_8	−0.187	−0.029	0.326	−0.111	0.273	−0.291	0.269	0.063	1.000

Y: Gini's concentration ratio of family income.
X_1: Percentage of population nonwhite.
X_2: Percentage of employed persons in middle-level occupations.
X_3: Median years of education among persons 25 years and older.
X_4: Ratio of family income as wages and salaries to property.
X_5: Ratio of persons not in labour force to persons in labour force.
X_6: Median family income.
X_7: Average hourly wage for production workers in manufacturing.
X_8: Number of persons per square mile of land area.

The Estimating Equations

In all, four alternative estimating configurations were used for testing the hypotheses outlined above, and these are presented as Tables 3, 4, 5 and 6 (referred to as equation Variants 1 through 4 for the remainder of this paper). It will be noticed that no equations have been estimated with X_5 included. This is due to the fact that in preliminary runs the explanatory power of this variable was demonstrated to be at best marginal, especially in comparison to the remaining seven variables which were all significant at the 0.01 level.[26] In Variant 1, using only quantitative variables X_1 through X_8 (excluding X_5), the multiple R-square is 0.840. In other words, these seven variables alone are capable of explaining approximately 84 percent of the SMSA-to-SMSA variation in Gini's concentration ratio of family income. In addition, five of the variables are shown by t-tests to be significant at the 0.01 level, while the remaining two are significant at the 0.05 level.

In Variant 2 the addition of the population dummy variables ($X_{9\text{-}11}$) to the seven original variables raises the multiple R-square to 0.860, while in Variant 3 the combination of the original seven variables plus just the region dummy variables ($X_{12\text{-}19}$) yielded an R-square of 0.868. The ultimate inclusion of both sets of dummies together with the original seven variables in Variant 4 causes the

[26]Dropping this variable in this analysis is not meant to imply that we have shown *no* relationship between activity rates and income inequality, although this is a distinct possibility. This action says, rather, that X_5's effects in *this* model are of such small magnitude that for all intents and purposes its presence or absence is of no consequence whatever, except that it costs one degree of freedom and affects the significance tests accordingly.

TABLE 3

REGRESSION VARIANT 1—COEFFICIENTS AND SIGNIFICANCE TESTS FOR RELATION BETWEEN GINI'S CONCENTRATION RATIO AND 7 QUANTITATIVE EXOGENOUS VARIABLES FOR 208 SMSAs, 1960

Variable	Regression Coefficient	Standard Error	t-value	Level of Significance	R^2 Without Variable
X_1	0.1065	0.0113	9.442	**	0.769
X_2	−0.3159	0.0219	−14.405	**	0.674
X_3	−0.0084	0.0015	−5.653	**	0.814
X_4	−0.0014	0.0007	−2.067	*	0.836
X_6[a]	−0.0068	0.0020	−3.374	**	0.831
X_7	−0.0140	0.0031	−4.479	**	0.823
X_8	0.0000019	0.0000010	1.899	*	0.837
Y-intercept (a)	0.6669	0.0191	35.002	**	—

R^2, unadjusted = 0.840
R^2, adjusted = 0.834
Standard Error of Estimate = 0.0138
F-value (7; 200 d.f.) = 149.893**

[a]Measured in thousands of dollars.
*Significant at the 0.05 level.
**Significant at the 0.01 level.
X_1–X_4, X_6–X_8: Same as in Table 2.

multiple R-square to jump to 0.890. This means that the entire group of variables explains all but 11 percent of the variation in Gini's R. This unexplained variation may be due to the exclusion of relevant explanatory variables, imperfect measurement, or the use of an inappropriate estimating form, such as by specifying a linear model as we have done rather than curvilinear, which might have occasioned a better fit.

Among the original seven quantitative variables, in all equation variants each coefficient possessed the sign which was expected of it in accordance with the theoretical model specified above. That is, the coefficients of variables X_1 and X_8 were positive as hypothesized, while the coefficients of variables X_2, X_3, X_4, X_6 and X_7 were negative as hypothesized. Also, as measured by the F-values, each of the four regression lines as a whole is overwhelmingly significant.

The addition to the equation of first the population dummies (Variant 2) and then in their stead the region dummies (Variant 3) led to some interesting results. The technique for including the dummies involved the use of zero-one dichotomous variables according to class as specified above. (Note that in order to preserve independence the technique calls for $n-1$ dummies for n qualitative classes.) In Variant 2, which tests for the effect of population size, the missing class is the largest SMSAs with populations exceeding one million persons. This then effectively presents the problem as asking whether or not there is a systematic difference between the level of Gini's R in the largest SMSAs as compared to each of the three other smaller population classes, in addition to that already explained

TABLE 4

REGRESSION VARIANT 2—COEFFICIENTS AND SIGNIFICANCE TESTS FOR RELATION BETWEEN GINI'S CONCENTRATION RATIO AND 7 QUANTITATIVE VARIABLES PLUS POPULATION DUMMIES FOR 208 SMSAs, 1960

Variable	Regression Coefficient	Standard Error	t-value	Level of Significance	R^2 Without Variable
X_1	0.0923	0.0111	8.323	**	0.811
X_2	−0.3113	0.0207	−15.017	**	0.701
X_3	−0.0081	0.0014	−5.780	**	0.837
X_4	−0.0017	0.0007	−2.610	**	0.856
X_6[a]	−0.0096	0.0020	−4.740	**	0.844
X_7	−0.0143	0.0030	−4.771	**	0.844
X_8	0.0000012	0.0000010	1.190		0.859
X_9	−0.0170	0.0036	−4.675	**	
X_{10}	−0.0105	0.0036	−2.930	**	0.840
X_{11}	−0.0066	0.0033	−1.993	*	
Y-intercept (a)	0.6920	0.0187	37.079	**	—

R^2, unadjusted = 0.860
R^2, adjusted = 0.853
Standard Error of Estimate = 0.0130
F-value (10; 197 d.f.) = 121.391**

[a]Measured in thousands of dollars.
*Significant at the 0.05 level.
**Significant at the 0.01 level.
X_1–X_4, X_6–X_8: Same as in Table 2.
X_9: Dummy for SMSA with total population 50,000 to 150,000.
X_{10}: Dummy for SMSA with total population 150,000 to 250,000.
X_{11}: Dummy for SMSA with total population 250,000 to 1,000,000.

by the earlier inclusion of the seven original variables. The significant jump in R-square from 0.840 to 0.860 is a partial indication of the explanatory power added by the inclusion of this factor, but perhaps more useful is another observation which can be made: Our hypothesis of a positive relationship between city size and Gini's R is explicitly supported by the increasingly large negative slopes of the regression coefficients as one moves from X_{11}, SMSAs with populations between 250,000 and one million, to X_9, SMSAs with populations between 50,000 and 150,000 as shown in Variant 2. This is as if to say, our equation, without taking into account the effects of SMSA population size, predicts without bias for SMSAs with populations larger than one million persons; but for SMSAs with populations less than one million persons, our predicted Y must be increasingly corrected downward, the smaller is the SMSA, in order to adjust for this size effect.

In like manner, equation Variant 3 is designed to test for a region effect. The same restrictions for independence hold here as in Variant 2; hence, for nine regions, eight dummy variables are specified. In order to substantiate the hypothesis that there is some region effect, it is merely necessary to demonstrate

TABLE 5

REGRESSION VARIANT 3—COEFFICIENTS AND SIGNIFICANCE TESTS FOR RELATION BETWEEN GINI'S CONCENTRATION RATIO AND 7 QUANTITATIVE VARIABLES PLUS REGION DUMMIES FOR 208 SMSAs, 1960

Variable	Regression Coefficient	Standard Error	t-value	Level of Significance	R^2 Without Variable
X_1	0.0594	0.0140	4.228	**	0.856
X_2	−0.3165	0.0239	−13.224	**	0.749
X_3	−0.0075	0.0014	−5.250	**	0.850
X_4	−0.0012	0.0007	−1.766	*	0.866
X_6[a]	−0.0042	0.0021	−2.010	*	0.866
X_7	−0.0121	0.0033	−3.620	**	0.859
X_8	0.0000019	0.0000010	1.980	*	0.866
X_{12}	−0.0014	0.0041	−0.337		
X_{13}	0.0005	0.0042	0.121		
X_{14}	0.0111	0.0053	2.105	*	
X_{15}	0.0204	0.0064	3.210	**	0.840
X_{16}	0.0141	0.0053	2.646	**	
X_{17}	−0.0055	0.0049	−1.118		
X_{18}	−0.0124	0.0059	−2.093	*	
X_{19}	−0.0034	0.0055	−0.624		
Y-intercept (a)	0.6379	0.0203	31.349	**	—

R^2, unadjusted = 0.868
R^2, adjusted = 0.858
Standard Error of Estimate = 0.0128
F-value (15; 192 d.f.) = 84.474**

[a]Measured in thousands of dollars.
*Significant at the 0.05 level.
**Significant at the 0.01 level.
X_1–X_4, X_6–X_8: Same as in Table 2.
X_{12}: Dummy for SMSA located in Middle Atlantic States.
X_{13}: Dummy for SMSA located in East North Central States.
X_{14}: Dummy for SMSA located in West North Central States.
X_{15}: Dummy for SMSA located in South Atlantic States.
X_{16}: Dummy for SMSA located in East South Central States.
X_{17}: Dummy for SMSA located in West South Central States.
X_{18}: Dummy for SMSA located in Mountain States.
X_{19}: Dummy for SMSA located in Pacific States.

that the dependent variable of SMSAs in some one (or more) region will not necessarily be well-estimated by an equation based on some other region (i.e., normally the base region is the one which is not coded—the nth). In Variant 3 the comparison (i.e., un-coded base) region is Region 1, the New England states.

Table 5 indicates by the t-values and regression coefficients for the region variables X_{12-19} that the strongest regional effects are to be found in Regions 5 and 6 (X_{15} and X_{16}), the Southern states. That is, an equation designed to estimate the dependent variable most accurately for SMSAs in the New England states, other things being equal, needs the greatest correction for estimates of the dependent variable in SMSAs in the South. This "correction" is implied in the equations by the relatively large and significant regression coefficients for these variables (disregarding sign).

TABLE 6

REGRESSION VARIANT 4—COEFFICIENTS AND SIGNIFICANCE TESTS FOR RELATION BETWEEN GINI'S CONCENTRATION RATIO AND 7 QUANTITATIVE VARIABLES PLUS POPULATION AND REGION DUMMIES FOR 208 SMSAs, 1960

Variable	Regression Coefficient	Standard Error	t-value	Level of Significance	R^2 Without Variable
X_1	0.0477	0.0132	3.603	**	0.883
X_2	−0.3230	0.0221	−14.590	**	0.767
X_3	−0.0077	0.0013	−5.785	**	0.871
X_4	−0.0015	0.0006	−2.361	**	0.887
X_6[a]	−0.0074	0.0020	−3.646	**	0.883
X_7	−0.0128	0.0031	−4.093	**	0.881
X_8	0.0000011	0.0000010	1.265		0.889
X_9	−0.0197	0.0035	−5.598	**	
X_{10}	−0.0127	0.0034	−3.767	**	0.860
X_{11}	−0.0085	0.0031	−2.760	**	
X_{12}	−0.0083	0.0039	−2.130	*	
X_{13}	−0.0010	0.0039	−0.253		
X_{14}	0.0053	0.0049	1.073		
X_{15}	0.1609	0.0059	2.714	**	0.868
X_{16}	0.0093	0.0050	1.882	*	
X_{17}	−0.0087	0.0046	−1.889	*	
X_{18}	−0.0128	0.0055	−2.345	**	
X_{19}	−0.0108	0.0052	−2.087	*	
Y-intercept (a)	0.6819	0.0202	33.803	**	—

R^2, unadjusted = 0.890
R^2, adjusted = 0.880
Standard Error of Estimate = 0.0118
F-value (18; 189 d.f.) = 85.160**

[a]Measured in thousands of dollars.
*Significant at the 0.05 level.
**Significant at the 0.01 level.
X_1 : Percentage of population nonwhite.
X_2 : Percentage of employed persons in middle-level occupations.
X_3 : Median years of education among persons 25 years and older.
X_4 : Ratio of family income as wages and salaries to property.
X_6 : Median family income.
X_7 : Average hourly wage for production workers in manufacturing.
X_8 : Number of persons per square mile of land area.
X_9 : Dummy for SMSA with population between 50,000 and 150,000.
X_{10}: Dummy for SMSA with population between 150,000 and 250,000.
X_{11}: Dummy for SMSA with population between 250,000 and 1 million.
X_{12}: Dummy for SMSA located in Middle Atlantic States.
X_{13}: Dummy for SMSA located in East North Central States.
X_{14}: Dummy for SMSA located in West North Central States.
X_{15}: Dummy for SMSA located in South Atlantic States.
X_{16}: Dummy for SMSA located in East South Central States.
X_{17}: Dummy for SMSA located in West South Central States.
X_{18}: Dummy for SMSA located in Mountain States.
X_{19}: Dummy for SMSA located in Pacific States.

Using t-values as a quantitative measure,[27] the relative significance of the independent variables can be seen to fluctuate very little as a function of the presence or absence of the dummy variables. For example, as Table 7 shows, in equation Variant 1, although nearly all the coefficients were significant at the 0.01 level, the most significant coefficients were those of X_2, the percentage of persons employed in middle-level occupations, and X_1, the percentage of the population who are nonwhite. The remaining variables in Variant 1 could be viewed as falling into two additional subsets: X_3, X_6 and X_7; and X_4 and X_8—moderate and lower significance respectively.

TABLE 7

t-STATISTICS OF QUANTITATIVE VARIABLES FOR EQUATION VARIANTS 1–4

Variable	Variant 1	Variant 2	Variant 3	Variant 4
X_1	9.442	8.323	4.228	3.603
X_2	−14.405	−15.017	−13.224	−14.590
X_3	−5.653	−5.780	−5.250	−5.785
X_4	−2.067	−2.610	−1.766	−2.361
X_6	−3.374	−4.740	−2.010	−3.646
X_7	−4.479	−4.771	−3.620	−4.093
X_8	1.899	1.190	1.980	1.265

Source: Tables 3, 4, 5 and 6.
Variable identities as in Table 2.

Table 7 also shows that in Variant 2, when the population dummies are added, the relative importance of the exogenous variables is little changed. In fact, although the t-statistics do fluctuate slightly, the rank order of the variables is identical for Variants 1 and 2. The minor changes which do occur are not so large as to influence the interpretation of the regression coefficients significantly; hence most of the effect of adding population variables to the equation may be considered pure extra explanatory power. This is not so much the case in Variant 3, where region effects alone are added. In this instance two major changes are observed: First, the t-statistic for X_1 drops sharply to 4.228, from 9.442 in Variant 1. Also the t-statistic for X_6 drops to less than two-thirds of its value in Variant 1. From this one may infer that much of the increment in explanatory power obtained by the addition of region dummies is achieved at a cost of significance in variables X_1 and X_6. In other words, the distribution of nonwhites in SMSAs in different regions, plus the general variation among regions in level of development, account for some part of the explanatory power of the region dummies, although not all of it. Furthermore, the relative stability of the magnitude of the other t-statistics, given the high partial correlation of the region dummies on top of an already high R-square, is indicative of their non-redundancy.

In Variant 4, through the inclusion of both sets of dummies and their interactions with each other and the quantitative variables, the final relationships

[27] t-values, rather than the often-used "beta values," are being used here due to the difficulty of attributing meaning to the "standardised units" with which beta values are concerned. It may be of incidental interest to note, however, that the beta values in fact correspond rather faithfully in rank order to the t-statistics in this study.

are observed. In this case, as shown in Table 7, X_1 is now diminished to a relatively low t-value, X_8 is not significant at all, and X_2 is a relatively super-significant -14.590. In other words, a certain amount of the pure effect of occupational mix, which was obscured before by the absence of the population and region variables, is now revealed. Simultaneously, the close relationship among nonwhite presence, population and region is similarly revealed, but in a non-clarifying manner. Underlying some unspecified amount of the explanatory power of population and region is very simply the distribution of nonwhites.

VI. Concluding Remarks

On the basis of the results outlined in the preceding section, several conclusions may be drawn regarding SMSA-to-SMSA variation in family income inequality: First, it has been shown that a relatively simple model employing seven "quantitative" variables and two sets of dummy variables is capable of explaining 89 percent of this variation. Second, we have been able to identify those factors which are most closely associated with this variation. Finally, within this family of associated factors the incremental estimating technique used has revealed certain relevant interactions, the recognition of which may be useful for policy purposes. This last assumes its importance only insofar as it is accepted by government(s) that income inequality may be a surrogate for welfare and that it wishes to pull the relevant economic, social, and institutional policy strings in order to enhance this condition, however defined.

Bibliography

Aigner, D. J. and A. J. Heins, "On the Determinants of Income Inequality". *American Economic Review*: Vol. 57, No. 1 (March 1967), pp. 175–81.

Al-Samarrie, A. and H. P. Miller, "State Differentials in Income Concentration". *American Economic Review*: Vol. 57, No. 1 (March 1967), pp. 59–72.

Ashenfelter, O. and A. Rees, *Discrimination in Labor Markets*. Princeton: Princeton University Press, 1973.

Atkinson, A. B., "On the Measurement of Inequality". *Journal of Economic Theory*: Vol. 2, No. 3 (September 1970), pp. 244–63.

Bowman, M. J., "A Graphical Analysis of Personal Income Distribution in the United States". *American Economic Review*: Vol. 35, No. 4 (September 1945), pp. 607–28.

Burns, L. S. and H. E. Frech, "Human Capital and the Size Distribution of Income in Dutch Cities". *De Economist*: Vol. 118, No. 6 (November-December 1970), pp. 598–618.

Chiswick, B. R., "Racial Discrimination in the Labor Market: A Test of Alternative Hypotheses". *Journal of Political Economy*: Vol. 81, No. 6 (November-December 1973), pp. 1330–52.

Farbman, M., "Income Concentration in the Southern United States". *Review of Economics and Statistics*: Vol. 55, No. 3 (August 1973), pp. 333–40.

Fitzwilliams, J. M., "Size Distribution of Income in 1963". *Survey of Current Business*: Vol. 43, No. 4 (April 1964), pp. 3–12.

Frech, H. E. and L. S. Burns, " 'Metropolitan Interpersonal Income Inequality': A Comment". *Land Economics*: Vol. 47, No. 1 (February 1971), pp. 104–6.

Gwartney, J., "Employment Discrimination, Productivity Factors, and Income Differentials Between Whites and Nonwhites". *American Economic Review*: Vol. 61, No. 3 (June 1970), pp. 396–408.

Haley, B. F., "Income Distribution in the U.S.", in Jean Marchal and Bernard Ducros, eds., *The Distribution of National Income*. New York: St. Martin's Press, 1968.

Hoyt, H., *Where the Rich and the Poor People Live*. New York: Urban Land Institute, 1966. Bulletin No. 55.

Kmenta, J., *Elements of Econometrics*. New York: Macmillan, 1971.

Kravis, I. B., "International Differences in the Distribution of Income". *Review of Economics and Statistics*: Vol. 42, No. 4 (November 1960), pp. 408–16.
Kravis, I. B., *The Structure of Income, Some Quantitative Essays*. Philadelphia: University of Pennsylvania Press, 1962.
Kuznets, S., "Economic Growth and Income Inequality". *American Economic Review*: Vol. 45, No. 1 (March 1955), pp. 1–28.
Kuznets, S., "Quantitative Aspects of the Economic Growth of Nations: VIII. Distribution of Income by Size". *Economic Development and Cultural Change*: Vol. 11, No. 2, part II (January 1963), pp. 1–80.
Mattila, J. M. and W. R. Thompson, "Toward an Econometric Model of Urban Economic Development", in Harvey Perloff and Lowdon Wingo, eds., *Issues in Urban Economics*. Baltimore: Johns Hopkins Press, 1968.
Miller, H. P., *Trends in the Income of Families and Persons in the United States: 1947 to 1960*. Washington: G.P.O., 1963. U.S. Bureau of the Census Technical Paper No. 8.
Murray, B. B., "Coefficient of Interarea Variation as a Measure of Spatial Income Inequality". *Journal of the American Statistical Association*: Vol. 65, No. 330 (June 1970), pp. 598–601.
Murray, B. B., "Metropolitan Interpersonal Income Inequality". *Land Economics*: Vol. 45, No. 1 (February 1969), pp. 121–25.
Newberry, D., "A Theorem on the Measurement of Inequality". *Journal of Economic Theory*: Vol. 2, No. 3 (September 1970), pp. 264–66.
Newhouse, J. P., "A Simple Hypothesis of Income Distribution". *Journal of Human Resources:* Vol. 6, No. 1 (Winter 1971), pp. 51–74.
Projector, D. S., G. S. Weiss, and E. T. Thoresen, "Composition of Income as Shown by the Survey of Financial Characteristics of Consumers", in Lee Soltow, ed., *Six Papers on the Size Distribution of Income and Wealth*, Studies in Income and Wealth, Vol. 33. New York: National Bureau of Economic Research, 1969.
Reder, M. W., "A Partial Survey of the Theory of Income Size Distribution", in Lee Soltow, ed., *Six Papers on the Size Distribution of Income and Wealth*, Studies in Income and Wealth, Vol. 33. New York: National Bureau of Economic Research, 1969.
Report of the Presidential Advisory Commission on Civil Disorders. New York: Bantam Books, 1968.
Scully, G. W., "Interstate Wage Differentials: A Cross-Section Analysis". *American Economic Review*: Vol. 59, No. 5 (December 1969), pp. 757–73.
Sen, A. K., *On Economic Inequality*. Oxford: Clarendon Press, 1973.
Soltow, L., "The Distribution of Income Related to Changes in the Distribution of Education, Age and Occupation". *Review of Economics and Statistics*: Vol. 42, No. 4 (November 1960), pp. 450–53.
Stark, T., *The Distribution of Personal Income in the United Kingdom*. Cambridge: Cambridge University Press, 1972.
Thompson, W. R., "Urban Economic Development", in Werner Z. Hirsch, ed., *Regional Accounts for Policy Decisions*. Baltimore: Johns Hopkins Press, 1966.
U.S. Bureau of the Census. *U.S. Census of Population, 1960*.
U.S. Department of Commerce. Office of Business Economics. *Income Distribution in the United States by Size, 1944–50*. Washington: G.P.O., 1953.
Weisskoff, R., "Income Distribution and Economic Growth in Puerto Rico, Argentina, and Mexico". *Review of Income and Wealth*: Series 16, No. 4 (December 1970), pp. 303–32.
Williamson, J. G., "Regional Inequality and the Process of National Development: A Description of the Patterns". *Economic Development and Cultural Change*: Vol. 13, No. 4, part II (July 1965), pp. 3–84.
Yntema, D., "Measures of the Inequality in the Personal Distribution of Wealth or Income". *Journal of the American Statistical Association*: Vol. 28 (December 1933), pp. 423–33.