

Income Inequality In Depressed Areas: A Principal Component Approach

WILLIAM D. GUNTHER* AND C. MICHAEL ELLIS**

I.

INTRODUCTION

A large body of literature has been accumulating which supports the hypothesis that an inverse relationship exists between the level of development and the degree of inequality present in the size distribution of income. [3, 12, 13, 18] This relationship, generally referred to as the "Kuznets Hypothesis" has been offered as a long-run phenomena, although sufficient exceptions have been noted to temper any tendency to assume that growth brings *pari passu* improved equity. [9, 17, 21] Given the possibility of short and long-run exceptions to the Kuznets Hypothesis, growth may distort the existing distribution of income in a variety of ways. Thus it seems that a cost/benefit question regarding the growth-distribution nexus seems to be an appropriate avenue of inquiry. Even in cases where growth can be expected to reduce inequality, alternative growth policies may have alternative distributional impacts which add a new dimension to policy analysis. [15, 16]

Although there exists a considerable amount of literature on the development-distribution question, the level of analysis frequently relates to cross-cultural comparisons and inter-state variations. [2, 5, 7, 20] The rapid development of regional economics has largely skirted the policy aspects of the distributional question although the literature implies that depressed or lagging regions would, other things being equal, have a greater distributional "problem" than the more affluent regions. [22] Thus, the distribution of income and its relationship to policy variables would seem to be as relevant an issue in depressed area economic planning as it is in developing countries, particularly when we are dealing with such large areas as Appalachia and the Ozark region.

The purpose of this paper is to examine the relationship between distributional equity and the level of development in a depressed region of the United States. The findings of the paper shed some light on the degree and direction in which distributional equity can be expected to vary and suggests the need for greater concern with the distributional consequences of accelerated development programs.

*Professor, The University of Alabama.

**Assistant Professor, The University of Central Arkansas.

The specific questions posed in this paper are: (1) In comparison with the nation, is the distribution of income more or less equally distributed in a depressed region such as the Ozarks?; and (2) Is the relationship between development and income inequality in the Ozark along the lines implied by the Kuznets Hypothesis?

II.

ESTIMATING THE SIZE DISTRIBUTION OF INCOME

To examine the distribution-development relationship, a cross-sectional analysis of the 340 counties of the Ozark region for the Census year 1970 was undertaken. At the county level, complete information of the size distribution of income does not exist and consequently was estimated. This was accomplished by using the distribution of families by income class and estimating the mean income for each class. The means of each interval were estimated by assuming a particular distribution of families within each interval, and based on these density functions, means were derived. In the first interval, (\$0-\$999), families were assumed to be distributed in a parabolic fashion, clustering toward the upper end of the interval. Families in nine income classes above the first interval are assumed to follow a linear distribution, while families in income classes of \$10,000 or more were assumed to follow the familiar Pareto distribution. [14] A summary measure of these distributions was then sought.

The selection of specific summary measures of the size distribution was a considerable undertaking in itself, complicated by the fact that different measures weight various parts of a distribution differently. The choice was made to select four inequality measures which have appeared with some frequency in the literature. The measures chosen were the Gini concentration ratio, the relative mean deviation, the coefficient of variation, and the standard deviation of common logarithms. These four measures were calculated for each of the counties of the Ozark region, for the region as a whole, and for the United States in 1970.

III.

MEASURING THE LEVEL OF DEVELOPMENT

A prerequisite for testing the hypothesis that development and inequality are inversely related is the selection of a method whereby different levels of economic development can be measured and distinguished from one another. In previous studies this need has been partially met by the arbitrary selection of proxy measures of development — in most cases per capita, mean, or median income. This proxy approach to measuring development can be criticized on the grounds that it fails to take into account the multi-dimensional nature of economic development. The multi-dimensional approach to measuring development has steadily gained converts in recent years and has been aided immeasurably by the

growing popularity and application of principal component and factor analytic methods. [1, 4, 6, 10, 11] This study takes this broader and more explicit account of the multi-dimensional nature of economic development and its relationship to income inequality by adopting a principal components approach to measuring levels of development.

The measurement of Ozark county development in 1970 utilizes a principal components analysis of twenty-seven selected economic, social, and demographic characteristics for each county. Census data in most cases served as the input source. A listing of these variables and their symbols is shown in Table 1.

The principal components technique is especially useful for studies which involve rather large numbers of variables since the number of principal components extracted is usually fewer than the number of original variables. In addition, the principal components are orthogonal — they are uncorrelated and are statistically independent of each other

TABLE 1

TWENTY-SEVEN DEVELOPMENT VARIABLES AND THEIR SYMBOLS

Variable	Symbol
1. Percent of families with poverty level income	PCTPOV
2. Percent of families with income of \$15,000 or more	PCT150VR
3. Median family income	MEDIAN
4. Percent of population urban	URBAN
5. Total population percent change, 1960-1970	TOTPOPPC
6. Total urban population change, 1960-1970	URBANPCT
7. Total rural population percent change, 1960-1970	RURALPCT
8. Percent net migration, 1960-1970	PCTMIG
9. Percent of fourteen to seventeen year olds in school	PCTSCH
10. Median school years completed, males age twenty-five and over	MEDSCH
11. Nonworker-worker ratio	WRKRAT
12. Percent of employed females, age sixteen and over	PCT16
13. Percent of employed males, age sixty-five and over	PCT650VR
14. Percent civilian labor force unemployed	PCTUNPL
15. Percent of workers in primary industries	PCTPRMI
16. Percent of workers in manufacturing	PCTMFG
17. Percent of workers in white-collar occupations	PCTWCL
18. Percent of workers who worked fifty to fifty-two weeks in 1969	PCTWKRYR
19. Percent of families with wage and salary income	PCTWS
20. Percent of families with welfare income	PCTWEL
21. Percent of families with "other" income	PCTOTR
22. Percent of families with non-farm self-employment income	PCTNFRM
23. Percent of families with farm self-employment income	PCTFRMI
24. Percent of poverty families lacking some or all plumbing facilities	PCTPLM
25. Percent of county acreage in farms	PCTFRM
26. Value-added by manufacturing (1967)	VALADD
27. New capital expenditures in manufacturing (1967)	CAPEXP

— and they are unique linear combinations of the original variables. In short, a data set of a large number of variables is described as completely and efficiently as possible in terms of a given number of independent principal components.

After all the variables have been standardized, a correlation matrix is produced and from this, an eigenvector matrix is derived. An eigenvector corresponds to a solution of the characteristic equation of the correlation matrix. Each element in the eigenvector matrix is converted into a corresponding element of the factor matrix which is thereafter used to give interpretation to the components which emerge from the data.

Recalling that principal components are linear combinations of the standardized variables, the component scores for each county were constructed in the following manner: for each county, the component score for the first principal component was the summation of the products obtained by multiplying the value of the county's twenty-seven standardized variables by each variable's corresponding element in the first eigenvector. For the second component score, the standardized variables were multiplied by each corresponding element in the second eigenvector and summed. This process was repeated until each county had a predetermined number of component scores. These component scores were thereafter weighted and summed to yield a composite or overall score for each county. This composite number is then interpreted as the economic development "index." The mean value of all the counties' component scores is zero; therefore, a county which has an economic development index of zero can be considered "average" for the region in terms of economic development. Greater positive values of the index reflect higher levels of development, while negative values indicate relatively lower levels of development.

In symbolic terms, component scores for individual counties can be described in the following manner:

For p number of variables and m number of components and eigenvectors:

$$C_1 = a_{11}Z_1 + a_{12}Z_2 + \dots + a_{1p}Z_p$$

$$C_2 = a_{21}Z_1 + a_{22}Z_2 + \dots + a_{2p}Z_p$$

·
·
·

$$C_m = a_{m1}Z_1 + a_{m2}Z_2 + \dots + a_{mp}Z_p$$

where: C_i = the value of a component score for a given county,
 $i = 1, 2, \dots, m$;

Z_j = the standardized value of variables, $J = 1, 2, \dots, p$;

a_{ij} = the loading of eigenvector i with respect to variable Z_j ,
 $i = 1, 2, \dots, m$;
 $j = 1, 2, \dots, p$.

The economic development index for each county is expressed as:

$$I_j = w_1C_1 + w_2C_2 + w_mC_m$$

- where: I_j = a county's economic development index or composite score, $j = 1, 2, . . . 340$
- C_i = a component score, $i = 1, 2, . . . m$;
- w_i = the weight assigned to a particular component score, $i = 1, 2, . . . m$. [19, pp. 359-65]

IV.

RESEARCH FINDINGS

In this study eight principal components were successively extracted. Together they accounted for 83.3 per cent of the total variance of the twenty-seven development variables, resulting in a considerable economy of data with only minor sacrifices in explanatory power.

The eight major dimensions of development in the Ozark region are presented below in Table 3. These dimensions are ordered in their relative statistical importance as shown by their respective eigenvalues and cumulative percent of eigenvalues.

Interpretation of the components of Ozark development was largely successful in spite of the fact that an unrotated factor matrix was used to give subjective interpretation to the components. Generally, it is necessary to rotate the factor matrix until the components are more clearly interpretable. However, such rotation destroys the unique functional relationships provided by the unrotated principal components technique. Table 4 shows the unrotated factor matrix. The eight components were interpreted from the matrix on the basis of the magnitude of the loadings in the matrix.

TABLE 3
PRINCIPAL COMPONENTS OF OZARK DEVELOPMENT, 1970

Component	Eigenvalue	Cumulative Percent of Eigenvalues
1. Economic Resource Base	8.567	31.7
2. Primary Industry	5.342	51.5
3. Migration/Population Change	2.642	61.3
4. Manufacturing Employment/ Economic Stability	1.522	66.9
5. Manufacturing Base	1.258	71.6
6. Development Diversification	1.160	75.9
7. Urban Growth	1.000	79.6
8. Human Capital Investment	.990	83.3

TABLE 4
UNROTATED FACTOR MATRIX*

Symbol								
PCTPOV	.91208	.15536	— .11755	— .19427	.07696	— .02347	.01736	.03165
PCT150VR	— .80373	.17824	— .21323	— .05292	.22412	.08285	.07787	.13528
MEDIAN	— .89381	.19228	— .06128	.18611	.08242	.17515	.03690	.07208
URBAN	— .61081	.39589	— .42174	— .28547	— .11819	.07194	— .12265	— .12123
TOTPOPPC	— .36739	.54573	.67901	— .14989	.16845	.02974	.07031	.01605
URBANPCT	— .23760	.31895	.44681	— .34473	.25724	.27826	.38275	— .32880
RURALPCT	— .11317	.32531	.55279	.24879	— .00756	— .08259	— .16512	.38951
PCTMIG	— .38132	.32172	.79523	— .12947	.11984	— .10111	.06105	— .02574
PCTSCH	— .27354	— .18329	— .17656	.06391	.01973	.07632	.47993	.69329
MEDSCH	— .84851	— .19156	.04943	— .04632	— .05967	.05162	— .08299	— .01788
WRKRAT	.80010	.21726	— .09330	.01470	.26038	.05903	— .10346	.25401
PCT16	— .71632	.13173	— .23715	— .11956	— .39058	— .11295	.22188	— .06400
PCT650VR	— .55535	— .61387	— .10139	.04042	.10233	.16647	.16555	— .07786
PCTUNPL	.57737	.45310	— .17929	— .22111	— .10716	— .01335	.24449	.02196
PCTPRMI	.28467	— .80134	.01111	— .02120	.34489	.22085	.02868	— .03410
PCTMFG	.16482	.45054	.05987	.59895	— .38853	— .15015	.29854	— .21376
PCTWCL	— .70832	.34482	— .17850	— .42308	— .00884	— .10874	— .18094	.13625
PCTWKRYR	— .54373	— .20277	— .06298	.50314	.04720	.38633	— .25778	— .03394
PCTWS	— .52355	.64971	— .23833	— .00414	— .09353	.29045	.07429	.04613
PCTWEL	.76712	.29217	— .24220	— .24984	.10748	.01778	— .08979	.05843
PCTOTR	— .64970	— .36718	.04622	— .05169	.00481	— .39045	.07624	.12494
PCTNFRM	— .10534	— .70567	.22790	— .06775	— .09809	— .41617	.03389	.00714
PCTFRMI	.03248	— .91627	.12361	.02284	.13589	— .04587	.19687	— .10985
PCTPLM	.68535	.22525	— .10283	.14032	.02919	.19407	.39150	— .02874
PCTFRM	— .25550	— .71267	— .16447	— .31236	— .11152	.08614	.09294	.00580
VALADD	— .37274	.29702	— .40581	.17945	.43925	— .43551	.07870	— .12479
CAPEXP	— .32524	.36796	— .33912	.25296	.54720	— .26604	.04949	— .11693

*Components 1 and 6 were “reflected” in order to compute the indexes of development.

We hasten to point out that although the interpretation of the components is a subjective matter, there were in most cases such clear-cut indications of functional relationships in the variables that little imagination was needed to categorize them. For example, in the second component — Primary Industry — the variables which had the highest loadings in the unrotated factor matrix were the percent of employment in primary industries, percent of families with farm self-employment income, and percent of county acreage in farms. Similarly, in the fifth component, Manufacturing Base, the most significant variables were the percent of employment in manufacturing, new capital expenditure in manufacturing, and value-added by manufacturers. Thus, the multi-dimensional aspect of economic development was revealed by the principal components and incorporated into the indexes of development. Furthermore the relationships among the underlying variables and the interpretation of the components from those relationships were consistent with theoretical expectation and the reality of Ozark development.

The measurement of family income inequality presented fewer difficulties than those encountered in measuring economic development. The inequality of family income in the Ozark region in 1970 is indicated below in Table 5 where a comparison is made between family income inequality for the United States and the Ozark region.

Clearly, the Ozark region had greater family income inequality than the nation in 1970 as indicated by the larger values of the inequality measures for the region. This, in itself, is not surprising, since others have found a similar tendency for incomes to be more unequally distributed the lower the level of development (usually as measured by per capita income). The fact that the Ozark region's income was (and probably still is) more unequally distributed than the nation's suggests that the equity differences in the distribution of income between the two areas may be traceable to differences in their basic patterns of economic development. Indeed, the greater income inequality of the Ozark region may be a factor in efforts to raise income and employment levels up to the national average. Efforts to accelerate development in the region, for example, may intensify equity differences to a greater extent than that which already exists between the region and the rest of the country. A tradeoff of this sort would probably necessitate a policy

TABLE 5
COMPARATIVE MEASURES OF INCOME INEQUALITY
IN THE U. S. AND THE OZARK REGION, 1970

Area	Gini Ratio	Relative Mean Deviation	Coefficient of Variation	Standard Deviation of Logs
United States	.355	.497	.713	.348
Ozark Region	.383	.546	.774	.367

mixture which may not have been contemplated by regional policy-makers, if, indeed, avoiding greater equity differences is considered to be of positive social value.

Whether a tradeoff between greater development and a more equally distributed income is an unhappy prospect for the Ozark region is, of course, something which only time can tell. This is particularly true in view of the fact that the Ozark Regional Commission places primary emphasis on improved training and education of the area's inhabitants rather than on investment in traditional forms of economic infrastructure. [8, p. 117] Both the lack of adequate development funds and the time required to set up and coordinate the Commission's policies with state and local organizations have allowed only modest progress thus far in the region's development program. In short, the few years which have elapsed since the Commission's inception do not appear to be sufficiently long enough to determine whether the region's development program has itself fostered greater or lesser income inequality.

On the other hand, the prevailing patterns of economic development which have evolved over time within the region may provide at least a tentative appraisal of the likelihood of a future development-inequality trade-off. If we assume that a cross-section analysis of county development in 1970 incorporates significant growth determinants applicable to time series questions, correlation analysis may indicate the relationship between variables which, *ceteris paribus*, may provide at least a tentative working hypothesis.

Table 6 summarizes the correlation results of Ozark income inequality and economic development for the 340 counties of the region. Both zero-order correlation and Spearman rank correlation were employed.

Note from Table 6 that all four measures of inequality were inversely related to economic development, suggesting that at the county level, lower income inequality is associated with higher levels of development. Generalizing to time-series questions, it would follow that development in the Ozark region has apparently followed along the lines of more-developed nations rather than exhibiting greater inequality as development has risen. [9, pp. 70-84]

TABLE 6
DEVELOPMENT VS. INEQUALITY
(n = 340)

Inequality Measure	Zero-Order Correlation (r) with Index	Spearman Rank (r) with Index
Gini Ratio	— .644	— .67
Relative Mean Deviation	— .719	— .70
Coefficient of Variation	— .633	— .62
Standard Deviation of Logarithms	— .511	— .43

REFERENCES

1. Irma Adelman, and Cynthia T. Morris, "An Econometric Model of Socio-Economic and Political Change in Underdeveloped Countries," *American Economic Review*, 58: 1184-1281 (December, 1968).
2. A. Al-Samarrie, and Herman Miller, "State Differentials in Income Concentration," *American Economic Review*, 57: 59-72 (March, 1967).
3. Fuat M. Andic, *Distribution of Family Incomes in Puerto Rico: A Case Study of the Impact of Economic Development on Income Distribution*, Rio Piedras, Puerto Rico: University of Puerto Rico, 1964.
4. B. J. L. Berry, "Basic Patterns of Economic Development," *Atlas of Economic Development*, Chicago: University of Chicago Press.
5. John Conlisk, "Some Cross-State Evidence on Income Inequality," *Review of Economics and Statistics*, 49: 115-18 (February, 1967).
6. Ronald D. Gilbert, "Factor Analysis of Oklahoma Counties," *Oklahoma Business Bulletin*, 39: 9-15 (August, 1971).
7. Frank A. Hanna, "Analysis of Interstate Income Differentials: Theory and Practice," *Studies in Income and Wealth*, 21: 113-61, (1951).
8. Niles M. Hansen, *Rural Poverty and the Urban Crisis*, Bloomington, Indiana University Press, 1970.
9. Roger D. Hansen, *Mexican Economic Development*, Washington, National Planning Association, 1971.
10. Gerald Hodge, "The Identification of Growth Poles," *A Report to the Ontario Department of Economics and Development*, 1966.
11. Walson R. Kerns, and J. Dean Jansma, "Financial Need Priorities for Awarding Construction Grants-in-Aid," *Growth and Change*, 4: 17-22 (January, 1973).
12. Irving B. Kravis, "International Differences in Income Inequality," *Review of Economics and Statistics*, 44: 408-416, 1962.
13. Simon Kuznets, "Quantitative Aspects of the Economic Growth of Nations: VIII, Distribution of Income by Size," *Economic Development and Cultural Change*, 11: 1-80 (1963).
14. Maurice Liebenberg, and Hymann Kaitz, "An Income Size Distribution from Income Tax and Survey Data, 1944," *Studies in Income and Wealth*, 13 New York: National Bureau of Economic Research: 442-44, 1951.
15. V. C. Nwaneri, "Income Distribution and Project Selection," *Finance and Development*, 10: 27-29 (September, 1973).
16., "Income Distribution Criteria for the Analysis of Development Projects," *Finance and Development*, 10: 6-19,37 (March, 1973).
17. P. D. Ojha, and V. V. Bhatt, "Pattern of Income Distribution in an Underdeveloped Economy: A Case Study of India," *American Economic Review*, 54: 711-20 (September, 1964).
18. Harry Oshima, "The International Comparison of Size Distributions of Family Incomes with Special Reference to Asia," *Review of Economics and Statistics*, 44: 439, 1962.
19. R. J. Rummel, *Applied Factor Analysis*, Evanston, Northwestern University Press: 359-65, 1970.
20. Jeffrey Shaefer, "A Ranking of States by Inequality Using Census and Tax Data: A Comment," *Review of Economics and Statistics*, 49, No. 4: 636-37 (November, 1967).
21. Richard Weisskoff, "Income Distribution and Economic Growth in Puerto Rico, Argentina, and Mexico," *Review of Income and Wealth*, Series 16: 303-32 (December, 1970).
22. Jeffrey G. Williamson, "Regional Inequality and the Process of National Development," *Economic Development and Cultural Change*, Part 2: 3-84, (July, 1965).