DOCTORAL DISSERTATION SERIES

Publication No.: 21, 366

AUTHOR: David Wright Varley, Ph. D., 1956
University of Michigan

TITLE: A QUANTITATIVE ANALYSIS OF REGIONALISM IN THE UNITED STATES, 1940

University Microfilms, Ann Arbor, Michigan

A QUANTITATIVE ANALYSIS OF REGIONALISM IN THE UNITED STATES, 1940

by
David W. Varley

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy in the
University of Michigan
1956

Doctoral Committee:

Professor Amos H. Hawley, Chairman
Professor Robert C. Angell
Professor Paul S. Dwyer
Associate Professor Ronald Freedman
Associate Professor Leslie Kish

Abstract

A QUANTITATIVE ANALYSIS OF REGIONALISM
IN THE UNITED STATES, 1940

by David W. Varley

This study of regionalism in the United States is addressed to two main questions. First, to what extent may regionalism be regarded as an independent or "explanatory" variable for human behavior? Second, to what extent may this independent character of regionalism be attributed to structural and socio-economic features of regional society? These questions are derived from the work of regional scholars whose research has produced a large body of material describing regional differences appearing in the United States.

Unfortunately, most of this descriptive material takes the form of only summary measures for the several regions. The question of how representative these summary measures are of entire regional areas is usually not answered explicitly. This is particularly unfortunate because the essence of the concept "region" involves homogeneity within territorial areas as well as heterogeneity among these areas. The research undertaken in this study treats both these aspects of the regional concept.

The study focuses on data gathered for the year 1940. The regional delineation selected for investigation is that proposed by the late Professor Howard W. Odum. This classifies the country into six regions, each of which consists of four or more entire states. The basic unit of measurement is the State economic area. These areas are groups of counties, so that in effect each region represents a population of State economic areas.

The dependent variables of this study are five types of behavior: viz., fertility, educational achievement, religious affiliation, and two indexes of political voting behavior. Measures of these types of behavior are tabulated for each State economic area. In addition, each State economic area is classified by thirteen separate structural and socio-economic characteristics. These characteristics are used in the study as independent control variables and include such items as metropolitan status, rural farm standard of living, proportion of employed workers engaged in agriculture, and per capita dollar retail sales.

The statistical measure employed in the analysis is the intraclass correlation (rho). This analysis of variance technique indicates what proportion of the total variance of each type of behavior examined can be attributed to the regional locations of that behavior. The procedure is as follows: (a) Intraclass correlations are computed for each dependent variable and for the total number of State economic areas in all regions. (b) Control is then exercised by selecting from within each region those State economic areas exhibiting some specified characteristic or group of characteristics. (c) Intraclass correlations are recomputed for these selected control groups. It is expected that the correlations computed for the control groups will be smaller than those computed for the total group of State economic areas. The extent to which this reduction in correlation value does occur under controlled

conditions is interpreted as a measure of the extent to which regional behavior differences can be attributed to regional differences in social structure and socio-economic characteristics.

The general findings can be summarized briefly. The Odum regions are most distinctive with respect to the political behavior variables, but have negligible significance for fertility. None of the variables studied shows each region exhibiting a unique behavior pattern. Rather, regional distinctiveness in behavior appears to rest on the "deviant" position of only one or two regions while the other regions are little differentiated from one another. Generally speaking, controls for structural and socio-economic characteristics do effect a leveling of regional differences in behavior. This is particularly evident in the more urban or metropolitan areas. This suggests that as our society becomes more urbanized and industrialized regional behavior differences will either disappear or else be exhibited only as features of a regional division of labor.

PREFACE

A relatively quick and easy understanding of the material in this report can be gained by reading the first two chapters and the summaries at the end of each of the other chapters. This might also be profitably supplemented by reading the first section of Chapter IV and Chapter V for information on control procedures.

This report and the research upon which it is based are the product of many hands. Among those who provided helpful counsel and advice were John Aird, Hubert Blalock, Robert Hamblin, Gerhard Lenski, Leo Schnore, and Robert Schulze.

The members of the committee in charge of the study gave prompt and valuable assistance whenever the author called upon them. In this regard special acknowledgement is due Professor Leslie Kish, who bore the brunt of the author's frequent solicitation for statistical advice. His patience and interest were much appreciated.

To his wife the author owes a debt of gratitude beyond calculation. Her assistance was given at every stage of the research and the manuscript preparation.

TABLE OF CONTENTS

Page

PREFACE ii

LIST OF TABLES. v

LIST OF APPENDIX TABLES x

Chapter

I. THE PROBLEM AND ITS SETTING. 1

Summary

II. THE PROCEDURE. 14

Scope and Focus of the Study
Selection of Basic Units and the Problem of Control
Selection of a Regional Delineation
Selection of Dependent Variables
Selection of Independent Variables for Control
Collection of Data and Statistical Measurement
Summary

III. THE FINDINGS: THE UNIVERSE OF STATE ECONOMIC AREAS AND A COMPARISON OF ITS METROPOLITAN AND NONMETROPOLITAN SEGMENTS 38

Comparison of Regional Means
The Intraclass Correlation
Components of the Intraclass Correlation
Comparison of Internal Homogeneity among Regions
The Influence of Structure on Regional Rank Positions
Summary

IV. THE FINDINGS: METROPOLITAN STATE ECONOMIC AREAS WITH APPLIED CONTROLS. 79

Preliminary Remarks and the Specification of Controls
Net Reproduction Rate
Proportion of Eligible Voters Voting
Proportion of Voters Voting Democratic
Educational Achievement
Religious Affiliation
Summary and Evaluation

Chapter Page

V. THE FINDINGS: NONMETROPOLITAN STATE ECONOMIC AREAS WITH APPLIED CONTROLS 141

Preliminary Remarks and the Specification of Controls
Net Reproduction Rate
Proportion of Eligible Voters Voting
Proportion of Voters Voting Democratic
Educational Achievement
Religious Affiliation
Summary and Evaluation

VI. SUMMARY AND CONCLUSIONS. 195

The Problem, Research Design, and General Findings
Evaluation and Suggestions for Further Research

APPENDIX A 214

APPENDIX B 228

BIBLIOGRAPHY 242

LIST OF TABLES

Table		Page
1.	Distributions of Regional Means for Fertility by Various State Economic Area Groupings	39
2.	Distributions of Regional Means for "Proportion of Eligible Voters Voting" by Various State Economic Area Groupings	41
3.	Distributions of Regional Means for "Proportion of Voters Voting Democratic" by Various State Economic Area Groupings	43
4.	Distributions of Regional Means for "Proportion of Males Age 25 Years Plus with No Formal Schooling" by Various State Economic Area Groupings.	44
5.	Distributions of Regional Means for "Proportion of Population with Religious Affiliation" by Various State Economic Area Groupings	46
6.	Values of Rho for Regional Distributions of Selected Variables by State Economic Area Groupings	48
7.	Values of Interregional and Intraregional Variance for Distributions of Selected Variables by Various State Economic Area Groupings	53
8.	Proportion of Total Population in Each Region which is Roman Catholic, by Various State Economic Area Groupings	57
9.	Standard Deviations of Selected Variables for Each Region and for the United States by State Economic Area Groupings.	61
10.	Ratios of Regional Standard Deviations for Selected Variables to Average Regional Standard Deviation for all Regions by State Economic Area Groupings	66
11.	Ratios of Regional Standard Deviations for Selected Variables to Standard Deviation for Entire United States by State Economic Area Groupings	69

Table		Page
12.	Rank Positions of Regions with Respect to Mean Values and Standard Deviations for Each Variable by Metropolitan and Nonmetropolitan Groupings	71
13.	Means and Sigmas of Regional Distributions of Intraregional Standard Deviations for Selected Variables by Metropolitan and Nonmetropolitan Areas	74
14.	Number of Metropolitan State Economic Areas in Each Region for the Total Group of Metropolitan State Economic Areas and Twelve Selected Control Groups.	87
15.	Minimum Values of Rho for Significance Levels of Five Per Cent and One Per Cent; Absolute Minimum Values of Rho; Total N for Total Group of Metropolitan State Economic Areas and Twelve Selected Control Groups.	88
16.	A Listing for Fertility of the Values of Rho, the Interregional Variance, the Intraregional Variance, and N for the Total Group of Metropolitan State Economic Areas and Twelve Selected Control Groups . .	90
17.	Distributions of Regional Means for Fertility by the Total Group of Metropolitan State Economic Areas and Twelve Selected Control Groups.	92
18.	Distributions of Regional Standard Deviations for Fertility by the Total Group of Metropolitan State Economic Areas and Twelve Selected Control Groups . .	93
19.	A Listing for Proportion of Eligible Voters Voting of the Values of Rho, the Interregional Variance, the Intraregional Variance, and N for the Total Group of Metropolitan State Economic Areas and Twelve Selected Control Groups.	100
20.	Distributions of Regional Means for Proportion of Eligible Voters Voting by the Total Group of Metropolitan State Economic Areas and Twelve Selected Control Groups	102
21.	Distributions of Regional Standard Deviations for Proportion of Eligible Voters Voting by the Total Group of Metropolitan State Economic Areas and Twelve Selected Control Groups.	103
22.	Mean Values and Standard Deviations for Proportion of Eligible Voters Voting in the Metropolitan State Economic Areas of the Southeast Region by the Proportion of the Total Population which is Nonwhite.	106

Table | Page

23. A Listing for Proportion of Voters Voting Democratic of the Values of Rho, the Interregional Variance, the Intraregional Variance, and N for the Total Group of Metropolitan State Economic Areas and Twelve Selected Control Groups 109

24. Distributions of Regional Means for Proportion of Voters Voting Democratic by the Total Group of Metropolitan State Economic Areas and Twelve Selected Control Groups 111

25. Distributions of Regional Standard Deviations for Proportion of Voters Voting Democratic by the Total Group of Metropolitan State Economic Areas and Twelve Selected Control Groups. 113

26. Mean Values and Standard Deviations for Proportion of Voters Voting Democratic in the Metropolitan State Economic Areas of the Southeast Region by the Proportion of the Total Population which is Nonwhite. 115

27. A Listing for Education of the Values of Rho, the Interregional Variance, the Intraregional Variance, and N for the Total Group of Metropolitan State Economic Areas and Twelve Selected Control Groups . . 117

28. Distributions of Regional Means for Education by the Total Group of Metropolitan State Economic Areas and Twelve Selected Control Groups. 119

29. Distributions of Regional Standard Deviations for Education by the Total Group of Metropolitan State Economic Areas and Twelve Selected Control Groups . . 122

30. A Listing for Religious Affiliation of the Values of Rho, the Interregional Variance, the Intra-regional Variance, and N for the Total Group of Metropolitan State Economic Areas and Twelve Selected Control Groups 126

31. Distributions of Regional Means for Religious Affiliation by the Total Group of Metropolitan State Economic Areas and Twelve Selected Control Groups. 128

32. Distributions of Regional Standard Deviations for Religious Affiliation by the Total Group of Metro-politan State Economic Areas and Twelve Selected Control Groups. 131

33. Number of Nonmetropolitan State Economic Areas in Each Region for the Total Group of Nonmetropolitan State Economic Areas and Ten Selected Control Groups. 145

Table		Page
34.	Minimum Values of Rho for Significance Levels of Five Per Cent and One Per Cent; Absolute Minimum Values of Rho; Total N for Total Group of Nonmetropolitan State Economic Areas and Ten Selected Control Groups	147
35.	A Listing for Fertility of the Values of Rho, the Interregional Variance, the Intraregional Variance, and N for the Total Group of Nonmetropolitan State Economic Areas and Ten Selected Control Groups . . .	149
36.	Distributions of Regional Means for Fertility by the Total Group of Nonmetropolitan State Economic Areas and Ten Selected Control Groups.	151
37.	Distributions of Regional Standard Deviations for Fertility by the Total Group of Nonmetropolitan State Economic Areas and Ten Selected Control Groups	153
38.	A Listing for Proportion of Eligible Voters Voting of the Values of Rho, the Interregional Variance, the Intraregional Variance, and N for the Total Group of Nonmetropolitan State Economic Areas and Ten Selected Control Groups.	156
39.	Distributions of Regional Means for Proportion of Eligible Voters Voting by the Total Group of Nonmetropolitan State Economic Areas and Ten Selected Control Groups	158
40.	Distributions of Regional Standard Deviations for Proportion of Eligible Voters Voting by the Total Group of Nonmetropolitan State Economic Areas and Ten Selected Control Groups.	161
41.	Mean Values and Standard Deviations for Proportion of Eligible Voters Voting in the Nonmetropolitan State Economic Areas of the Southeast Region by the Proportion of the Total Population which is Nonwhite	163
42.	A Listing for Proportion of Voters Voting Democratic of the Values of Rho, the Interregional Variance, the Intraregional Variance, and N for the Total Group of Nonmetropolitan State Economic Areas and Ten Selected Control Groups.	166
43.	Distributions of Regional Means for Proportion of Voters Voting Democratic by the Total Group of Nonmetropolitan State Economic Areas and Ten Selected Control Groups.	168

Table		Page
44.	Distributions of Regional Standard Deviations for Proportion of Voters Voting Democratic by the Total Group of Nonmetropolitan State Economic Areas and Ten Selected Control Groups	170
45.	Mean Values and Standard Deviations for Proportion of Voters Voting Democratic in the Nonmetropolitan State Economic Areas of the Southeast Region by the Proportion of the Total Population which is Nonwhite.	172
46.	A Listing for Education of the Values of Rho, the Interregional Variance, the Intraregional Variance, and N for the Total Group of Nonmetropolitan State Economic Areas and Ten Selected Control Groups. . . .	174
47.	Distributions of Regional Means for Education by the Total Group of Nonmetropolitan State Economic Areas and Ten Selected Control Groups	176
48.	Distributions of Regional Standard Deviations for Education by the Total Group of Nonmetropolitan State Economic Areas and Ten Selected Control Groups.	178
49.	A Listing for Religious Affiliation of the Values of Rho, the Interregional Variance, the Intraregional Variance, and N for the Total Group of Nonmetropolitan State Economic Areas and Ten Selected Control Groups	181
50.	Distributions of Regional Means for Religious Affiliation by the Total Group of Nonmetropolitan State Economic Areas and Ten Selected Control Groups.	185
51.	Distributions of Regional Standard Deviations for Religious Affiliation by the Total Group of Nonmetropolitan State Economic Areas and Ten Selected Control Groups.	187

LIST OF APPENDIX TABLES

APPENDIX A

Table		Page
1.	Frequency Distributions of Metropolitan State Economic Areas for Each Region by the Proportion of the Total Population which is Urban	215
2.	Frequency Distributions of Metropolitan State Economic Areas for Each Region by the Proportion of the Total Population which is Rural Nonfarm . . .	216
3.	Frequency Distributions of Metropolitan State Economic Areas for Each Region by the Proportion of the Total Population which is Rural Farm.	217
4.	Frequency Distributions of Metropolitan State Economic Areas for Each Region by the Proportion of the Total Population which is Nonwhite.	218
5.	Frequency Distributions of Metropolitan State Economic Areas for Each Region by the Proportion of Employed Persons Working in Agriculture	219
6.	Frequency Distributions of Metropolitan State Economic Areas for Each Region by the Proportion of the Employed Labor Force in Manufacturing	220
7.	Frequency Distributions of Metropolitan State Economic Areas for Each Region by the Proportion of the Employed Labor Force in Wholesale and Retail Trade	221
8.	Frequency Distributions of Metropolitan State Economic Areas for Each Region by Per Capita Dollar Retail Sales.	222
9.	Frequency Distributions of Metropolitan State Economic Areas for Each Region by Per Capita Dollar Receipts from Services.	223
10.	Frequency Distributions of Metropolitan State Economic Areas for Each Region by the Rural Farm Level of Living Index.	224

Table | Page

11. Frequency Distributions of Metropolitan State Economic Areas for Each Region by the Proportion of the Rural Farm Labor Force who are Unpaid Family Workers 225

12. Frequency Distributions of Metropolitan State Economic Areas for Each Region by the Proportion of the Rural Farm Labor Force in Nonfarm Occupations. 226

13. Percentage Distributions of Metropolitan State Economic Areas among Regions for the Total Group of Metropolitan State Economic Areas and Twelve Selected Control Groups. 227

APPENDIX B

1. Frequency Distributions of Nonmetropolitan State Economic Areas for Each Region by the Proportion of the Total Population which is Urban 229

2. Frequency Distributions of Nonmetropolitan State Economic Areas for Each Region by the Proportion of the Total Population which is Rural Nonfarm . . . 230

3. Frequency Distributions of Nonmetropolitan State Economic Areas for Each Region by the Proportion of the Total Population which is Rural Farm. 231

4. Frequency Distributions of Nonmetropolitan State Economic Areas for Each Region by the Proportion of the Total Population which is Nonwhite. 232

5. Frequency Distributions of Nonmetropolitan State Economic Areas for Each Region by the Proportion of Employed Persons Working in Agriculture 233

6. Frequency Distributions of Nonmetropolitan State Economic Areas for Each Region by the Proportion of the Employed Labor Force in Manufacturing 234

7. Frequency Distributions of Nonmetropolitan State Economic Areas for Each Region by the Proportion of the Employed Labor Force in Wholesale and Retail Trade 235

8. Frequency Distributions of Nonmetropolitan State Economic Areas for Each Region by Per Capita Dollar Retail Sales. 236

Table Page

9. Frequency Distributions of Nonmetropolitan State Economic Areas for Each Region by Per Capita Dollar Receipts from Services 237

10. Frequency Distributions of Nonmetropolitan State Economic Areas for Each Region by the Rural Farm Level of Living Index 238

11. Frequency Distributions of Nonmetropolitan State Economic Areas for Each Region by the Proportion of the Rural Farm Labor Force who are Unpaid Family Workers. 239

12. Frequency Distributions of Nonmetropolitan State Economic Areas for Each Region by the Proportion of the Rural Farm Labor Force in Nonfarm Occupations 240

13. Percentage Distributions of Nonmetropolitan State Economic Areas among Regions for the Total Group of Nonmetropolitan State Economic Areas and Ten Selected Control Groups 241

CHAPTER I

THE PROBLEM AND ITS SETTING

One of the tools which has been widely used in social science research both as an heuristic device and as a simple classification device is the concept—"region." Nearly every researcher working with data that are national in scope presents somewhere in his research report a regional breakdown of his results. This apparent compulsion to present data in a form that permits regional comparison presumably rests on the assumption that regions are important, distinctive subdivisions of the larger society. And, indeed, this assumption is supported at least provisionally by considerable evidence showing substantial differences among regions with respect to many characteristics.

Unfortunately, much of this evidence is presented in the form of averages or some other summary measures which necessarily conceal information about the distribution of characteristics <u>within</u> regional areas. This neglect of intraregional differences is particularly unfortunate because the significance of the regional concept rests fundamentally on the idea that each region is a homogeneous territorial unit—homogeneous, that is, with respect to one or more characteristics which serve to set apart the region from its surrounding area. In short, the taxonomic cornerstone of any regional delineation is homogeneity

within and heterogeneity among territorial units. It should be noted, however, that these necessary conditions for adequate regional delineation are independent of one another. That is to say, one may not safely infer intraregional homogeneity from evidence demonstrating interregional heterogeneity.

The fact that there has often been a lack of attention paid to intraregional similarities and differences should not be construed to mean that there has been no attention paid to this facet of regionalism. Indeed, a number of scholars have devoted a great deal of effort to the study of regionalism in all its aspects. These efforts have produced an impressive body of research data describing the regional character of the United States. Furthermore, some of these scholars have themselves set up regional delineations of the country[1]—a task requiring explicit, operational treatment of intraregional homogeneity.

In addition to the wealth of descriptive material provided by these studies another important aspect of the regionalists' work involves theoretical matters. Unfortunately, the writer has been unable to find in the literature a concise and unambiguous statement of the regionalists' theoretical position. Indeed, there seems to be a variety of opinions about what the concept "region" means and what significance it has for social

---

[1]See: Howard W. Odum, Southern Regions of the United States (Chapel Hill: The University of North Carolina Press, 1936); A. R. Mangus, Rural Regions of the United States (Washington: U.S. Government Printing Office, 1940); Donald J. Bogue, "An Outline of a Complete System of Economic Areas," American Journal of Sociology, LX, No. 2 (September, 1954), 136-139.

research. This is clearly illustrated in the work of the National Resources Committee during the 1930's when it conducted an investigation of regional factors and their importance for national planning. In the published report of this study an entire chapter was devoted to a recital of various views about regionalism expressed by prominent regional scientists.[2] Despite a lack of consensus among these scholars there are some areas of more or less general agreement which have theoretical implications worth noting.

First, regions are regarded as areas exhibiting homogeneity in one or more aspects. This regional homogeneity is not an abstract, or artificial, construct imposed arbitrarily by man but is rather a "spontaneous expression of physical and psychological differences. Regions are genuine entities, each of which expresses both natural and cultural differentiation from its neighbors."[3]

With this view of the region as an entity sui generis it would appear that we have here something resembling a comprehensive natural area of no small significance for the study of society. And, indeed, this is affirmed in a statement by the late regional scholar Howard W. Odum, who asserted that "the smallest and basic unit for the complete scientific study of society is found in the region."[4] This idea of the comprehensive

---

[2] U.S., National Resources Committee, Regional Factors in National Planning and Development (Washington: U.S. Government Printing Office, 1935), chap. xii.

[3] Ibid., p. 141.

[4] H. W. Odum and K. C. Jocker (eds.), In Search of the Regional Balance (Chapel Hill: University of North Carolina Press, 1945), p. 3.

significance of the concept-region-is expressed even more succinctly in a statement by the Natural Resources Committee, to wit., "In one sense, and perhaps the best one, regionalism is a way of life."[5]

Fortunately, regional scholars have done more than simply assert the equation of regionalism and "way of life." They have also presented a rationale which not only accounts for the reality of regions but also justifies the importance they attach to regionalism as a conceptual tool for the study of society.

This rationale is given rather clear expression in the report of the National Resources Committee. In a section devoted to a discussion of the "composite" or "multi-factor" region the Committee notes: "It is a known fact of human ecology or geography that there is an active interrelation between the patterns of human culture and activity on the one hand and the natural environment on the other. In no sense is this meant to imply that one is causal to the other; rather it is a way of saying that the structures and occupance patterns of man tend to vary with the character of the natural environment because the former are molded and adjusted to, and in large measure derived from, materials of the latter. The ecological relationships are so fundamental that they may be said to be the real determinants of regions. Moreover, they surround man with different conditions for living, different degrees of success in winning a livelihood, and different factors which challenge

---

[5]National Resources Committee, op. cit., p. 138.

attention, engross interest, and set social psychology and public opinion. Thus, regional differences may be said to be also psychological realities."[6]

In a brief but more recent statement of this point of view R. H. Knapp notes: "The essence of the regional idea in geography, then, is that land is varied, and that it can best be studied in terms of its variety. Also, the activities of man are closely related to environment, and this means that human activities ought to be studied in terms of regions."[7] Knapp also comments on the cultural and psychological aspects of regionalism by saying: "The regional approach emphasizes cultural phenomena which cannot be measured exactly, such as folkways, the spirit of peoples, the folklore, and personality."[8]

This emphasis on regionally distinctive cultures and psychological properties appears time and time again in the literature. R. E. Dickinson mentions that regions are, among other things, areas of "social feeling,"[9] or "areas of marked regional consciousness among inhabitants."[10] Odum lays great stress on the distinctiveness of regional folk-cultures and asserts: "The supreme climax of nature is in the personality of the individual and in the folk-culture of the people."[11]

---

[6] Ibid., p. 154.

[7] R. H. Knapp, American Regionalism and Social Education (Cambridge: Harvard University Press, 1947), p. 20.

[8] Ibid., p. 5.

[9] R. E. Dickinson, City, Region, and Regionalism (London: K. Paul, Trench, Trubner and Co., 1947), p. 8.

[10] Ibid., p. 254.

[11] H. W. Odum and K. C. Jocker (eds.), op. cit., p. 22.

He goes on to say: "Regional attitudes and mores are so definite and powerful that they constitute rights and wrongs; they determine the nature of behavior and institutions."[12] In a recent book on regionalism by Carle C. Zimmerman and Richard E. DuWors the authors discuss at great length the psychosocial structure of the United States as it relates to regional subdivisions of the country. Indeed, at the conclusion of their Introduction to the book Zimmerman and DuWors state: "Every region has its character making ideals—its regional personality forming processes."[13]

Although it has been remarked earlier that it is difficult to find in the literature a clear and concise statement of the regionalists' theoretical position, the preceding discussion does contain material which at least adumbrates such a position. To make this more explicit the writer has tried to arrange the more salient ideas in the form of several propositions which, considered together, represent what might be called the regionalists' theoretical approach to the study of society.

1. Society is the product of an interactive process in which man adjusts to his environment.

2. The basis of this adjustment activity is to be found in the cultural resources available to man.

3. While these cultural resources are partly determined by historical accidents (such as migration of ethnic groups), they also affect and are affected by the physical environment.

---

[12] Ibid., p. 33.

[13] C. C. Zimmerman and R. E. DuWors, Graphic Regional Sociology (Cambridge: The Phillips Book Store, 1952), p. 13.

4. Through the interaction of man, his culture, and his environment, there is produced a distinctive area of human group living which may be identified as a regional society.

5. While some regions may resemble each other more than others, each is discernibly different because each represents a unique constellation of factors based on differences in physical environment or in cultural heritage or both.

6. The consequences of this regional distinctiveness cannot be dismissed lightly because they have ramifications extending throughout the entire fabric of society. The behavior, attitudes, personalities, and institutions of the inhabitants are, so to speak, all stamped with a distinctive regional brand.

Stated in this manner, the regionalists' theoretical position in effect assumes the form of an hypothesis—an hypothesis in which regionalism represents an independent factor with respect to a large body of social phenomena. It would appear, therefore, that the utility of the regionalists' theoretical approach to the study of society is amenable to empirical examination to the extent that comparative data for different regions are available.

As mentioned earlier, however, most of the comparative data presented for regional subdivisions of the United States take the form of summary measures of some kind or other. Interesting as this evidence is, it does not represent a very adequate test of the theoretical validity or research utility of regionalism as a particular approach to the study of social life. The unitary quality of regions as distinct entities sui generis must be demonstrated by analysis not only of diversity among regions but also of similarity within such territorial units. It is quite possible to arrange by random procedure various groupings of differentiated elements comprising a universe

and still find what may seem to be substantial differences among the groupings with respect to summary measures. Only by employing techniques which can assess the influence of chance factors on the distribution of values can one have confidence that revealed differences among groups are not chance occurrences. To this writer's knowledge there have been published no empirical studies addressing this basic problem of regionalism. The research reported in this monograph represents one effort, albeit a limited one, to meet this problem vis-à-vis.

Although the problem of measuring the significance of regions as territorial units for social analysis is central to this study, there is also a collateral problem of considerable interest to the writer and, presumably, to anyone interested in regionalism. This is the problem of specifying what features of regionalism are most important as influences on behavior. In a sense the very fact that this collateral problem is raised assumes that regional distinctiveness does in some measure genuinely influence behavior. The question of why or how such influence is exercised requires for its answer some theoretical rationale buttressed by empirical evidence. Accordingly, it would seem appropriate here to comment on the kind of theoretical explanation the writer has outlined as representative of the thinking one finds in much of the literature on regionalism.

If regional diversity does in fact arise from varying conditions of the natural environment and from differences in the cultural heritage of the inhabitants, one would expect to find regional distinctiveness maintained and perhaps even increased wherever regions are relatively isolated or self-contained

areas. The self-sufficiency dictated by conditions of isolation requires that workable solutions to the problems of life be fashioned out of limited resources. The habitat with its local peculiarities and the inhabitants' prevailing culture with its idiosyncrasies-these represent the raw materials which man, as the active agent in an essentially passive environment, must use to meet the problem of survival. Although these natural and cultural resources do impose limitations, there is probably a variety of solutions possible within the range of these limitations. That is to say, the natural and cultural resources represent not only limiting but also permissive conditions. The important thing to be noted, however, is that once a workable solution has been contrived, it probably tends to be perpetuated, because first, it is workable, and second, isolation reduces the likelihood that functioning alternatives will be perceived. It would seem, therefore, that under conditions of isolation a regional population is likely to develop and maintain a way of life which is culturally distinctive.

A different situation is presented, however, when isolation is broken down and when technology becomes more elaborate. Under these conditions one might expect to find some leveling of regional differences. Wherever interregional social contact is sustained and is extensive, cultural diffusion is inevitable. Since cultures are patterns of interrelated traits, diffusion is a process that tends to be self-reinforcing. That is to say, diffusion begets diffusion. It is apparent that in this kind of situation cultural diversity is reduced in proportion to the amount of social contact and accompanying diffusion that occur.

This leveling of regional differences is also influenced by the prevailing technological development. The term-technology-is used here to refer to the tools and techniques by which man manipulates his natural environment. It can be seen, therefore, that the more highly developed the technology, the greater is man's control over his habitat and its energy resources. This reduces the limiting effects of environment on cultural development. Consequently, a well-developed technology will tend to reduce regional diversity to the extent that such diversity is a product of peculiarities of the local habitat.

The pertinence of these comments for the theoretical position taken by investigators of regionalism is obvious. In the United States today interregional transportation and communication are extensive, and the technology is well developed. Furthermore, the future seems to promise an expansion and intensification of interregional social contact and exchange. Accordingly, it is suggested that the picture of regional diversity as an independent factor of great importance in the study of social life may have been limned too boldly by regional scholars.

Unfortunately, the research design of this study of regionalism is not set up to test the hypothesis that regionalism is declining in importance for social analysis. An answer to this question must await further study. However, regardless of whether such a decline is or is not occurring, the fact remains that current data do show some regional diversity with

respect to behavior.[14] Thus, there still remains the question which is central to the collateral problem mentioned earlier--namely, what specific features of regionalism are influential with respect to behavior?

The literature on regionalism does not ignore this problem, but it does tend to treat it obliquely. Regional scholars are prone to identify regional diversity as cultural diversity and then to "explain" regional variations in behavior as being products of cultural differences among regions. Furthermore, when these cultural differences are discussed, they are often revealed to be motivational differences stemming from such things as personality types, philosophies of life, the spirit of the people, attitudes, mores, etc. When expressed by writers endowed with literary felicity, this kind of explanation can be rather persuasive. However, for the social scientist who seeks empirical verification, obvious difficulties arise because respectable, comprehensive data are not available for this aspect of social life. The importance of this problem was noted by Louis Wirth, who was prompted to remark: "The failure to discriminate the many distinct factors that underlie the emergence and persistence of regions is a serious fault of present day scholarship and research. . . . As a tool for the discernment of interrelations between habitat and culture, the regional concept has great value, provided we do not assume what needs to be proved--namely, that these correlations actually exist, and proceed

---

[14] It should perhaps also be noted that this study does not attempt to investigate the historical question of why or how regions have become the way they are.

to analyze the processes that account for these correlations."[15]

Since the problem of specifying significant regional differences cannot be handled very satisfactorily on a motivational level, some alternative must be devised. The alternative chosen for this study is both simple and sociologically pertinent. It consists of selecting several structural features of organized society which presumably vary as among regions and focusing attention on how such variations are associated with behavioral differences among regions. The thought here, of course, is that these structural features represent objective conditions to which people must adjust. Differences in these objective conditions may be regarded as differences in problems confronting people, and in so far as these problems require behavioral solutions, one would expect to find differences in behavior exhibited. In short, the assumption underlying the analysis here is that particular motivations and their behavioral expressions are derivative products of the basic structural characteristics of social organization. Therefore, the more similar the basic structures of different societies, the more similar will be the behavior patterns found. With reference to regionalism per se these remarks imply that much, if not most, of the behavioral differences among regions may represent adjustments of the regional populations to different structural conditions for collective life.

Summary

The importance of regionalism for the study of social

---

[15]Louis Wirth, "Limitations of Regionalism," Regionalism in America, ed. M. Jensen (Madison: University of Wisconsin Press, 1951), p. 386.

life has been clearly acknowledged by the frequent appearance in research literature of regional comparison used as a simple descriptive device, and by the intensive investigations of regionalism pursued by some scholars who regard it as a socially significant phenomenon per se to be explained. There is abundant evidence in the form of summary measures which attests to the significance of regions as distinctive subdivisions of the larger society. This demonstration of heterogeneity among regions represents only partial verification of the significance of regionalism. Of equal importance is the need to demonstrate homogeneity within regional areas. This latter problem has been somewhat neglected in empirical research. There has also been some neglect in specifying what features of regionalism appear to be most significantly related to behavioral differences among the inhabitants of the various regions.

The research undertaken in this study is focused on these two neglected problems of regionalism. The effort here is to measure with some precision first, the extent to which regionally located behavior is distinctive behavior and second, the extent to which behavioral diversity among regions is associated with specific features of regional social organization.

The specific features of organization involved in the second problem represent conditions presumably requiring behavioral adaptation. It is believed that the more similar these conditions, the more similar will be the adaptive behavior patterns. The implication here is that much if not most of the behavioral diversity among regions can be attributed to differences in the structural conditions of social life among the various regions.

CHAPTER II

THE PROCEDURE

The task of setting up a research design addressed to the problems of regionalism described in the foregoing chapter involved a number of questions which had to be answered more or less simultaneously. That is to say, solutions to the several problems of research design were somewhat contingent upon one another. In the interests of clear exposition, however, no attempt will be made here to describe in detail the procedural steps actually taken in setting up the design because they involved a great deal of "backing and filling." Nevertheless, the broad problems of design were handled rather systematically. The balance of this chapter is devoted to a discussion of these problems arranged roughly in the order in which they were addressed.

Scope and Focus of the Study

The decision to focus this study on regionalism in the United States was made for various reasons. First, with a land area of nearly three million square miles the United States exhibits considerable variation in physiography and thereby provides one important variable of potential significance for the emergence of regional subdivisions. Second, a good deal of the previous work on regionalism has been focused on the United

States, so that this study serves to extend an already well established line of investigation. Third, a variety of fairly reliable data is readily available from registration sources. By restricting the study to one national state, problems involving the accessibility and comparability of data, which so often beset studies of international scope, were greatly reduced.

Selection of Basic Units and the Problem of Control

Since regionalism from a schematic point of view represents spatial clusterings of the elements of a population, it was necessary in setting up the research design to specify what basic units were to be regarded as the constituent elements of the population being studied. This problem of selecting the basic units of measurement was especially important because a central feature of the proposed analysis was the measurement of homogeneity within and heterogeneity among the several regions. Although this sort of analysis could be pursued by treating the individual person as the basic population element, the sociological focus of the study prompted a somewhat more group-oriented conceptualization of the basic elements. That is to say, the regions were conceived as being populations of subgroups each of which was a more or less organized collectivity of individuals. Viewed in this manner, individuals become simply items of enumeration useful for describing properties of groups.

As a general guide in making a decision about the basic units, three criteria were kept in mind. First, the units must be of the type for which sufficient pertinent data are easily available. Second, the units must be few enough in number to

permit the analysis to be carried out by one person in a reasonable length of time. Third, the number of units must be sufficiently large to permit the exercise of controls in part of the analysis. This third criterion perhaps needs some elaboration.

It will be recalled that one of the two basic problems addressed in this study is that of investigating what structural features of regionalism appear to be closely related to behavioral differences among the regions. The technique to be used in this investigation was planned as follows: (a) Taking into account all units in all regions, measure the amount of variation in behavior exhibited by these units which seems to be associated with differences in the regional location of that behavior. (b) Select from each region basic units which are similar to one another with respect to certain structural characteristics and repeat the measure of behavioral variation associated with regional location. This represents a control for structural peculiarities among the regions. (c) Any difference between the magnitudes of the uncontrolled and controlled measures may be regarded as an index of the influence on behavior exercised by regional differences in structure.

It is certainly apparent that this method of control through selection involves risks of a high case mortality rate, especially if multiple controls are applied with any precision. Consequently, the greater the number of basic units eligible for this control selection, the more adequate would be the study's analysis of regional social structure and its influence on behavior. However, as indicated above, the study design must be tailored to fit the capacities of a one-person-research-operation.

Since the situation posed conditions clearly requiring some compromise, a search was made of data sources in order to find what alternatives were available for this decision about basic units.

The results of this search revealed that pertinent data were available for four different kinds of subregional units. These were counties, State economic areas, States, and urban places. Assessing the adequacy of these alternatives in terms of the criteria set forth above, it was quickly seen that the State economic area provided the best solution to the problem of selecting basic units. This decision was really made out of default rather than because of any intrinsic value attaching to the State economic area for the purposes of this study. Counties, of which there are over 3,000, were too numerous to be handled by one researcher in the time available. States were too few in number to permit the application of controls. And, like the county units, urban places, numbering over 3,000, also represented too great a burden for one person to work with. In addition, a decision to use urban places as basic units would have precluded any analysis of behavioral differences among the rural populations of the several regions. Accordingly, the State economic area was chosen to serve as the basic, subregional unit for this study. A few comments about this kind of unit are perhaps in order here.[1]

State economic areas are spatial units intermediate in size between counties and States. They are relatively homogeneous subdivisions of States because each State economic area

---

[1] For a more extended discussion, see Donald J. Bogue, State Economic Areas (Washington: U.S. Government Printing Office, 1951).

consists of one county or a group of counties which are similar with respect to certain social and economic characteristics. This grouping of county units serves to subdivide the forty-eight States into 501 State economic areas. For purposes of tabulating certain kinds of data, some of these areas are combined. Consequently, in this study the number of State economic areas used is 443.

One particularly useful feature of these intermediate areal units is that they are classified with respect to whether or not they are metropolitan in character. That is to say, with a few modifications the standard metropolitan areas established by the Division of Statistical Standards have been adopted as metropolitan State economic areas. This permits one to use State economic areas to analyze, among other things, the large and growing aggregates of metropolitan population in the country and yet permits the results of such analysis to be compared with findings from studies exclusively focused on standard metropolitan areas themselves. As will be seen, this study makes extensive use of this metropolitan-nonmetropolitan classification of State economic areas.

Although, at the time it was made, the decision to use the State economic area as a basic unit seemed best, in retrospect the arguments in favor of that decision do not appear quite so compelling. The discussion of the reasons for that decision set forth above perhaps gives the impression that the researcher was confronted with something resembling a Hobson's choice. That is to say, if counties and urban places were too numerous to be handled adequately and if States were too few in number to permit

the exercise of adequate controls, the only alternative, presumably, was the intermediate size unit--the State economic area. In point of fact, this is not precisely true because there was another alternative--an alternative which was considered and rejected perhaps too hastily. The reference here is to the possibility of using a sample of county units.

This idea of selecting the county as the basic unit and using sampling to keep the number of units within manageable limits was rejected because it seemed to involve some complications which would be difficult and time-consuming to resolve. For example, the sample would have had to be drawn in such a way as to assure adequate representation of rural and urban populations in each region and to keep the selection of sample counties within reasonable limits in terms of population size. Although these problems could probably have been met satisfactorily, the work-cost involved seemed inordinately high, especially when the availability of what appeared to be an adequate number of State economic areas obviated such problems of sample selection and subsequent problems of adjustment in the statistical measures used. Accordingly, the State economic area was selected as the basic unit of measurement for the study.

Selection of a Regional Delineation

One of the most striking features of the literature on regionalism is the uncertainty, or at least lack of consensus, about the number and distribution of regional areas in the United States. There are literally scores of different regional patterns which have been set up by various persons and agencies. Indeed,

the report of the National Resources Committee on regionalism includes no fewer than 145 different regional delineations of the country.[2] To be sure, most of these represent national subdivisions more or less arbitrarily set up for the administrative convenience of various federal agencies. However, there still is quite a variety of delineations which purport to represent genuinely emergent natural areas. Some of these are based on a single-factor criterion such as climate, topography, or types of farming, while others have been constructed from multiple-factor criteria which combine such things as physical, demographic, economic, and social characteristics into regionally discriminating indexes.

The task of selecting from this array of possibilities one regional pattern which is pertinent and researchable from the standpoint of the problems addressed in this study is not an easy one. The criteria finally used as a rough guide for this selection involved the following considerations: (a) The regional boundaries must not cut across State economic areas because these are the basic measuring units. (b) The several regions must be few enough in number so that each constitutes a sufficiently large grouping of basic units to permit control selection within each region. (c) The several regions must be large enough in number so that none encompasses widely divergent areas and populations. (d) The regional delineation should be fairly well-known, preferably one which has attracted interest and has been actually

---

[2] U.S., National Resources Committee, Regional Factors in National Planning and Development (Washington: U.S. Government Printing Office, 1935).

used in empirical research. (e) The plan followed in delineating the regional pattern should be one involving some methodological sophistication; that is, a plan in which intuitive judgments are subordinate to respectable, quantitative procedures.

With these criteria in mind several alternative regional patterns were considered, and it was concluded that the classification proposed by Howard W. Odum provided an acceptable solution to this problem of selecting a regional delineation for analysis. Professor Odum, together with a large staff of associates, developed this sixfold regional division of the United States after an exhaustive study involving some 700 social, economic, and geographic characteristics. This represents by far the most ambitious effort to construct a comprehensive regional pattern which the writer has found reported in the literature.

Each of the six regions is composed of four or more entire States, so that there is no fracturing of State economic areas. The smallest number of basic units comprising a regional grouping is 26, while the largest is 129. As the product of a prominent exponent of regional study, Professor Odum's work merits serious attention, and this it has received from chorographers and other interested persons. Indeed, the Office of Business Economics in the United States Department of Commerce adopted (with one small modification) the Odum sixfold regional pattern for its impressive analysis of Regional Trends in the United States Economy.[3]

Testimony as to the comprehensiveness of this regional

---

[3]Charles A. R. Wardwell, Regional Trends in the United States Economy (Washington: U.S. Government Printing Office, 1951).

pattern is afforded both directly and indirectly by Professor Odum himself. In the monograph reporting the results of this collaborative effort to study regionalism in the United States, Professor Odum sketches the requirements of a satisfactory regional subdivision as follow: "The number of regions must not be too large, nor, on the other hand, must the regions themselves be too large for effective study and unity. Each region should combine the largest number of geographic, economic, and cultural factors possible for the purpose of classification and study."[4] Although explicitly noting the artificiality of observing State boundary lines, Professor Odum with perhaps pardonable pride summarizes the major contributions of the study as including this "definitive sixfold regional division of the United States"[5] which he hopes "may contribute something toward a more uniform basis for regional study and planning."[6] In short, the Odum classification seems to fulfill at least passably well the criteria set forth earlier as a guide in selecting the regional delineation most appropriate for this study. The six regions with their respective State complements are listed below:

Northeast—Connecticut, Delaware, District of Columbia, Maine, Maryland, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, Vermont, West Virginia.

Southeast—Alabama, Arkansas, Florida, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, Virginia.

---

[4] Howard W. Odum, Southern Regions of the United States (Chapel Hill: University of North Carolina Press, 1936), p. 263.

[5] Ibid., p. 5. Italics added by the writer.

[6] Ibid., p. 7.

Middle States—Illinois, Indiana, Iowa, Michigan, Minnesota, Missouri, Ohio, Wisconsin.

Northwest— Colorado, Idaho, Kansas, Montana, Nebraska, North Dakota, South Dakota, Utah, Wyoming.

Southwest— Arizona, New Mexico, Oklahoma, Texas.

Far West— California, Nevada, Oregon, Washington.

Selection of Dependent Variables

It was observed in the preceding chapter that regional scholars frequently write about the essence of regionalism with a literary flair which may be attractive for the reader but which presents operational difficulties for the researcher. With no reliable and comprehensive data on philosophies of life, personality traits, and various spiritual qualities of the human soul, the problem of measuring directly the regional distinctiveness of such phenomena must be abandoned. This does not mean, however, that research interest in this aspect of life must be entirely forsaken. If one proceeds from a somewhat modified pragmatic position that such phenomena have importance in direct proportion to their behavioral consequences, the problem becomes at least indirectly researchable. That is to say, an attempt to operationalize these psychic aspects of life can be made by searching for behavioral expressions of them.

This search was undertaken, and it met with limited success because the study design requires data national in scope. An investigation of Census and other resources finally yielded five variables which seemed promising. They cover a fairly wide variety of human activity. Each represents a type of behavior which has attracted the interest of researchers and laymen alike.

And the reliability of the data (with one possible exception) is reasonably good. These variables, which are viewed as "dependent" in the context of this study, are as follow:

1. Native white net reproduction rate, 1935-40.
2. Proportion of eligible voters[7] who voted in the Presidential Election of 1940.
3. Proportion of voters who voted for the Democratic Party's candidate in 1940.
4. Proportion of males age 25 years and older in 1940 who reported no formal schooling.
5. Proportion of the total population officially recorded as being affiliated with a formal religious organization, 1936.[8]

Each of these dependent variables may be regarded as the product of behavior motivated in some measure by the ideas, interests, desires, and attitudes which people acquire through collective living. Decisions must be made with respect to having a family, voting, getting an education and joining a church. If personality factors, philosophies of life, and other similar phenomena constitute the mainsprings motivating these kinds of decisions, and if these motivational mainsprings differ among regions, then the consequent behavior should also exhibit regional differences. This at least is the logic of the argument for presenting this study as a partial test of the regionalists' basic hypothesis--that regionalism is an important, independent

---

[7] "Eligible voters" are here defined as all persons of voting age; viz., 21 years old for every State except Georgia, where 18 is the lower age limit for voting. It is recognized that this overestimates the size of the eligible voter population. However, it is the best convenient estimate available.

[8] The figures for the 1936 total population were estimated by interpolating from 1930 and 1940 Census data.

factor for a wide variety of social phenomena.

It will be noted that data for the dependent variables listed above all relate to 1940 with the exception of religious affiliation, which relates to 1936.[9] As will perhaps be surmised, the decision to focus this study on data gathered more than a decade ago was prompted primarily by expediency. Insufficient data from the 1950 Census were available in the form of State economic area tabulations to permit focusing the analysis on more current evidence. In addition, however, there are two

---

[9]The sources of data for the dependent variables are as follow:

a) Net reproduction rate: Donald J. Bogue, State Economic Areas (Washington: U.S. Government Printing Office, 1951), Table A.

b) Proportion of eligible voters voting, 1940: For data on number of eligible voters (i.e., persons age 21 plus, 1940), see: U.S., Bureau of the Census, U.S. Census of Population, 1950, Vol. II, Characteristics of the Population (Washington: U.S. Government Printing Office, 1952), Table 41. For data on number of voters, see: U.S., Bureau of the Census, Vote Cast in Presidential and Congressional Elections, 1928-1944 (Washington: U.S. Government Printing Office, 1946), Table 7.

c) Proportion of voters voting Democratic: U.S., Bureau of the Census, Vote Cast in Presidential and Congressional Elections, 1928-1944 (Washington: U.S. Government Printing Office, 1946), Table 7.

d) Proportion of males age 25 years plus with no formal schooling: U.S., Bureau of the Census, Sixteenth Census of the United States. Population, Vol. II, Characteristics of the Population (Washington: U.S. Government Printing Office, 1943), Table 21.

e) Proportion of population affiliated with religious organization, 1936: For data used to interpolate 1936 total population, see: Donald J. Bogue, State Economic Areas (Washington: U.S. Government Printing Office, 1951), Table A. For data on number of persons with religious affiliation, see: U.S., Bureau of the Census, Religious Bodies, 1936 (Washington: U.S. Government Printing Office, 1941), Vol. I, Table 32.

positive advantages derived from using the older material. First, 1940 was a Presidential Election year, so that the analysis of relationships between voting behavior and characteristics of social structure are temporally comparable. Second, there are no conveniently assembled data on religious affiliation subsequent to 1936. Consequently, focusing the study on 1950 material would certainly have meant discarding this dependent variable. Indeed, there may still be some question raised about the legitimacy of including this variable even in a 1940 analysis in view of the four year time lapse between the recording of religious behavior and the recording of structural characteristics to which that behavior is presumably related. It was felt, however, that the ambiguity introduced by this temporal disjuncture is more than compensated for by the extension which a religious behavior item provides to the range of dependent variables studied.

There is actually a more damaging criticism which can be made of this religious behavior variable. Stated briefly, it questions the reliability of the data. The 1936 Census of Religious Bodies collected its data from mail questionnaires which were sent to the officials of every established church congregation in the country. Although the Bureau of the Census exerted considerable effort to get complete returns, there is reason to believe that the 1936 survey suffered from poor reporting. For example, the total church membership increased only 2.3% between 1926 and 1936, although the increase in total population was considerably larger during this period.[10] Furthermore, there was a

---

[10] U.S., Bureau of the Census, Religious Bodies, 1936 (Washington: U.S. Government Printing Office, 1941), I, 85.

net decrease of nearly one-half billion dollars in the reported value of church edifices during this decade. While some of this decrease may have been genuine in the sense that it represents the abandonment of church property due to consolidations of congregations or the reassessment of church property during the early years of the depression decade, the Bureau of Census people believe that the greater part of the decrease can be attributed to incomplete reporting.[11]

The decision to retain in the study this religious behavior item despite apparent inadequacies of the data may have been unwise. However, since there was no good evidence of an important regional selectivity in poor reporting and since inclusion of data on religious affiliation does add an interesting item to the inventory of dependent behavior variables, it was decided to retain it. The reader should keep in mind, however, that religious affiliation is the least reliable of the five dependent variables analyzed.

Although the appropriateness of the fertility and voting variables is perhaps self-evident, the datum selected to measure educational achievement deserves some clarifying comment. The proportion of adult males with no formal schooling is perhaps the best estimate of illiteracy that can be devised from 1940 Census data. While this may seem to represent a somewhat negative approach to educational achievement, it is actually this quality of "negativeness" which makes it appropriate for this study. In other words, it is regarded as a measure of the extent to which a

---

[11] Ibid., p. 17.

population has failed to take advantage of educational opportunity. This, of course, rests on the assumption that everyone, regardless of regional location, has had the opportunity to get at least some formal schooling.

It was felt here that the thing to be avoided was a measure of educational achievement which clearly reflected regional differences in wealth or poverty, because one of the intended controls to be exercised was that of standard of living. Since educational opportunities involve costs which are more burdensome the poorer the region, the decision was made to minimize the influence of this cost burden by reducing the educational requisite to its barest essentials—namely, any formal schooling at all. In other words, the researcher did not want to "stack the deck" in his favor by selecting a measure of educational achievement which clearly discriminated to the advantage of the economically well-favored region.

These considerations, of course, represent only one side of the coin. It was also necessary to make sure the measure finally selected was not so minimal that it failed to provide at least some discrimination among regions. A spot check revealed that differentiation on the basis of some-or-no-formal-schooling does provide small but noticeable differences among regional populations with respect to educational achievement. Accordingly, the proportion of adult males with no formal schooling was adopted as the dependent education variable.

Selection of Independent Variables for Control

It has already been remarked that a concept frequently

appearing in the literature on regionalism is that of culture. Sometimes the concept is treated as a dependent factor which must be explained. At other times it assumes independent character and is used to explain other phenomena. It seems to the writer that resort to the culture concept as an independent factor leaves much to be desired. The concept is usually not sufficiently discriminating to be very useful as an explanatory device because it appears to include just about everything except the physiological organism and the physical environment. Furthermore, as has already been noted, even when culture is discussed in the literature, little of value for empirical research emerges because the discussion so often focuses on motivational phenomena for which good, comprehensive data are conspicuously lacking. Basically, the problem here is that of specifying what features of regionalism have importance as influences on behavior.

As befits a study with a sociological focus, the particular features of regionalism selected for analysis here are those pertaining to social structure. Accordingly, a selection of thirteen specific, independent factors has been made in an effort to identify more precisely some basic determinants of behavioral differences among regions. This list of variables, although limited, of course, by the availability of data, does provide an interesting variety of social and economic characteristics. All of these data are for 1940. They are as follow:[12]

---

[12] The source of data for all independent variables is as follows: Donald J. Bogue, State Economic Areas (Washington: U.S. Government Printing Office, 1951), Table A.

1. Metropolitan status.
2. Proportion of total population which is urban.
3. Proportion of total population which is rural nonfarm.
4. Proportion of total population which is rural farm.
5. Proportion of employed workers in manufacturing.
6. Proportion of employed workers in agriculture.
7. Proportion of employed workers in wholesale and retail trade.
8. Rural farm level of living index.
9. Per capita dollar retail sales.
10. Per capita dollar receipts from services.
11. Proportion of rural farm labor force employed in nonfarm occupations.
12. Proportion of rural farm labor force employed as unpaid family workers.
13. Proportion of total population which is nonwhite.

Each of these characteristics, it may be observed, is an index representing more or less specific conditions for collective living. In so far as behavior patterns represent solutions or adjustments to problems posed by these conditions of life, one would expect to find differences in behavior associated with differences in the index values of the characteristics listed above. Looking at it from a broader perspective, if regionally diverse behavior can be attributed to regionally diverse social structure, then a manipulation of these social and economic characteristics should be accompanied by changes in the behavior patterns observed. It is, of course, recognized that demonstration of an association is not tantamount to demonstrating a cause-and-effect relationship. However, it is also recognized that the former is an inevitable accompaniment of the latter.

Collection of Data and Statistical Measurement

With the State economic area designated as the basic unit of measurement in this study design, the task of gathering data involved primarily combining tabulations for the 3,101 counties

or equivalent subdivisions of the United States into multiple-county groupings numbering 443. Fortunately, this rather arduous job had to be done for only four of the dependent variables. Data for all other variables, both dependent and independent, are available in the form of tabulations for State economic areas.

Since it had been decided that manipulation and measurement of the data would be facilitated by the use of IBM punch cards, all of the data were prepared for coding sheets. The exact value of each dependent variable was entered on the coding sheets for each State economic area. In addition, blank columns were left in order that the value of each dependent variable could be squared and entered in the punch cards by a tabulating punch machine. Data on the independent variables were classified by coded intervals on frequency distribution scales, none of which exceeded a two-column code and most of which were contained in a one-column code.

A few comments about the reliability of this work are perhaps in order here. The problem of checking reliability was depressing to contemplate, because it involved for the most part simple duplication of the work. However, some effort was made to check each stage of the work. With respect to the accuracy of copying and combining county data for the four dependent variables not tabulated by State economic areas, a six per cent duplication check of the original work was made by a person other than the original recorder. This involved examining approximately 1200 of the over 18,000 county entries copied from Census sources. The error reported was negligible. A ten per cent

check was made by the original coder of coding reliability which showed less than a one per cent error; i.e., less than one per cent of the code items were in error. Card punching was done by the Tabulating Service of the University of Michigan. Part of the "service package" provided by this organization for this type of work is a 100 per cent duplication check. Since the coding sheets were clean and the punching involved no peculiarities, there is reason to believe that error in this operation was negligible. The only check made on the IBM machine tabulations was that of comparing the various runs for discrepancies in the total N's for each region. No errors were detected here. With respect to measurement procedures, several internal checks were possible and were used. However, in addition, each measure was computed twice, and a few which appeared intuitively to be peculiar were computed three or more times. In summary, therefore, it is the writer's opinion that the accuracy of the work reported here is fairly respectable.

It will be recalled from the discussion in the first chapter that an adequate measure of regional distinctiveness must comprehend both difference among regions and similarity within regions. Accordingly, the question of what statistical device would be appropriate for this study prompted a search for an available measure in the field of analysis of variance. After several alternatives were considered, it was finally decided that the intraclass correlation (rho) would be an interesting and effective measuring instrument to use.[13]

---

[13] For a discussion and illustrative application of this statistical measure, see: Leslie Kish, "Differentiation in Metropolitan Areas," American Sociological Review, XIX, No. 4 (1954), 388-398.

Stated generally, the intraclass correlation is a ratio which shows what proportion of the total variance among the elements of a population can be accounted for or "explained" by the clustering of those elements in various subgroups of the population. Translating this into terms pertinent for this study design, the intraclass correlation can tell us for each of our dependent variables how much of the total variance among State economic units is due to the location of these units in the Odum sixfold regional pattern.

The formula used for this intraclass correlation is as follows:

$$\text{rho} = \frac{\sigma_b^2}{\sigma^2} - \left[\frac{M}{N-M} \cdot \frac{\sigma_w^2}{\sigma^2} \right]$$

where: N = the number of State economic areas.
M = the number of regions.
σ^2 = the total variance among all State economic areas.
σ_b^2 = the variance among the individual means of regions.
σ_w^2 = the variance among State economic areas but within regional groupings.

It will be observed that the first term of the formula represents the ratio referred to in the definition above. The second term in brackets following the minus sign is simply a corrective factor designed to reduce rho to zero whenever the magnitude of the difference among regional means is within the range of chance fluctuation.[14]

Since the basic mechanism of this measure is one which involves breaking down the total variance into two components—the variance within and the variance among groups or regions—

---

[14] For a sophisticated statistical discussion of the intraclass correlation, see: Leslie Kish, "On the Differentiation of Ecological Units" (unpublished Ph.D. dissertation, Department of Sociology, University of Michigan, 1952).

it can be seen that changes in either or both of the components may affect the value of rho. The word "may" is used here because changes in both components can be complementary with the result that the effects of each change are cancelled out and the value of rho is left unaltered.[15] In short, rho is a summary measure which reflects net influences. To get a real picture of regional distinctiveness one must not stop with the summary measure, but rather investigate why it assumes the magnitude it does. This requires examination of its components. The pertinence of these remarks will become more evident when the substantive results of the study are presented.

One additional feature of rho which the reader should be aware of involves its relationship to the well-known F-test of significance in analysis of variance. This relationship can be shown rather clearly in the following formula:[16]

$$F = \left[\text{rho} + \frac{M}{N-M} \right] \frac{N-M}{M} \cdot \frac{N-1}{N}$$

By transposing terms to solve for rho one may derive a formula which permits the specification of minimum rho values necessary to achieve statistical significance at different probability levels; viz.,

$$\text{rho} = \frac{M}{N-M} \left[F \cdot \frac{N}{N-1} - 1 \right]$$

Inspection of this formula shows clearly that the larger the size of N the smaller the value of rho necessary to achieve any given

---

[15] Other features of rho of pertinence for this study are presented later in this report where the substantive findings are discussed. See pages 49, 50, 80, 105-107.

[16] See: Kish, loc. cit., p. 122.

level of significance.

Before the conclusion of this discussion of research procedure, one cautionary note should be recalled to the reader's attention. This involves the nature of the substantive findings to be presented in the following chapters of this report. Although the dependent variables in this study constitute measures of types of behavior which individuals engage in, the form in which these data appear in this research does not pertain to individuals per se but rather to groups of individuals—i.e., the inhabitants of State economic areas. Accordingly, analysis of the behavioral homogeneity or heterogeneity within and among regions is really an analysis of similarity or differences in the behavior of regional subgroups. In other words, in terms of analysis each region constitutes a population of subgroups, not individuals.

With this in mind one other feature of the substantive findings and their analysis should also be noted. This involves the question of whether the data of this study are to be regarded as universe or sample data. If the data are viewed as constituting the universe of phenomena being investigated, then the use of significance tests and corrective factors for chance fluctuation may be regarded as irrelevant. On the other hand, if the data are viewed as a sample of a universe of phenomena, then the specific values of the measures computed become less interesting, and the significant focus of attention becomes directed to tests of a null hypothesis or perhaps to the establishment of confidence limits for the particular values computed.

It is the writer's belief that arguments can be presented

in support of both points of view. For example, since the 443 State economic areas used in this study include every foot of land area and every inhabitant within the continental limits of the United States, it would seem that the data are comprehensive and therefore constitute a genuine universe. On the other hand, given these comprehensive data, one may still raise the question of whether the classification of these data into a particular, six-fold, regional pattern (e.g., Odum) produces regional behavior differences which are greater than one would expect from a purely random six-fold classification. The pertinence of this question is readily apparent when it is recalled that the entire work of the regional scholars is predicated on the assumption that regions are significantly different subdivisions of the larger society. Since the validity of this assumption can in part be tested statistically, it would seem appropriate within the limits of this research to apply such a test through the use of sampling statistics. In summary, therefore, the writer feels that the findings of this study can be profitably examined as representative of both universe and sample phenomena.

The substantive material of this report is organized and presented in the three chapters which immediately follow. Chapter III reports on the results obtained when the entire complement of State economic areas is analyzed in terms of its totality and in terms of its two major subdivisions—the metropolitan and nonmetropolitan areas. Chapter IV reports on the results obtained when controls are applied to the metropolitan segment of the population. Chapter V reports on the results obtained when the nonmetropolitan segment is examined under various controlled

conditions.

Summary

This is a study of the extent to which regionalism in the United States is associated with diverse patterns of behavior. The regional pattern examined is that devised by Professor Howard W. Odum, who, on the basis of exhaustive analysis, subdivided the entire country into six regional areas each of which consists of four or more States. The basic units of measurement for the study are 443 State economic areas. These units are relatively homogeneous subdivisions of States and have been constructed by combining counties which have similar social or economic characteristics. Five dependent variables representing a variety of types of behavior have been selected for analysis. The first task is to measure how much of the total variation in each type of behavior can be "explained" by the regional locations of the populations exhibiting this behavior. The second task is to measure how much of this "explained" variation can be attributed to differences in the social structures of the six regions. Accordingly, thirteen independent variables representing structurally pertinent characteristics have been selected for use as control variables. All data--for both the dependent and independent variables--relate to 1940 except for one dependent variable which is reported for 1936. The statistical measure used is the intra-class correlation (rho).

CHAPTER III

THE FINDINGS: THE UNIVERSE OF STATE ECONOMIC AREAS AND A COMPARISON OF ITS METROPOLITAN AND NONMETROPOLITAN SEGMENTS

The substantive findings reported in this chapter are arranged under five major headings as follow: (a) a comparison of regional means for each of the dependent variables measured,[1] (b) a comparison of intraclass correlations for each dependent variable, (c) an analysis of the total interregional and intraragional variance components of the intraclass correlations, (d) a comparison of regions with respect to their internal homogeneity for each dependent variable, and (e) a comparison of regional rank positions for metropolitan and nonmetropolitan areas. The data under the first four headings are presented for the total complement of State economic areas and also for the metropolitan and nonmetropolitan segments of this universe of State economic areas.

Comparison of Regional Means

Data for the fertility variable are presented in Table 1,

---

[1] The regional mean values for all the dependent variables throughout this study have been computed by summing the separate State economic area values and dividing by the number of State economic areas involved. In short, all State economic areas have been treated as equivalent units for statistical purposes, and no attempt has been made to weight these unit areas on the basis of their population size.

which contains mean values for the native white net reproduction rate, 1935-40. A comparison of the regional means for the total group of State economic areas found in the first column clearly shows considerable difference between the two extreme regions represented by the high fertility Southeast and the low fertility Far West. The range of 243 in this distribution of means appears to be substantial and is not the result of any one highly deviant region. These data suggest that the reservoirs for population growth in the United States are to be found in the two Southern regions and in the Northwest.

TABLE 1

DISTRIBUTIONS OF REGIONAL MEANS FOR FERTILITY* BY VARIOUS STATE ECONOMIC AREA GROUPINGS

Regions	State Economic Area Groupings					
	Total Econ. Areas	N	Metrop. Econ. Areas	N	Nonmetrop. Econ. Areas	N
Northeast	998	86	869	43	1127	43
Southeast	1195	129	852	35	1322	94
Middle States	1066	109	876	41	1181	68
Northwest	1148	45	817	6	1199	39
Southwest	1168	40	880	11	1278	29
Far West	952	34	814	13	1037	21

*As a measure of fertility the net reproduction rate indicates the number of female babies which would be born to 1,000 females born "today" if these 1,000 females were subject to specified birth and death rate patterns during their generation life span. In short, it is a measure of the intrinsic replacement potential of a population on a per generation basis.

An examination of the third and fifth data columns of Table 1 reveals some interesting and perhaps not surprising

differences in fertility rates between metropolitan and nonmetropolitan segments of each regional population. In every region metropolitan fertility falls considerably below nonmetropolitan fertility levels. Indeed, in no region do the metropolitan areas have fertility rates sufficient to replace their populations. This, of course, means that the population potential for replacement and growth must reside in the nonmetropolitan segments of the country, and this is in fact indicated by the mean values found in column five.

The most dramatic differences between metropolitan and nonmetropolitan fertility appear in the Southeast and Southwest regions, both of which exhibit metropolitan rates over 30 per cent lower than those for nonmetropolitan areas. Furthermore, not only is metropolitan fertility lower than nonmetropolitan fertility in every region, but the ranges and regional distributions of these rates are considerably altered in these two population segments. There is very little variation among metropolitan regional means. The total range of mean values is only 66, as compared with 243 reported above for the total group of State economic areas and 285 for the nonmetropolitan areas.

On the basis of these findings it would seem advisable to distinguish between metropolitan and nonmetropolitan segments of the population before generalizing about regional fertility patterns. Both the magnitude of fertility differences and the rank ordering of regions with respect to these differences are materially altered by whether or not the regional populations are organized in a metropolitan type of social structure.

In Table 2 regional means are presented for the proportion

of eligible voters who voted in the 1940 Presidential Election. These data show some striking differences and similarities among regions with respect to voter turnout. For the total grouping of State economic areas the Northwest and Middle States regions are very similar in having a high proportion of eligible voters voting, while the Southwest and particularly the Southeast are quite low in their voter turnout. The remaining two regions--the Northeast and Far West--occupy an intermediate position between the extremes. It is interesting to note that this pairing of regions with respect to the proportion of voters voting resembles what seems to be an incipient pairing pattern among our six regions for fertility. That is to say, the same regions tend to be paired on the basis of the similarity of their values, although their rank positions are quite different for the two variables.

TABLE 2

DISTRIBUTIONS OF REGIONAL MEANS FOR "PROPORTION OF ELIGIBLE VOTERS VOTING" BY VARIOUS STATE ECONOMIC AREA GROUPINGS

Regions	State Economic Area Groupings					
	Total Econ. Areas	N	Metrop. Econ. Areas	N	Nonmetrop. Econ. Areas	N
Northeast	68.1%	86	67.5%	43	68.6%	43
Southeast	27.8	129	26.6	35	28.2	94
Middle States	74.3	109	70.8	41	76.4	68
Northwest	75.7	45	72.8	6	76.2	39
Southwest	41.4	40	34.5	11	44.1	29
Far West	64.6	34	64.5	13	64.6	21

This pattern of regional means for voter turnout based on the total number of State economic areas remains virtually

unchanged for both the metropolitan and nonmetropolitan segments of the regional populations. In terms of actual values, however, the distributions of metropolitan and nonmetropolitan means are not identical. Although the differences are with one exception rather small, a comparison of regional means for metropolitan and nonmetropolitan areas does show a metropolitan voter turnout which is smaller in every region than the turnout in nonmetropolitan areas. This similarity among regions of metropolitan-nonmetropolitan voter differentials tends to parallel the patterns observed for fertility, and thus supports in some measure the suggestion noted earlier that the metropolitan type of social structure appears to have similar effects on the populations of every region. It should be pointed out, however, that the stability of this regional pattern in voter turnout and the rather small metropolitan-nonmetropolitan differentials exhibited seem to indicate a genuine regional distinctiveness in voting which is not greatly affected by the metropolitan type of social structure.

Distributions of regional means for the proportion of voters voting Democratic in the Presidential Election of 1940 are presented in Table 3. The most conspicuous feature of these data is the overwhelming predominance of Democratic voters in the two southern regions. There is very little difference among the Northeast, Middle States, and Northwest regions, which show relatively low proportions of Democratic voters. The Far West maintains an intermediate position between these extremes—a position similar to that for the voter turnout variable discussed above. Once again there appears to be a pairing of regions with respect to mean values, although this is somewhat modified by a splitting

of the Northeast-Far West pair with the former region exhibiting a mean value quite similar to those of the Middle States and Northwest regions.

TABLE 3

DISTRIBUTIONS OF REGIONAL MEANS FOR "PROPORTION OF VOTERS VOTING DEMOCRATIC" BY VARIOUS STATE ECONOMIC AREA GROUPINGS

Regions	State Economic Area Groupings					
	Total Econ. Areas	N	Metrop. Econ. Areas	N	Nonmetrop. Econ. Areas	N
Northeast	48.5%	86	52.8%	43	44.2%	43
Southeast	79.0	129	78.9	35	79.1	94
Middle States	49.1	109	54.7	41	45.7	68
Northwest	47.3	45	53.0	6	46.4	39
Southwest	70.7	40	72.4	11	70.1	29
Far West	56.9	34	58.4	13	55.9	21

This regional rank order pattern persists relatively unchanged for both the metropolitan and nonmetropolitan segments of the population. However, the magnitude of the differences in Democratic voting between metropolitan and nonmetropolitan areas is quite noticeable for all regions except the two southern ones. For every region except the Southeast the metropolitan vote is more heavily Democratic than the nonmetropolitan. Indeed, the data show differences sufficiently large to give the Democratic Party a majority of votes cast in the metropolitan areas of every region.

This propensity for metropolitan populations in all regions to vote Democratic suggests that metropolitan-nonmetropolitan

distinctions are of more than passing interest when regional patterns of voting preference are considered. Furthermore, it also suggests that the apparently small differences between metropolitan and nonmetropolitan voter turnout in the several regions may have important consequences for the final, total outcome of voter preference.

In Table 4 there are presented distributions of regional

TABLE 4

DISTRIBUTIONS OF REGIONAL MEANS FOR "PROPORTION OF MALES AGE 25 YEARS PLUS WITH NO FORMAL SCHOOLING" BY VARIOUS STATE ECONOMIC AREA GROUPINGS

Regions	State Economic Area Groupings					
	Total Econ. Areas	N	Metrop. Econ. Areas	N	Nonmetrop. Econ. Areas	N
Northeast	3.5%	86	3.8%	43	3.2%	43
Southeast	6.9	129	4.4	35	7.8	94
Middle States	1.8	109	2.0	41	1.8	68
Northwest	1.5	45	1.4	6	1.5	39
Southwest	6.1	40	4.4	11	6.7	29
Far West	2.6	34	3.0	13	2.4	21

means for the proportion of adult males with no formal schooling. It is evident at first glance that none of the regions has what in an absolute sense can be regarded as a really high proportion of unschooled adult males. However, the differences among regions are quite noticeable and relatively substantial. Once again, we find a pairing of regions on the basis of their mean values. The highest proportions are found in the two southern regions, the lowest in the Middle States and Northwest, while the

Northeast and Far West occupy intermediate positions. This rank order pattern persists without change for both the metropolitan and nonmetropolitan segments of the population.

The differences in mean values between the metropolitan and nonmetropolitan areas are rather interesting. The two southern regions have substantially lower proportions of unschooled males in their metropolitan areas than in their nonmetropolitan areas. The Northeast and Far West show precisely the opposite change with larger proportions in the metropolitan areas, while the Middle States and Northwest show very little difference. The effect of all this is to reduce the total range of regional differences for metropolitan areas to something less than one-half that for nonmetropolitan areas.

It seems to the writer that these regional differences in educational achievement between metropolitan and nonmetropolitan areas can most probably be attributed to migration. It does not appear reasonable to view the metropolitan areas of the Northeast and the Far West as having less educational opportunities than the nonmetropolitan areas of these regions. Furthermore, it is well known that these metropolitan centers, particularly in the Northeast, have served as ports of entry and residential destinations for foreign immigrants and internal migrants. In addition, it is also well known that the South has been a source for a great deal of internal migration, especially of the rural to urban variety. In short, it would seem reasonable to infer that a large proportion of the unschooled adult males residing in the metropolitan areas of the Northeast and Far West (and probably even the Middle States) are immigrants from foreign countries and

migrants from rural sources within the United States. In any event, the fact remains that regional diversity as measured by these summary figures of educational achievement is considerably altered by metropolitan-nonmetropolitan distinctions.

The fifth dependent variable examined in this study is that of religious affiliation. Table 5 presents regional means

TABLE 5

DISTRIBUTIONS OF REGIONAL MEANS FOR "PROPORTION OF POPULATION WITH RELIGIOUS AFFILIATION" BY VARIOUS STATE ECONOMIC AREA GROUPINGS

Regions	State Economic Area Groupings					
	Total Econ. Areas	N	Metrop. Econ. Areas	N	Nonmetrop. Econ. Areas	N
Northeast	47.9%	86	53.4%	43	42.3%	43
Southeast	37.9	129	41.4	35	36.6	94
Middle States	41.3	109	40.6	41	41.7	68
Northwest	40.3	45	42.3	6	40.0	39
Southwest	36.3	40	40.2	11	34.8	29
Far West	26.0	34	29.2	13	23.9	21

for the proportion of the total population affiliated with an established church. Perhaps the most impressive occurrence noted for the total group of State economic areas is the difference between the Northeast as the highest and the Far West as the lowest regions with respect to religious affiliation. The range of almost 22 percentage points is quite substantial. The other four regions once again tend to cluster in two pairs, with the two southern regions falling somewhat below the Middle States and Far West regions.

The extremes in this rank order pattern persist for both the metropolitan and nonmetropolitan groupings of State economic areas. However, in line with the findings for most of the other variables, religious affiliation exhibits noticeable differences between metropolitan and nonmetropolitan areas. With but one exception (the Middle States) every region shows greater affiliation in its metropolitan segment than in its nonmetropolitan segment. The magnitude of this difference varies among the regions, with the Northeast showing a difference more than twice as large as that exhibited by any other region. It should also be noted that, except for the Northeast and Far West which retain their extreme rank positions, religious affiliation in metropolitan areas shows remarkably similar values in the other four regions.

In a later section of this chapter some further comments will be made about the deviant positions of the Northeast and Far West with respect to religious affiliation. However, on the basis of the data in Table 5 one may tentatively conclude that in most of the regions religious affiliation is noticeably related to the metropolitan-nonmetropolitan status of regional population subgroups.

The Intraclass Correlation

The comparison of regional means discussed in the previous section represents a limited kind of analysis of regional diversity. Although some interesting differences were observed among regions and between the metropolitan and nonmetropolitan segments of regions, there still remains the question--how much of this apparent regional diversity is genuine? That is to say,

to what extent are regions real areas of distinctive group behavior? An answer to this question requires not only a consideration of how much regions differ from one another in terms of summary measures such as regional means but also a consideration of how accurate these regional means are as descriptions of the group behavior of all regional inhabitants. Fortunately, the intraclass correlation discussed in Chapter 2 comprehends both these considerations and provides a quantitative measure of the extent to which behavior is regionally distinctive.

Table 6 presents intraclass correlations for each dependent variable as it appears in each of our three State economic

TABLE 6

VALUES OF RHO FOR REGIONAL DISTRIBUTIONS OF SELECTED VARIABLES BY STATE ECONOMIC AREA GROUPINGS*

Selected Dependent Variables	State Economic Area Groupings		
	Total Economic Areas	Metropolitan Economic Areas	Nonmetropolitan Economic Areas
Fertility	.093	-.028	.142
Voters Voting	.753	.774	.749
Democrats	.611	.589	.652
Education	.368	.194	.439
Religion	.194	.353	.133
N	443	149	294

*On the basis of F-tests, all values of rho in this table are statistically significant at the one per cent level except that for fertility in metropolitan areas.

area groupings. For the total regional complements of these State economic areas it can be seen at a glance that there are

great differences among the variables in the amount of behavioral variation associated with regional location. Only nine per cent of the total variation in fertility among State economic areas is associated with regional subdivisions of the country. This means that over 90 per cent of the fertility variation is attributable to factors other than regional location. On the other hand, the two voting variables show high correlations, with 75 per cent of the variation in voter turnout and 61 per cent of the variation in Democratic voters being associated with regions. These rho values are equivalent to r's of .868 and .782, respectively. The two variables with intermediate rho values show regionalism "explaining" a little over one-third of the variation in educational achievement and a bit less than one-fifth of the variation in religious affiliation.

On the basis of these measures it is clearly evident that for the country as a whole the most regionally distinctive types of behavior studied here are those of voting followed by educational achievement, religious affiliation, and finally fertility. Furthermore, considering the correlation values often found in social science research, at least the first three and possibly the fourth types of behavior show regional relationships of more than passing interest.

Extending the analysis to the metropolitan and nonmetropolitan segments of the total population, several findings of further interest appear. An examination of the last two columns of Table 6 shows quite clearly that the intraclass correlations for the total group of State economic areas discussed above are really net products of measures quite different in magnitude.

For every variable except those pertaining to voting behavior the difference between metropolitan and nonmetropolitan rho values is rather large—the higher value being over twice the magnitude of the lower value in each case. Only one of the five variables—namely, religious affiliation—shows a substantially higher correlation for metropolitan areas, while two of the variables—fertility and education—show a much smaller correlation for this metropolitan segment of the population.

The negative value of the metropolitan fertility correlation perhaps merits some comment. This rho value of -.028 indicates no significant fertility variation associated with regions. Although Table 1 does show some interregional diversity with respect to mean fertility values, this diversity is so small that the corrective factor in the correlation formula which adjusts for chance fluctuations serves to reduce rho to a meaningless value. It should also be pointed out that while the fertility variation associated with regions is over half again as large for nonmetropolitan areas as for the total group of State economic areas, it still leaves over 85 per cent of the total variation "unexplained." In short, the relationship of regionalism and native white fertility appears from these data to be rather small.

Another notable feature of Table 6 is the stability of the regional correlations for voting behavior. With respect to voter turnout a difference of only 2.5 percentage points distinguishes metropolitan from nonmetropolitan areas. For Democratic voters this difference is greater but is still relatively small, amounting to only 6.3 percentage points. It should be pointed out here that, while these two voting variables are rather stable

for the three State economic area groupings, they are not identical. Approximately three-quarters of the variation in voter turnout is associated with regions. However, less than two-thirds of the variation in Democratic voting is so "explained." The differences in rho value between these two variables are of a magnitude which equals or exceeds some of the total rho values for other variables in Table 6.

Turning to the variables for education and religion, it is interesting to observe that the metropolitan-nonmetropolitan distinction affects these variables in opposite ways. The regional distinctiveness of educational achievement is greater in nonmetropolitan areas while that of religious affiliation is greater in metropolitan areas. Indeed, the rho value for religious affiliation in metropolitan areas is the highest rho computed for the three non-political variables, while in nonmetropolitan areas it is the lowest, even falling below the fertility correlation which has already been referred to as rather small.

In general, these intraclass correlations in Table 6 indicate fairly clearly that to a substantial degree the regional distinctiveness of at least three and possibly four of the five types of behavior measured here is contingent upon whether that behavior occurs in metropolitan or nonmetropolitan areas of organization. Since the metropolitan-nonmetropolitan distinction serves to identify population groups involved in different types of social structure, it would seem that these data alone tend to support at least presumptively the general hypothesis that regional differences in social structure are significantly related to regional differences in behavior.

Components of the Intraclass Correlation

In discussing the character of the intraclass correlation in the previous chapter it was remarked that the intraclass correlation is really a measure of the effects of two factors operating simultaneously--namely, the variation of unit values around their respective regional means and the variation of regional means around a grand mean for the whole country. Since the sum of these two sources of variance represents the total variance for the entire distribution of unit values, it can be seen that the value of rho is contingent upon the magnitudes of both variance components. In other words, since rho is basically a ratio of the interregional variance to the total variance, the value of rho will vary directly with the size of the interregional variance and inversely with the size of the intraregional variance. It would, therefore, be of interest to examine the components of the rho values presented above in Table 6 because they exhibit considerable variety in their magnitudes when the various groupings of State economic areas are compared. The reasons for this can only partly be inferred from data on the regional means, themselves, as they appear in Tables 1 through 5. Accordingly, in Table 7 there are presented all of the components of variance used in computing each intraclass correlation discussed above and found in Table 6.

The most striking thing immediately seen in Table 7 is that metropolitan areas have quite different components of variance from those for the total and the nonmetropolitan groupings of State economic areas. Indeed, comparing the metropolitan and nonmetropolitan segments of the regional populations, two general

TABLE 7

VALUES OF INTERREGIONAL AND INTRAREGIONAL VARIANCE FOR DISTRIBUTIONS OF SELECTED VARIABLES BY VARIOUS STATE ECONOMIC AREA GROUPINGS*

Selected Variables	Interregional Variance			Intraregional Variance		
	Total Econ. Areas	Metrop. Econ. Areas	Nonmetrop. Econ. Areas	Total Econ. Areas	Metrop. Econ. Areas	Nonmetrop. Econ. Areas
Fertility	7,144	305	7,527	60,772	22,993	39,639
Voters	407.7	351.5	440.8	131.3	97.2	143.8
Democrats	187.4	117.1	236.3	116.7	76.3	122.4
Education	4.7	1.0	7.4	7.9	3.3	9.0
Religion	30.7	49.5	22.4	118.9	81.2	126.3

*See Table 6 for the values of rho based on these components of variance. The total variance values have been omitted for reasons of space. They can, however, be easily estimated by summing within each State economic area grouping the interregional and intraregional variance values listed above.

observations should be noted. First, for every behavior variable except religion the metropolitan interregional variance is smaller than the nonmetropolitan interregional variance. Second, for every behavior variable, excepting none, the metropolitan intraregional variance is smaller than the nonmetropolitan intraregional variance.

The significance of these two observations for the value of rho is fairly clear. The fact that the diversity of regional mean values (i.e., interregional variance) is less for metropolitan areas indicates that rho should be lower for metropolitan areas than for nonmetropolitan areas, assuming other things are equal. Of course, in point of fact this assumption of ceteris paribus conditions cannot be made, because Table 7 also reveals that metropolitan areas have a smaller intraregional variance than do nonmetropolitan areas. This tends to make the metropolitan rho higher than the nonmetropolitan correlation.

Stating this point somewhat differently, the data in Table 7 show that the internal homogeneity of regions is greater for the metropolitan segment than the nonmetropolitan segment of the population. This is true for each of the five variables measured. However, accompanying this increased internal, regional homogeneity for metropolitan areas is a reduction in the amount of external heterogeneity among regions. This is true for all variables except religious affiliation. Keeping in mind that the value of rho varies directly with the degree of internal homogeneity and with the degree of external heterogeneity, it can be seen that when metropolitan and nonmetropolitan rho values are compared, their differences are the net products of changes in

variance which are exerting opposite influences. That is to say, the variance changes tend to cancel one another as far as the values of rho are concerned. Regions become at one and the same time more similar both internally and externally when their metropolitan areas are compared. Since, however, the comparison of metropolitan and nonmetropolitan rho values does reveal some substantial differences, one can attribute this only to differences between the two components of variance in their amounts of change. This is clearly evident in the data of Table 7.

With respect to fertility, there can be no doubt that the negative rho of -.028 for metropolitan areas is due to the negligible amount of interregional variance, viz., 305. This negligible value more than compensates for the greater internal homogeneity of the regions' metropolitan areas as reflected in the intraregional variances of 39,639 for nonmetropolitan areas and 22,993 for metropolitan areas.

Turning to the voter turnout variable, it will be recalled that metropolitan and nonmetropolitan segments of the population have rho values of .774 and .749, respectively. This slightly higher value for metropolitan areas can be seen to occur because of a greater intraregional homogeneity of metropolitan areas and despite a greater interregional homogeneity of metropolitan areas.

This same pattern of metropolitan-nonmetropolitan differences holds for Democratic voters, except that the increased similarity among regions (i.e., greater interregional homogeneity) more than compensates for the greater intraregional homogeneity of metropolitan areas. Accordingly, the rho value for metropolitan

areas is somewhat lower than that for nonmetropolitan areas.

The picture of metropolitan-nonmetropolitan differences in regard to educational achievement is rather dramatic. Here, the patterns of change in the two components of variance are the same as those already witnessed for the fertility and the two voting variables. However, the magnitude of the changes is quite large. Although the intraregional variance for metropolitan areas is only a little over one-third that for nonmetropolitan areas, this still represents but a modest change compared to the reduction in interregional variance which shows the metropolitan value to be only 13 per cent of the nonmetropolitan variance. In this situation it is not surprising to find that the value of rho for metropolitan areas is far less than one-half that for nonmetropolitan areas.

It is only with respect to religious affiliation that we find for metropolitan-nonmetropolitan comparisons a noncontradictory pattern of variance changes so far as rho values are concerned. Here, the variance within regions shows greater homogeneity for metropolitan areas, while the variance among regions shows greater heterogeneity for metropolitan areas. It is in this latter variance component—greater interregional heterogeneity for metropolitan areas—that we find the feature which distinguishes religious affiliation from all the other dependent variables. Some effort to account for this would seem to be in order here.

In Table 5, discussed earlier, the distributions of regional means for religious affiliation were presented. It will perhaps be recalled that in every region except the Middle States the metropolitan areas show higher proportions of religious

affiliation than nonmetropolitan areas. Furthermore, the greatest difference by far is exhibited by the Northeast, which has mean values of 53.4 per cent and 42.3 per cent for metropolitan and nonmetropolitan areas, respectively. There is little doubt that a large part of the rho value for metropolitan distributions of this religion variable can be attributed to this rather deviant Northeast region. This is attested by the fact that a computation of rho which excludes the Northeast yields a value of .137, as compared to .353 when the Northeast is included.

The question of why the Northeast shows such a disproportionately large amount of religious affiliation especially in its metropolitan areas may be answered at least tentatively. It is the writer's suggestion that this high degree of affiliation is attributable to a heavy representation of Roman Catholics in the population. Evidence in support of this is provided by the data in Table 8. These data clearly show that, not only does the

TABLE 8

PROPORTION OF TOTAL POPULATION IN EACH REGION WHICH IS ROMAN CATHOLIC, BY VARIOUS STATE ECONOMIC AREA GROUPINGS

Regions	State Economic Area Groupings		
	Total Economic Areas	Metropolitan Economic Areas	Nonmetropolitan Economic Areas
Northeast	26.1%	29.0%	16.7%
Southeast	4.0	8.3	2.8
Middle States	16.0	19.9	12.0
Northwest	9.6	11.0	9.3
Southwest	10.0	11.5	9.4
Far West	12.8	13.9	10.1

Northeast as a whole have a disproportionate representation of Roman Catholics in its population, but that this disproportion is most noticeable in the metropolitan segment of the Northeast population. If the assumption being made here is correct--namely, that the loyalty of the Roman Catholic Church membership tends to exceed that for any other major denomination--it may be expected that populations with heavy representations of Roman Catholics will exhibit a high degree of religious affiliation.[2]

Parenthetically, it is interesting to note that this disproportionate representation of Roman Catholics in the Northeast is not apparently reflected in fertility behavior. This may in part be due to the highly urbanized character of life in the Northeast which counteracts the high fertility potential of the Roman Catholic element in the population. Indeed, it is possible that the mean fertility value of metropolitan areas in the Northeast is as high as it is because of the large number of Roman Catholic residents in these metropolitan centers. These counteracting forces of urbanism and Catholicism may also be operating to influence fertility patterns in the Middle States region. Although the evidence supporting this idea is not entirely persuasive, it is interesting to observe that with respect to urban development the three regions which rank highest by far are the Northeast, the Middle States, and the Far West. The expectation

---

[2]Quite apart from this assumption about membership loyalty, it should be noted that Roman Catholic membership includes all persons baptized "in the Faith." This means that the proportion of affiliated persons in a total population which has a large representation of Roman Catholics will be inflated by the inclusion of baptized children. Most Protestant denominations do not report in their membership records children under age 14.

that this high degree of urbanization will be reflected in a depressed fertility measure is not borne out by the mean net reproduction rates for the metropolitan areas of the Northeast and Middle States (see Table 1). In fact, if anything, these data show differences which place the Northeast and Middle States regions rather high on the metropolitan fertility scale. Something would seem to be counteracting the fertility-depressing forces of urbanism. The suggestion here is that the data in Table 8 may throw some light on this apparent anomaly, for it can be clearly seen that the only regions which have disproportionately large complements of Roman Catholics in their metropolitan areas are the Northeast and the Middle States.

To summarize the general results of this examination of the intraclass correlation components, it can be said that for every variable the metropolitan segment of the population shows a relatively high degree of intraregional homogeneity. Furthermore, for every variable except religious affiliation the differences among regions are reduced when only the metropolitan areas are considered. The larger degrees of intraregional homogeneity tend to inflate the values of rho for metropolitan areas while the smaller degrees of interregional heterogeneity tend to deflate these correlations. Consequently, the differences in rho values between metropolitan and nonmetropolitan areas emerge as net products of contradictory changes in the variance components. These results seem to indicate that metropolitan areas are in some respects extra-regional in character. That is, they tend to reduce regional distinctiveness. The only substantial contradiction of this appears for the religious affiliation variable. The

major source of deviation here resides in the Northeast, which exhibits a very high degree of religious affiliation that seems to be attributable to a large Roman Catholic element in the population.

Comparison of Internal Homogeneity among Regions

The analysis in the previous section was restricted to a consideration of the two main components of variance with particular reference to how changes in these components affect the intraclass correlation. Although several insights were derived from this analysis, nothing was learned about the degrees of internal homogeneity for the individual regions themselves. Since one of the important criteria for delineating a regional pattern is homogeneity within regions, it would be of interest to examine the extent to which this criterion is met in the regional pattern being studied here. The standard deviations presented in Table 9 provide basic evidence for this kind of assessment together with an opportunity for comparing the homogeneity of each region's metropolitan and nonmetropolitan segments.

It is apparent from even a casual examination of these data that metropolitan areas are, generally speaking, more homogeneous than either the total complement of State economic areas or the nonmetropolitan segments of the total population. This is true for every region except the Far West, where the relationships are somewhat erratic. It can also be seen, however, that considerable variety is exhibited with respect to which regions and variables manifest this high degree of metropolitan homogeneity relative to that obtaining for nonmetropolitan areas.

TABLE 9

STANDARD DEVIATIONS OF SELECTED VARIABLES FOR EACH REGION AND FOR THE UNITED STATES BY STATE ECONOMIC AREA GROUPINGS

Regions	Fertility	Voters	Democrats	Education	Religion
Total State Economic Areas					
Northeast	251	11.0%	10.3%	1.8%	13.3%
Southeast	295	15.2	13.8	4.0	10.1
Middle States	212	6.9	7.8	1.2	9.1
Northwest	216	4.4	8.7	.9	13.8
Southwest	249	17.0	13.1	4.7	11.3
Far West	156	3.7	5.6	1.7	7.0
U.S. Total	261	23.2%	17.4%	3.6%	12.3%
Metropolitan State Economic Areas					
Northeast	154	8.3%	6.4%	1.8%	10.8%
Southeast	154	15.0	12.4	2.1	8.7
Middle States	152	5.9	6.6	1.1	7.3
Northwest	123	4.5	7.2	.6	8.9
Southwest	172	13.2	12.4	2.8	11.0
Far West	127	4.1	5.7	2.0	5.8
U.S. Total	153	21.2%	13.9%	2.0%	11.5%
Nonmetropolitan State Economic Areas					
Northeast	262	13.5%	11.5%	1.7%	13.4%
Southeast	226	15.2	14.1	4.1	10.2
Middle States	151	6.6	6.6	1.2	10.1
Northwest	179	3.7	8.6	.9	14.4
Southwest	172	17.5	13.2	5.1	11.1
Far West	103	4.0	5.8	1.4	7.1
U.S. Total	217	24.2%	18.9%	4.1%	12.2%

A comparison of metropolitan and nonmetropolitan standard deviations in fertility reveals that a substantially greater metropolitan homogeneity is most evident in the Northeast, the Southeast, and the Northwest. It may be noted that these are the regions with the highest nonmetropolitan sigmas. For voter turnout, relative metropolitan homogeneity is most prominently displayed in the Northeast and Southwest regions, while the variables for Democratic voting finds the Northeast alone exhibiting what appears to be a substantial difference in homogeneity between metropolitan and nonmetropolitan areas. These regions also have rather large nonmetropolitan sigmas. With respect to educational achievement it is the two southern regions, which have large nonmetropolitan standard deviations, that show the greatest reduction in sigma value for metropolitan areas. For religious affiliation all regions except the Southwest exhibit noticeable increases in homogeneity for their metropolitan populations, with the Northwest's appearing most conspicuous in this respect. And again, it is this region—the Northwest—which has a high (indeed the highest) standard deviation for the nonmetropolitan population.

In general, when all variables are considered, it appears that all regions except for the Far West and perhaps the Middle States show substantially greater degrees of homogeneity in their metropolitan areas than in the nonmetropolitan segments of their populations. The most conspicuous regions in this regard are the Northeast and Southeast. It also appears that the greatest differences in homogeneity between metropolitan and nonmetropolitan areas occur in those regions which exhibit relatively great

diversity in the behavior of their nonmetropolitan populations.

The somewhat deviant position of the Far West has been noted before in several contexts, and is a phenomenon difficult to account for. The thought has occurred that the Far West might be an "odd" region because of the important role it has played as a destination for rather substantial migration from areas all over the country. For example, in 1940 only 44 per cent of the people living in the Far West region were residing in the States in which they were born. This would seem to suggest that the Far West's population may be a regional potpourri. While this may be true, it is not reflected in summary measures of behavior for State economic areas. In point of fact, the Far West exhibits standard deviation values (for nearly every variable) which rank it very high on the scale of intraregional homogeneity. This holds true regardless of whether the total population or its metropolitan or nonmetropolitan segments are the focus of study. It may be that the selectivity of migrants to the Far West has resulted in a resident population whose multi-regional composition is only a nominal one. That is, the migrants may not be genuinely representative of the regions from which they came. On the other hand, even a great diversity of regional population types in the Far West may not appear in our analysis as evidence of heterogeneity <u>if</u> the distribution of these diverse population types is fairly random. In other words, if there is no clustering or spatial concentration of such types, the kinds of basic unit measures used in this study will not reveal much diversity. In any event, it is interesting to observe that a region which, by virtue of its growth, is probably undergoing greater change than

any other region in the country still exhibits a very high degree of internal homogeneity with respect to the variables examined here.

Turning back for a moment to the findings for the other five regions, it should be pointed out that the general pattern which shows greater homogeneity for the metropolitan segments of each region's population is a pattern which, considered by itself, tends to accentuate the regional distinctiveness of metropolitan behavior. Stating the point in a different way, the size of the standard deviation may be viewed as a negative index of each region's intrinsic contribution to the value of rho. Thus, the smaller the sigma value, the "purer" is the region, so to speak, and hence (everything else being equal) the more distinctive is the region in the context of the total delineation. Of course, as it has been seen, everything else is not equal. The value of rho for metropolitan areas is substantially larger for only one variable—namely, religious affiliation. And this can be attributed mainly to a peculiarity in the population composition of one region—the Northeast. In short, it would seem that one can reasonably infer from the data so far presented that metropolitan types of social structure possess certain similarities which tend to level regional differences.

The discussion of the standard deviations presented in Table 9 has been focused primarily on comparisons of the three State economic area groupings with one another. These data, however, also provide an opportunity to compare the regions themselves with one another in terms of their degrees of internal homogeneity. To facilitate this kind of comparison, the standard

deviations in Table 9 have been converted to indexes which express each sigma value as a ratio to a common base. These ratios are presented in Table 10. The common base employed here is the standard deviation for the total intraregional variation for each variable—in other words, the average intraregional standard deviation for all regions considered together. Accordingly, the ratios vary in value around an average base value of 100. Any ratio which exceeds 100 indicates a region which is more heterogeneous than the average region, while a ratio falling below the base value of 100 indicates a region which is more homogeneous than the average with respect to a particular characteristic.

Examination of these ratios in Table 10 reveals rather clearly that the two southern regions together with the Northeast are, generally speaking, the least homogeneous of our six regions. This tends to hold true for all three of the State economic area groupings. On the other hand, of the three remaining regions—i.e., the Middle States, Northwest, and Far West—it is the Far West which is the most homogeneous area. It is only in the metropolitan segment that we find any index values for the Far West high enough even to approximate the average for all regions.

It is interesting also to note that the two southern regions appear to be the most heterogeneous of all the regions. This is particularly evident with respect to the two voting variables and education. Regardless of whether one looks at the total complement of State economic areas, the metropolitan segment, or the nonmetropolitan segment, the Southeast and Southwest exhibit much greater heterogeneity than the average region. That this should occur for the two voting variables, in particular,

TABLE 10

RATIOS OF REGIONAL STANDARD DEVIATIONS FOR SELECTED VARIABLES TO AVERAGE REGIONAL STANDARD DEVIATION FOR ALL REGIONS BY STATE ECONOMIC AREA GROUPINGS

Regions	Fertility	Voters	Democrats	Education	Religion
Total State Economic Areas					
Northeast	102	96	95	64	122
Southeast	120	132	128	143	93
Middle States	86	60	72	43	83
Northwest	88	38	80	32	127
Southwest	101	148	121	167	104
Far West	63	32	52	61	64
All Regions	100	100	100	100	100
Metropolitan State Economic Areas					
Northeast	101	84	74	100	120
Southeast	101	152	142	117	97
Middle States	100	60	76	61	81
Northwest	81	45	83	33	99
Southwest	113	133	142	156	122
Far West	84	41	66	111	64
All Regions	100	100	100	100	100
Nonmetropolitan State Economic Areas					
Northeast	132	112	104	57	120
Southeast	114	127	127	137	91
Middle States	76	55	59	40	90
Northwest	90	31	77	30	128
Southwest	86	146	119	170	99
Far West	52	33	52	47	63
All Regions	100	100	100	100	100

suggests that the so-called "Solid South" may perhaps more accurately be considered the least "solid" of all the regions. Of course, practically speaking, it may not make much difference under our present political system how much variation there is in southern voting behavior, just so long as more than 50 per cent of the vote goes Democratic. Nevertheless, it is interesting to observe that the South does show such great diversity in the degree of its "solid" Democratic affiliation.

Although the data in Table 10 provide for an interesting comparison of the regions, one to another, they are not well adapted to a comparison which permits an assessment of the success with which each variable assumes a regionally distinctive pattern. That is to say, the question of which variables exhibit the greatest intraregional homogeneity cannot be answered with confidence by an inspection of ratios whose base values are directly dependent upon the internal homogeneity of the several regions. Accordingly, ratio expressions of the intraregional standard deviations found in Table 9 have been recomputed by using as a base of 100 the standard deviation for the entire United States rather than the average intraregional sigma for all regions which is used in Table 10. These ratios, based on the distribution of values for the entire country, are presented in Table 11. Since these ratios express a relationship between two measures of dispersion which have no intrinsic relation to one another, they may be regarded as indexes of the extent to which any particular region is a significantly more homogeneous area than the country as a whole with respect to particular variables. In other words, if one of the criteria for delineating regions is internal homogeneity,

TABLE 11

RATIOS OF REGIONAL STANDARD DEVIATIONS FOR SELECTED VARIABLES TO STANDARD DEVIATION FOR ENTIRE UNITED STATES BY STATE ECONOMIC AREA GROUPINGS

Regions	Fertility	Voters	Democrats	Education	Religion
Total State Economic Areas					
Northeast	96	47	59	50	108
Southeast	113	66	79	111	82
Middle States	81	30	45	33	74
Northwest	83	19	50	25	112
Southwest	95	73	75	130	92
Far West	60	16	32	47	57
U.S. Total	100	100	100	100	100
Metropolitan State Economic Areas					
Northeast	101	39	46	90	94
Southeast	101	71	89	105	76
Middle States	99	28	47	55	63
Northwest	80	21	52	30	77
Southwest	112	62	89	140	96
Far West	83	19	41	100	50
U.S. Total	100	100	100	100	100
Nonmetropolitan State Economic Areas					
Northeast	121	56	61	41	110
Southeast	104	63	75	100	84
Middle States	70	27	35	29	83
Northwest	82	15	46	22	118
Southwest	79	72	70	124	91
Far West	47	16	31	34	58
U.S. Total	100	100	100	100	100

these ratios permit an assessment of how successfully this criterion has been met for each variable in each region.

Inspection of the ratios in Table 11 reveals that the Odum regional classification used here is most successful for voter turnout and least successful for the fertility variable. This holds true regardless of the different groupings of State economic areas, although among the nonmetropolitan population segments the ratios for religious affiliation tend to be as high if not higher than those for fertility. Generally speaking, the two voting variables exhibit the greatest amount of intraregional homogeneity. In no region does the standard deviation for either of these variables ever equal that for the entire country. The closest approximation of this occurs in the two southern regions where, for metropolitan areas, the amount of intraregional dispersion falls just 11 per cent below that for the country as a whole.

It is a little difficult to generalize about the variables for education and religion because they differ somewhat for metropolitan and nonmetropolitan areas. For education the metropolitan ratios for each region exceed the nonmetropolitan ratios, while for religious affiliation the contrary situation (with one exception) obtains. Both variables, however, tend to fall in between the extremes represented by fertility and voting behavior. Consequently, a general rank ordering of our five variables in terms of increasing degrees of intraregional homogeneity relative to the dispersion of values for the country as a whole would find fertility at the bottom, followed by religion and education, Democratic voters, and finally voter turnout.

The Influence of Structure on Regional Rank Positions

The whole question of whether regions constitute areas of distinctive behavior involves not only a consideration of the general extent to which this is true but also a consideration of what the specific pattern of differences is. In an effort to throw some light on this latter facet of the regionalism question, each region has been assigned rank positions based on the values of its mean and standard deviation for each variable and for each of the two major subgroupings of State economic areas—namely, the metropolitan and nonmetropolitan population segments. Thus, for example, with respect to fertility the regions are ranked from one to six, first, on the basis of their metropolitan mean fertility values and, second, on the basis of their nonmetropolitan mean fertility values. This same ranking procedure is carried through for the standard deviations of fertility. It can be seen, therefore, that each region receives four rankings for each variable. The point of interest here is to see if the regions shift their rank positions when metropolitan and nonmetropolitan types of social structure are compared. These data are presented in Table 12.

Turning first to the rankings of regional means, an inventory of the thirty metropolitan-nonmetropolitan comparisons reveals that eighteen of them (or 60 per cent) show no change. Of the twelve which are altered, only five shift more than one rank position. Furthermore, all but one of these five cases are found in a shift of the fertility pattern.

If one may regard a shift of only one rank position as a relatively small change, it appears that over 80 per cent of the

regional mean rankings remain fairly stable despite the structural differences which accompany metropolitan-nonmetropolitan types of living. This relative stability in regional rankings is quite impressive if one discards the fertility variable and focuses on the remaining four types of behavior. Here, 23 out of the 24 regional mean rankings do not change more than one rank position when metropolitan and nonmetropolitan areas are compared.

TABLE 12

RANK POSITIONS OF REGIONS WITH RESPECT TO MEAN VALUES AND STANDARD DEVIATIONS FOR EACH VARIABLE BY METROPOLITAN AND NONMETROPOLITAN GROUPINGS

Regions		Fertility		Voters		Democrats		Education		Religion	
		MA	NMA	MA	NMA	MA	NMA	MA	NMA	MA	NMA
Northeast	$\bar{X}$	3	5	3	3	6	6	3	3	1	1
	σ	3	1	3	3	5	3	4	3	2	2
Southeast	$\bar{X}$	4	1	6	6	1	1	1	1	3	4
	σ	2	2	1	2	2	1	2	2	4	4
Middle States	$\bar{X}$	2	4	2	1	4	5	5	5	4	2
	σ	4	5	4	4	4	5	5	5	5	5
Northwest	$\bar{X}$	5	3	1	2	5	4	6	6	2	3
	σ	6	3	5	6	3	4	6	6	3	1
Southwest	$\bar{X}$	1	2	5	5	2	2	2	2	5	5
	σ	1	4	2	1	1	2	1	1	1	3
Far West	$\bar{X}$	6	6	4	4	3	3	4	4	6	6
	σ	5	6	6	5	6	6	3	4	6	6

Although this rank order pattern of regions seems to persist in a more or less rough fashion despite differences in the structural conditions of life for each regional population, one may not assume that these structural differences are unimportant

for behavior. It should be recalled that in Table 6 the measures of dispersion for regional means (i.e., the interregional variance) show metropolitan values that are smaller for every variable except one than those found for nonmetropolitan areas. The one exception is religious affiliation, the deviation of which can be attributed to the population composition of the Northeast. It would seem, therefore, that what has occurred in the behavior patterns of metropolitan and nonmetropolitan populations is a convergence of regional differences in metropolitan areas. However, except for fertility behavior this convergence has not been great enough to alter radically the rank positions of the several regions. Thus, the pattern of metropolitan regional mean differences appears to be substantially that of nonmetropolitan differences "writ small."

With respect to the regional rankings of standard deviations, a comparison of metropolitan and nonmetropolitan areas shows essentially the same patterns as those just discussed for regional means. Eighty per cent of the comparisons show no change greater than one rank position, while 40 per cent of the rankings are identical for the two population segments. And once again it is in fertility behavior that most of the large rank changes appear.

Although the stability of this regional pattern of standard deviations is not so great as that for regional means, one cannot reasonably regard the over-all shifts as being revolutionary. As with the case of regional means, however, the persistence in metropolitan and nonmetropolitan areas of the same rank order pattern for intraregional homogeneity does not compromise

the significance of structural differences for behavior. It will be recalled that in Table 9 the sigma values for metropolitan areas are lower than those for nonmetropolitan areas in nearly every region and for nearly every variable. Although this shows that the dispersion of each region's values around its respective regional means is relatively small for metropolitan areas, the question still remains as to whether metropolitan and nonmetropolitan areas differ with respect to the interregional distribution pattern of their sigma values. In other words, it has already been demonstrated that the standard deviations for each region tend to be reduced in size for metropolitan areas. It would now be of interest to know whether this reduction in each region's sigma value is accompanied by a tendency for these regional values to converge, thereby reducing differences among regions in the extent of their internal dispersion.

To determine whether metropolitan areas exhibit this convergence of regional standard deviations requires the computation of a measure of dispersion for metropolitan and nonmetropolitan distributions of regional sigma values. Since there are six regions, we have a distribution comprised of six units, each of which has a sigma value for its metropolitan areas and a sigma value for its nonmetropolitan areas. Both the means and the sigmas for each of these distributions of sigma values have been computed and are presented separately for each variable in Table 13.

It can be clearly seen that for each variable the mean sigma value for the six regions is lower in metropolitan than in nonmetropolitan areas. Furthermore, for each variable the sigma measures find metropolitan areas exhibiting less interregional

dispersion than do nonmetropolitan areas. The conclusion appears to be obvious; viz., when metropolitan and nonmetropolitan areas are compared, there is a reduction in the amount of intraregional variance for metropolitan areas and also a reduction in the amount of interregional dispersion of this distribution of intraregional variances. In short, not only does each individual region tend to become more homogeneous internally when metropolitan areas alone are considered, but there also tends to be for all regions a greater similarity in the amount of internal homogeneity exhibited by each. This seems to testify strongly to the leveling influence of metropolitan types of social structure on interregional differences.

TABLE 13

MEANS AND SIGMAS OF REGIONAL DISTRIBUTIONS OF INTRAREGIONAL STANDARD DEVIATIONS FOR SELECTED VARIABLES BY METROPOLITAN AND NONMETROPOLITAN AREAS

Selected Variables		Metropolitan Economic Areas	Nonmetropolitan Economic Areas
Fertility	$\bar{X}$	147	182
	σ	17	51
Voting	$\bar{X}$	8.5%	10.1%
	σ	13.3	17.4
Democrats	$\bar{X}$	8.4%	10.0%
	σ	8.9	10.0
Education	$\bar{X}$	1.7%	2.4%
	σ	2.2	5.1
Religion	$\bar{X}$	8.8%	11.0%
	σ	5.8	7.5

Summary

The analysis of regional means for each of the types of behavior examined in this study reveals in each case differences among some of the regions which appear quite substantial. Furthermore, the general rank order pattern of regional means tends to hold up for both the metropolitan and nonmetropolitan segments of the population. However, for every variable except religious affiliation the differences among these regional means (as measured by interregional variance) are smaller for the metropolitan than for the nonmetropolitan areas. This indicates that regional differences tend to converge in metropolitan areas, although this convergence is not so great as to rearrange the rank positions of the regions with respect to one another. The exceptional case of religious affiliation rests on the deviation of one region--the Northeast, which exhibits a disproportionately large amount of religious affiliation in metropolitan areas. This may perhaps be attributed to a large representation of Roman Catholics in the metropolitan population of the Northeast.

Summary measures of the extent to which regions constitute areas of distinctive behavior are provided by intraclass correlations (rho). These measures exhibit considerable variety for the different types of behavior. The two voting variables show the greatest amount of association with regions while fertility shows the least.

When metropolitan and nonmetropolitan areas are compared with respect to their intraclass correlations, the variables for fertility, Democratic voting, and education show reductions for the metropolitan population. Reasons for the differences between

metropolitan and nonmetropolitan rho values can be found by examining the two major components of this correlation measure—namely, the interregional variance and the intraregional variance. As a reference to the correlation formula will clearly show, the value of rho varies directly with the magnitude of the interregional variance and inversely with the magnitude of the intraregional variance.

An examination of these two components of the total variance reveals that for every variable except religion the difference between metropolitan and nonmetropolitan rho values is a net product of variance changes which contradict one another. For every type of behavior the metropolitan intraregional variance is smaller than the internal variance for nonmetropolitan areas. This tends to inflate the value of rho for metropolitan areas because intraregional homogeneity is increased. However, except for religion, every variable also exhibits an interregional variance which is smaller in metropolitan areas than in nonmetropolitan areas. This tends to deflate the value of rho for metropolitan areas, because interregional heterogeneity is decreased. The only noncontradictory pattern of variance change is found with respect to religious affiliation, for here interregional heterogeneity is relatively great for metropolitan areas while intraregional homogeneity is reduced in these areas. As noted before, the increased variance for metropolitan regional means can be attributed to the Northeast with its high proportion of Roman Catholics.

In an effort to evaluate the relative success with which the delineation of each region meets the criterion of internal

homogeneity, standard deviations were computed for the distributions of every variable in each region and for metropolitan and nonmetropolitan areas separately. The results clearly show the Far West as the most homogeneous of the six regions, while the two southern regions generally exhibit the least internal homogeneity. The Northeast is somewhat erratic, but tends to resemble the South in this respect, while the Middle States region falls only a little below the high level of homogeneity recorded for the Far West. The Northwest exhibits a rather erratic pattern which defies generalization.

A comparison of the degrees of homogeneity in metropolitan and nonmetropolitan areas reveals that metropolitan areas tend to be more homogeneous than nonmetropolitan areas for nearly every region. The only noticeable exception to this is the Far West, where the distinction between metropolitan and nonmetropolitan structure seems to have negligible consequences for measures of homogeneity. In line with the findings for regional means, the general rank order pattern of regional sigmas tends to remain substantially unchanged for metropolitan and nonmetropolitan areas. However, there does seem to be a tendency for convergence of these regional sigma differences among the several regions' metropolitan areas. In other words, regional measures of dispersion tend to become both reduced in size and more similar to one another when metropolitan areas alone are considered.

These results viewed together with the convergence of regional mean values for metropolitan areas suggest that metropolitan areas are organizational units which tend to level regional differences in behavior. Since these metropolitan

concentrations of population represent the most highly urbanized and industrialized types of social structure in the larger society, it would appear that these findings tend to support the hypothetical expectations presented in the first chapter--namely, that similar types of social structure pose similar types of problems which tend to be solved by similar types of behavior; and further, that regional differences which emerge under conditions of isolation and geographical distinctiveness tend to dissolve when interregional contact becomes extensive and technology becomes more highly developed.

CHAPTER IV

THE FINDINGS: METROPOLITAN STATE ECONOMIC AREAS WITH APPLIED CONTROLS

The findings in this chapter represent an extension of the analysis of metropolitan State economic areas reported in Chapter III. Essentially, the task here is to examine the regional distinctiveness of behavior in metropolitan State economic areas when various conditions of life in these areas are controlled. The reader is reminded of the basic hypothesis which guides this investigation--namely, that similar conditions of life pose similar problems which tend to be solved by similar types of behavior. Since our measure of regional differences in behavior is the intraclass correlation (rho), it is expected that the application of controls will bring about a reduction in the value of this correlation measure.

Preliminary Remarks and the Specification of Controls

It will be recalled from the discussion of methodology in Chapter II that data on thirteen independent variables were tabulated for purposes of control. Furthermore, it will be recalled that control is to be exercised by selecting from within each region those State economic areas which have certain characteristics. Since this procedure involves discarding areas with inappropriate characteristics, care must be taken not to control with such rigor

that the case mortality rate threatens the respectability of the sample N. The point of repeating this caution (which was discussed in Chapter II) is simply to draw attention to the fact that the representation of metropolitan units in each of the three Western regions is rather small. Indeed, the Northwest has a total of only six metropolitan areas, which means there is not much opportunity for control selection within this grouping.

Another point which should be noted as pertinent for the control procedure involves use of the intraclass correlation measure. Stated briefly, the smaller the sample size, the more important is the correction factor for chance fluctuation in the determination of the value of rho.[1] This can easily be seen from the formula:

$$\text{rho} = \frac{\sigma_b^2}{\sigma^2} - \left[\frac{M}{N-M} \cdot \frac{\sigma_w^2}{\sigma^2} \right]$$

Since the correction term (in brackets) constitutes a downward adjustment of the basic rho ratio, it is possible for our control procedure to bring about a reduction in the value of rho that is entirely due to a reduction in sample size. Needless to say, the

---

[1] It may also be seen from this formula that the absolute minimum value of rho is determined by the size of the ratio $\frac{M}{N-M}$. In the extreme case where the means of all M classes are identical, σ_b^2 will equal zero and σ_w^2 will equal σ^2. Following the formula, the basic rho ratio of $\frac{\sigma_b^2}{\sigma^2}$ will equal zero and $\frac{\sigma_w^2}{\sigma^2}$ will equal one. We will have, therefore, a formula whose values read as follow: $\text{rho} = 0 - \frac{M}{N-M} \cdot 1$. Obviously, the minimum value of rho will approximate zero in direct proportion to the size of N.

researcher did not desire any support for his hypothesis to be contingent on this kind of statistical artifact.

The import of these general introductory remarks about the application of controls is perhaps self-evident. The researcher was compelled to compromise the rigor of his controls in order to assure an adequate representation of units in each region and a total sample size which does not discriminate greatly in favor of the hypothesis being tested.

It is probably apparent from the remarks above that a good deal of the control procedure had to be "played by ear," so to speak. Not until the data had been tabulated and IBM runs made did the researcher know how the units in each region were distributed with respect to each independent variable. Furthermore, once this information was available there still remained the problem of deciding specifically what control groupings to select. The range of possibilities is enormous, and obviously no attempt could be made to exhaust them. Accordingly, somewhat arbitrary selections were made which at the time seemed reasonable but which should properly be regarded as frankly exploratory efforts.

Twelve control groups were finally selected. These groups involve eight independent variables which are used either singly or in conjunction with one another. The decisions as to which variables to use and how to use them were based in large part on an examination of the distributions of each variable in each region. These data are presented in Appendix A. Even the most casual glance at these data reveals noticeable differences in the distributions of most of these independent variables. The eight variables finally selected for use as controls are listed below,

together with the appendix reference for their regional distributions:

1. Proportion of the total population which is urban (Appendix A, Table 1).
2. Proportion of the total population which is rural farm (Appendix A, Table 3).
3. Proportion of the total population which is nonwhite (Appendix A, Table 4).
4. Proportion of the employed labor force in manufacturing (Appendix A, Table 6).
5. Proportion of the employed labor force in wholesale and retail trade (Appendix A, Table 7).
6. Rural farm level of living index (Appendix A, Table 10).
7. Proportion of rural farm labor force who are unpaid family workers (Appendix A, Table 11).
8. Proportion of rural farm labor force in nonfarm occupations (Appendix A, Table 12).

The actual application of these independent variables as controls took the form of establishing three variables as base controls and then applying in serial fashion other variables as additional controls. The base controls established were two in number, viz., (a) less than 85 per cent of the total population urban and less than 25 per cent of the total population rural farm; (b) less than 35 per cent of the employed workers in manufacturing.

The urban-rural farm control is designed to insure the selection of units which are not overwhelmingly urban and yet are clearly not rural farm in any extremely important degree. This appears to be consistent with the concept of the metropolitan area as consisting of a large central city which mediates relationships for an extended hinterland. Furthermore, this control, by

establishing an urban ceiling, functions indirectly as a population size control. All of the extremely large metropolitan centers such as New York, Chicago, Los Angeles, etc. have populations exceeding 85 per cent urban. This urban control, therefore, avoids the intuitive absurdity of combining these giant metropolises with such communities as Wichita Falls, Texas, Bay City, Michigan, Waterloo, Iowa, and Albuquerque, New Mexico.

The control on manufacturing employment is designed to remove the extremely heavily industrialized communities from the Northeast and Middle States regions. None of the three Western regions exhibits much industrial manufacturing, so that to render the regions somewhat comparable in this respect requires bringing the Northeast and Middle States down to the level of Western manufacturing employment.

The specific control ranges for the other independent variables can be described simply and briefly. With respect to the nonfarm occupation variable, at least 35 per cent of the rural farm labor must be employed in nonfarm occupations. This means that less than two-thirds of the rural farm population can be regarded as agricultural in character when this variable is used. For unpaid family workers, less than 10 per cent of the rural farm labor force may be so employed. This represents a control which minimizes marginal, subsistence farming and by implication places most of the rural farm population in a commercial nexus. The rural farm level of living index poses some difficulties because of the great discrepancies in the regional distributions of this variable. It was decided to select an upper-middle range here whose limits run from an index value of 85 at the bottom to 124 at

the top. This means that the areas with the highest levels of living (particularly in the Northeast and Middle States) are discarded together with the areas of lowest living standard (particularly in the Southeast).

As can well be anticipated, controlling for the nonwhite proportion of the population means discarding a large part of the Southeast region. Some representation in this region is maintained, however, by rejecting only those areas with 15 per cent or more of their populations classified as nonwhite. Incidentally, "nonwhite" does not include persons of Mexican ancestry. Because of this, the case mortality rate in the Southwest region is not too damaging.

The last variable to be considered is that referring to wholesale and retail trade employment. The only control which could assure some representation in each region was one that selected areas with high proportions of their workers so employed. Accordingly, all areas with less than 20 per cent of their workers employed in wholesale and retail trade were rejected. This involved a rather high loss of cases in the Northeast, but the other regions are fairly well represented.

With these brief remarks in mind about the kinds of "control breaks" used for each independent variable, we may now consider how the final control groups were set up. Ideally, the greater the number of controls which are applied simultaneously, the greater is the structural similarity among the units of the several regions, and, hypothetically, the greater is the similarity in behavior. Unfortunately, as has been noted before, the loss of cases severely limits both the number of controls and the

rigor with which they can be applied. Since some criteria had to be established to guide the control procedure, it was rather arbitrarily decided that no region's representation could fall below three units, and the total representation for all regions would not be permitted to fall below 60 units. The term, unit, of course, refers to a State economic area. With these conditions serving as minimum standards to be met, the following twelve control groups were set up:

S-1: Less than 85 per cent urban, less than 25 per cent rural farm.

S-2: Less than 85 per cent urban, less than 25 per cent rural farm, 35 per cent or more of rural farm labor force in nonfarm occupations.

S-3: Less than 85 per cent urban, less than 25 per cent rural farm, less than 10 per cent of rural farm labor force who are unpaid family workers.

S-4: Less than 85 per cent urban, less than 25 per cent rural farm, rural farm level of living index between 85 and 124, inclusive.

S-5: Less than 85 per cent urban, less than 25 per cent rural farm, rural farm level of living index between 85 and 124, less than 15 per cent nonwhite.

S-6: Less than 85 per cent urban, less than 25 per cent rural farm, less than 35 per cent of employed workers in manufacturing.

S-7: Less than 35 per cent of employed workers in manufacturing.

S-8: Less than 35 per cent of employed workers in manufacturing, 35 per cent or more of rural farm labor force in nonfarm occupations.

S-9: Less than 35 per cent of employed workers in manufacturing, less than 10 per cent of rural farm labor force who are unpaid family workers.

S-10: Less than 35 per cent of employed workers in manufacturing, rural farm level of living index between 85 and 124, inclusive.

S-11: Twenty per cent or more of employed workers in wholesale and retail trade.

S-12: Less than 15 per cent nonwhite.

The success with which these twelve control groups meet the requirements of regional representation and adequate sample size set forth above can be judged from the data presented in Table 14. In only four of the control groups (S-2, S-3, S-5, and S-8) does a region's representation fall to the specified minimum of three units, and only one control group (S-3) reaches the minimum of 60 units for the total sample size. This table can be used as a general reference table, for all of the data to be presented in this chapter consist of computations based on these distributions of State economic areas.

Before presenting data for the five dependent variables, one additional table for general reference would seem to be appropriate here. This is Table 15, which contains three basic, intraclass correlation measures for each separate group of State economic areas. The three basic values of rho are the following: (a) the minimum value of rho necessary to achieve significance at the five per cent level; (b) the minimum value of rho necessary to achieve significance at the one per cent level; (c) the absolute minimum value that can be achieved by rho.

The rest of this chapter is devoted to presenting the substantive findings for our five dependent behavioral variables as they appear in metropolitan State economic areas under controlled conditions. The presentation of these data will take the form of discussions of each variable separately. A summary section will attempt to tie together all the important general observations and

TABLE 14

NUMBER OF METROPOLITAN STATE ECONOMIC AREAS IN EACH REGION FOR THE TOTAL GROUP OF METROPOLITAN STATE ECONOMIC AREAS AND TWELVE SELECTED CONTROL GROUPS

Regions	Total Metrop. Areas	Control Groups*											
		S-1	S-2	S-3	S-4	S-5	S-6	S-7	S-8	S-9	S-10	S-11	S-12
Northeast	43	35	31	33	28	26	13	16	16	14	12	5	41
Southeast	35	28	20	19	16	5	26	32	20	18	17	13	7
Middle States	41	29	16	23	18	18	14	21	9	14	11	19	41
Northwest	6	5	3	3	4	4	5	6	3	4	5	6	6
Southwest	11	10	5	7	8	3	10	11	6	8	9	10	6
Far West	13	9	5	8	8	8	9	13	6	12	12	10	13
Total U.S.A.	149	116	80	93	82	64	77	99	60	70	66	63	114

*The specific controls used are as follow:

S-1: Urban < 85%, Rural farm < 25%.
S-2: " " " , " " " " , Nonfarm occupation = 35%+.
S-3: " " " , " " " " , Unpaid family workers < 10%.
S-4: " " " , " " " " , Farm level of living = 85-124.
S-5: " " " , " " " " , " " " " " " , Nonwhite < 15%.
S-6: " " " , " " " " , Manufacturing employment < 35%.

S-7: Manufacturing employment < 35%.
S-8: " " " " , Nonfarm occupation = 35%+.
S-9: " " " " , Unpaid family workers < 10%.
S-10: " " " " , Farm level living = 85-124.
S-11: Wholesale and retail trade employment = 20%+.
S-12: Nonwhite < 15%.

TABLE 15

MINIMUM VALUES OF RHO FOR SIGNIFICANCE LEVELS OF FIVE PER CENT AND ONE PER CENT; ABSOLUTE MINIMUM VALUES OF RHO; TOTAL N FOR TOTAL GROUP OF METROPOLITAN STATE ECONOMIC AREAS AND TWELVE SELECTED CONTROL GROUPS

Listed Items	Total Metrop. Areas	Control Groups*											
		S-1	S-2	S-3	S-4	S-5	S-6	S-7	S-8	S-9	S-10	S-11	S-12
5% Level	.052	.071	.110	.091	.107	.144	.114	.084	.156	.130	.140	.147	.073
1% Level	.086	.119	.186	.154	.181	.243	.192	.141	.266	.220	.236	.248	.123
Absolute Minimum Rho	-.042	-.054	-.081	-.069	-.079	-.103	-.084	-.064	-.111	-.094	-.100	-.105	-.056
Total N.	149	116	80	93	82	64	77	99	60	70	66	63	114

*The specific controls used are as follow:

S-1: Urban < 85%, Rural farm < 25%.
S-2: " " " , " " " " , Nonfarm occupation = 35%+.
S-3: " " " , " " " " , Unpaid family workers < 10%.
S-4: " " " , " " " " , Farm level of living = 85-124.
S-5: " " " , " " " " , " " " " " " " , Nonwhite < 15%.
S-6: " " " , " " " " , Manufacturing employment < 35%.

S-7: Manufacturing employment < 35%.
S-8: " " " " , Nonfarm occupation = 35%+.
S-9: " " " " , Unpaid family workers < 10%.
S-10: " " " " , Farm level living = 85-124.
S-11: Wholesale and retail trade employment = 20%+.
S-12: Nonwhite < 15%.

to assess the success of this control procedure in reducing the value of the intraclass correlation as a measure of behavioral diversity among regions.

Net Reproduction Rate

It will be recalled that the fertility measures for all metropolitan State economic areas considered together show no relationship to regional location. Indeed, the rho value of -.028 is just about as low as it can get--the absolute minimum value being -.042 as shown in Table 15. In this situation it would seem somewhat unreasonable to expect the rough controls applied in this study to effect any further reduction in the value of the intraclass correlation. It would, however, be of interest to see whether this lack of association between fertility and regional location is maintained when structural characteristics of metropolitan areas are manipulated. Conceivably, certain types of metropolitan areas may exhibit regional differences which are cancelled out when the total group is examined.

The data presented in Table 16 provide evidence needed to answer this question. The values of rho appearing in this table clearly show no significant relationship between metropolitan fertility behavior and regionalism. None of the control groups exhibits rho values that even come close to the minimum values necessary for statistical significance at the five per cent level (see Table 15).

The rather small fluctuations that do appear in the values of rho can be accounted for by the shifts in values of the interregional and intraregional components of variance which are also

TABLE 16

A LISTING FOR FERTILITY OF THE VALUES OF RHO, THE INTERREGIONAL VARIANCE, THE INTRAREGIONAL VARIANCE, AND N FOR THE TOTAL GROUP OF METROPOLITAN STATE ECONOMIC AREAS AND TWELVE SELECTED CONTROL GROUPS

Listed Items	Total Metrop. Areas	Control Groups*											
		S-1	S-2	S-3	S-4	S-5	S-6	S-7	S-8	S-9	S-10	S-11	S-12
Rho	-.028	.034	-.004	-.024	.050	.012	.043	-.025	-.043	-.078	-.064	.012	-.005
Interregional Variance . . .	305	1654	1393	829	2058	1917	2681	979	1807	410	791	1721	1211
Intraregional Variance . . .	22,993	17,977	17,995	18,846	15,131	16,455	20,091	25,455	27,604	28,054	23,276	14,454	24,184
N	149	116	80	93	82	64	77	99	60	70	66	63	114

*The specific controls used are as follow:

S-1: Urban < 85%, Rural farm < 25%.
S-2: " " " , " " " " , Nonfarm occupation = 35%+.
S-3: " " " , " " " " , Unpaid family workers < 10%.
S-4: " " " , " " " " , Farm level of living = 85-124.
S-5: " " " , " " " " , " " " " " " , Nonwhite < 15%.
S-6: " " " , " " " " , Manufacturing employment < 35%.

S-7: Manufacturing employment < 35%.
S-8: " " " " , Nonfarm occupation = 35%+.
S-9: " " " " , Unpaid family workers < 10%.
S-10: " " " " , Farm level living = 85-124.
S-11: Wholesale and retail trade employment = 20%+.
S-12: Nonwhite < 15%.

presented in Table 16. Although every control group has an interregional variance which exceeds that for the total group of metropolitan areas, the larger values are in no case of a magnitude sufficient to effect more than a modest change in rho, even when this greater interregional heterogeneity is accompanied by greater intraregional homogeneity (see control groups S-1 through S-6).

In order to see more specifically what is happening among the several regions when the various controls are applied, it is necessary to examine data on each region's fertility means and standard deviations. These data are found in Table 17 and Table 18, respectively.

Generally speaking, the data on regional means in Table 17 show remarkably little change under controlled conditions. This is particularly true of the two Eastern regions, the Middle States, and the Far West. The Northwest is also fairly consistent except when controls are applied to nonfarm occupations and unpaid, family farm workers (see S-2, S-3, S-8, and S-9). The Southwest is somewhat erratic, although even here the changes are not dramatically great.

There are, however, a few items of interest in Table 17 which may be pointed out. The control for wholesale and retail trade employment (S-11) has a quite noticeable, depressant effect on fertility in the two Eastern regions and in the Middle States. Furthermore, this is not due to one or two highly deviant metropolitan areas, because the standard deviations shown in Table 18 for these regions are very small when this control (S-11) is applied. It may be noted here that it is this large reduction in the internal variation of these three regions which accounts for

TABLE 17

DISTRIBUTIONS OF REGIONAL MEANS FOR FERTILITY BY THE TOTAL GROUP OF METROPOLITAN STATE ECONOMIC AREAS AND TWELVE SELECTED CONTROL GROUPS

Regions	Total Metrop. Areas	Control Groups*											
		S-1	S-2	S-3	S-4	S-5	S-6	S-7	S-8	S-9	S-10	S-11	S-12
Northeast	869	896	886	888	883	894	943	906	906	887	878	774	875
Southeast	852	825	852	847	821	880	818	834	844	850	815	747	969
Middle States	876	904	902	892	926	926	936	860	886	880	882	794	876
Northwest	817	825	893	893	835	835	825	817	893	864	823	817	817
Southwest	880	894	770	896	842	907	894	880	765	876	831	882	944
Far West.	814	784	793	798	792	792	784	814	816	826	822	794	814

*The specific controls used are as follow:

S-1: Urban < 85%, Rural farm < 25%.
S-2: " " " , " " " " , Nonfarm occupation = 35%+.
S-3: " " " , " " " " , Unpaid family workers < 10%.
S-4: " " " , " " " " , Farm level of living = 85-124.
S-5: " " " , " " " " , " " " " " " , Nonwhite < 15%.
S-6: " " " , " " " " , Manufacturing employment < 35%.

S-7: Manufacturing employment < 35%.
S-8: " " " " , Nonfarm occupation = 35%+.
S-9: " " " " , Unpaid family workers < 10%.
S-10: " " " " , Farm level living = 85-124.
S-11: Wholesale and retail trade employment = 20%+.
S-12: Nonwhite < 15%.

TABLE 18

DISTRIBUTIONS OF REGIONAL STANDARD DEVIATIONS FOR FERTILITY BY THE TOTAL GROUP OF METROPOLITAN STATE ECONOMIC AREAS AND TWELVE SELECTED CONTROL GROUPS

Regions	Total Metrop. Areas	Control Groups*											
		S-1	S-2	S-3	S-4	S-5	S-6	S-7	S-8	S-9	S-10	S-11	S-12
Northeast	154	156	158	150	122	119	200	199	199	198	147	97	155
Southeast	154	118	121	129	112	143	120	138	135	135	125	69	195
Middle States	152	109	116	107	122	123	105	176	231	193	223	102	152
Northwest	123	135	127	127	148	148	135	123	127	120	135	123	123
Southwest	172	174	55	205	145	207	174	172	52	200	140	182	205
Far West.	127	107	136	106	111	111	107	128	136	125	129	135	128
Total U.S.A..	153	140	139	139	131	135	153	164	171	169	154	127	159

*The specific controls used are as follow:

S-1: Urban < 85%, Rural farm < 25%.
S-2: " " " , " " " " , Nonfarm occupation = 35%+.
S-3: " " " , " " " " , Unpaid family workers < 10%.
S-4: " " " , " " " " , Farm level of living = 85-124.
S-5: " " " , " " " " , " " " " " " , Nonwhite < 15%.
S-6: " " " , " " " " , Manufacturing employment < 35%.

S-7: Manufacturing employment < 35%.
S-8: " " " " , Nonfarm occupation = 35%+.
S-9: " " " " , Unpaid family workers < 10%.
S-10: " " " " , Farm level living = 85-124.
S-11: Wholesale and retail trade employment = 20%+.
S-12: Nonwhite < 15%.

the fact that this control group exhibits in Table 16 the smallest total intraregional variance observed for any of the twelve control groups.

Another item of interest in Table 17 is the performance of the two Southern regions when controls are applied to the nonwhite variable (S-12). The highest mean fertility values in the entire table appear in the Southeast and Southwest regions when those metropolitan areas with less than 15 per cent of their populations classified as nonwhite are examined. It should be remembered here that the fertility measure being used is the native white net reproduction rate. The question, however, can be asked: Why should one find a high white fertility pattern in those Southern areas which have few Negroes?

Fortunately, the number of cases is small enough to permit a rather intensive examination of the data. It can be reported that these high mean fertility values in the Southeast and Southwest are due to a few highly deviant areas which have very large net reproduction rates. That these are "highly deviant" areas is testified to by the fact that in Table 18 these Southern regions have by far the largest standard deviations found for this control group (S-12). Further examination also reveals that these deviant, high fertility areas also are characterized by low rural farm level of living indexes. This seems to suggest that in the Southern regions there are "pockets" of poor white settlement around certain metropolitan centers which exhibit the high fertility pattern that has traditionally been regarded as typical of poor, rural folk. One can only conjecture about why such areas should be conspicuous exceptions to the pervasiveness of Negro settlement in the South.

It may be that such marginal areas do not attract Negroes; or it may be that Negroes are actively rejected by a population which probably views them as threatening competition for meagre resources. More than likely both factors are operative in this situation.

Several specific references have already been made to the data on regional standard deviations in Table 18. With respect to the table as a whole, however, some additional general remarks are perhaps in order. The hypothesis under test here prompts the expectation that controls for differences in social structure will result in a greater similarity of behavior both within and among regions. This means that the regional standard deviation as a measure of behavioral differences within a region should be reduced when controls are exercised. An inventory of the comparisons between the standard deviations of each control group and those of the total uncontrolled metropolitan grouping in Table 18 reveals only partial support for the hypothesis. Considering all six regions and all twelve control groups, the accounting of the 72 comparisons shows only 47 per cent of them reduced under controls, and some of these reductions are quite slight. In other words, in more than one-half of the cases the controls applied here fail to perform according to hypothetical expectations.

It should be noted, however, that this general accounting procedure is deceptive. That is to say, it does not really provide an accurate over-all picture of the results of these controls because in some instances the application of controls does not alter the State economic area composition of the region. Thus, for example, controlling for less than 35 per cent manufacturing

employment leaves the three Western regions completely unchanged. None of these regions has a metropolitan area with as much as 35 per cent of its labor force employed in manufacturing.[2] Revising the inventory procedure accordingly, the results show 52 per cent of the comparisons taking the form of reductions in the control groups' standard deviations. While this is not greatly different from the other inventory figure, it does represent a more accurate picture of the effects of controls.

A more detailed examination of the data in Table 18 reveals that the greatest frequency of reduction occurs in the two Eastern regions, the Middle States, and the Far West. In these four regions 64 per cent of the comparisons take the form of lower sigma values for the control groups. It is also interesting to observe that these four regions seem most susceptible to reductions in sigma value when the urban-rural farm base control is employed (see S-1 through S-6). For these six control groups 83 per cent of the comparisons in these four regions exhibit reductions in sigma value.

The question of why the Southwest and the Northwest in particular should fail to respond to controls in the same way the other regions do is puzzling. Actually, the Northwest's failure to exhibit sigma reductions is really a failure to respond much at all to the application of controls. Ten out of the twelve control groups for this region have sigma values that vary no more than ten per cent from the standard deviation for the entire group of the region's metropolitan areas. The Northwest is also a region whose

---

[2]See Table 14 for changes in the unit composition of each region when controls are applied.

metropolitan areas representation is very limited-namely, six areas in all. This does not provide much opportunity for variation.

The Southwest's deviant position is more difficult to comprehend. There is considerable fluctuation in the sigma values of the control groups here. There is also, however, some consistency in that controls for nonfarm occupation (S-2 and S-8) and controls for rural farm level of living (S-4 and S-10) both effect reductions in sigma values, while the controls for unpaid farm family labor (S-3 and S-9) show larger sigma values. It would seem from this that a low proportion of unpaid farm family labor in the Southwest is not a good index of rural conditions which, by minimizing marginal subsistence farming, throw most of the rural farm labor force into a commercial nexus and ipso facto constitute a control on rural farm level of living. Perhaps there are some unique peculiarities about the rural farm population ranged about the fringes of Southwestern metropolitan centers.

To summarize briefly, the analysis of regional differences in metropolitan fertility behavior under controlled conditions shows an insignificant relationship between fertility and regionalism. Although there is some fluctuation in the values of the interregional and intraregional components of variance when controls are applied, these changes are relatively small and simply serve to reinforce the conclusion drawn from an analysis of the total group of metropolitan areas-namely, that the fertility behavior of metropolitan area, native white residents is essentially the same regardless of differences in regional location.

With respect to the little change that does occur when

various controls are exercised, it appears that the urban-rural farm base controls are the most successful in reducing the internal variation of the individual regions. This is particularly true of the two Eastern regions, the Middle States, and the Far West. The Northwest does not conform here, but this appears due to the fact that none of the controls seems to alter very much the rather high degree of internal homogeneity of this region. The Southwest is rather susceptible to controls on nonfarm occupation, rural farm level of living, and unpaid, farm family workers. The first two controls reduce the internal variation of this region, while the third control increases the regional sigma value. This may be due to some peculiarities of the rural farm population residing on the fringes of the metropolitan areas in the Southwest.

Generally speaking, the attempt to reduce differences in metropolitan fertility behavior by controlling differences in metropolitan social structure meets with only partial success. Some of this failure may perhaps be attributed to the selection of inappropriate controls or to a lack of precision in applying the controls. This latter factor may be particularly crucial in view of the fact that the total group of metropolitan areas exhibits such small differences that very little opportunity remains for further reduction.

Proportion of Eligible Voters Voting

It will be recalled from the findings reported in Chapter III that this variable on voter turnout has the highest intraclass correlation coefficient attained by any of our five dependent variables—namely, .774. The general performance of this variable under

controlled conditions is shown in Table 19. It may be observed from the data on the values of rho that for all control groups save one (S-1) there is a reduction in the amount of variation associated with regional location. It may also be observed that most of these reductions are rather small. Indeed, one-half of the control groups have reductions of less than ten per cent of the value of the total uncontrolled rho.

While the magnitudes of these reductions may not be particularly impressive, the consistency is. In order to see more specifically what is bringing about these changes in the value of the intraclass correlation, attention is directed to the data on the interregional and intraregional components of variance in Table 19. Here it can be seen that for eight out of the twelve control groups there is a reduction in the value of the total, uncontrolled, interregional variance. It can also be seen that with but three exceptions (S-5, S-11, and S-12) the control groups exhibit larger intraregional variance values than does the total group of metropolitan areas. Since the value of rho varies directly with the magnitude of the intraregional variance and inversely with the magnitude of the interregional variance, it is apparent that the almost perfect consistency of the reduction in rho values for these control groups is due in good part to the greater internal heterogeneity of the regions. That is to say, for the four control groups which do <u>not</u> exhibit a reduction in their interregional variance components (S-1, S-6, S-7, and S-8) we should expect to find a larger interclass correlation. The fact that this does not occur can be explained by the performance of these control groups' intraregional variance values which become larger.

TABLE 19

A LISTING FOR PROPORTION OF ELIGIBLE VOTERS VOTING OF THE VALUES OF RHO, THE INTERREGIONAL VARIANCE, THE INTRAREGIONAL VARIANCE, AND N FOR THE TOTAL GROUP OF METROPOLITAN STATE ECONOMIC AREAS AND TWELVE SELECTED CONTROL GROUPS

Listed Items	Total Metrop. Areas	Control Groups*											
		S-1	S-2	S-3	S-4	S-5	S-6	S-7	S-8	S-9	S-10	S-11	S-12
Rho	.774	.774	.712	.710	.683	.378	.762	.768	.680	.689	.675	.755	.514
Interregional Variance . . .	351.5	371.3	309.0	273.6	240.1	59.5	453.3	432.9	356.3	313.7	274.5	302.5	86.2
Intraregional Variance . . .	97.2	101.3	112.3	102.0	100.1	76.8	127.3	120.8	144.3	124.8	115.1	86.1	73.5
N	149	116	80	93	82	64	77	99	60	70	66	63	114

*The specific controls used are as follow:

S-1: Urban < 85%, Rural farm < 25%.
S-2: " " " , " " " " , Nonfarm occupation = 35%+.
S-3: " " " , " " " " , Unpaid family workers < 10%.
S-4: " " " , " " " " , Farm level of living = 85-124.
S-5: " " " , " " " " , " " " " " " , Nonwhite < 15%.
S-6: " " " , " " " " , Manufacturing employment < 35%.

S-7: Manufacturing employment < 35%.
S-8: " " " " , Nonfarm occupation = 35%+.
S-9: " " " " , Unpaid family workers < 10%.
S-10: " " " " , Farm level living = 85-124.
S-11: Wholesale and retail trade employment = 20%+.
S-12: Nonwhite < 15%.

It is perhaps unnecessary to point out that the most striking changes which take place in the values of rho and the two variance components appear when control is exercised over the nonwhite proportion of the total population (S-5 and S-12). Some additional evidence relating to this independent variable will be presented later on.

A more detailed analysis of the consequences for regional differences in voting behavior when controls are applied is permitted by data on regional means and standard deviations. This material appears in Tables 20 and 21. It is immediately obvious from the data on regional means in Table 20 that it is the two Southern regions which account for the fluctuating values of the interregional variance. The other four regions are affected very little by the application of controls. Indeed, among these four regions no control group exhibits a mean value that differs as much as five per cent from the mean value of its respective total metropolitan area grouping. This consistency is quite remarkable, particularly in view of the rather large fluctuations appearing in the Southern regions.

The question of consistency in voter turnout within regions, however, is more accurately revealed by the data on standard deviations in Table 21. Here it is clearly evident that the Southeast and Southwest are far and away the most inconstant regions. These data also show considerable variation in the Northeast, but the remaining three regions are conspicuous in their relatively high degrees of homogeneity. These remarks pertain to all the control groups without noticeable exception.

With respect to the specific regions and control groups

TABLE 20

DISTRIBUTIONS OF REGIONAL MEANS FOR PROPORTION OF ELIGIBLE VOTERS VOTING BY THE TOTAL GROUP OF METROPOLITAN STATE ECONOMIC AREAS AND TWELVE SELECTED CONTROL GROUPS

Regions	Total Metrop. Areas	Control Groups*											
		S-1	S-2	S-3	S-4	S-5	S-6	S-7	S-8	S-9	S-10	S-11	S-12
Northeast	67.5%	67.7%	67.7%	67.2%	66.2%	66.3%	69.1%	68.6%	68.6%	67.7%	66.6%	69.5%	67.6%
Southeast	26.6	27.1	30.3	31.4	34.2	46.1	26.2	26.2	30.7	31.8	34.6	32.3	41.0
Middle States	70.8	71.5	69.7	71.2	70.8	70.8	73.3	73.0	72.8	73.5	73.6	70.5	70.8
Northwest	72.8	73.0	76.0	76.0	73.8	73.8	73.0	72.8	76.0	75.0	73.4	72.8	72.8
Southwest	34.5	32.4	33.2	34.2	34.5	46.6	32.4	34.5	37.0	37.0	36.9	35.3	40.7
Far West.	64.5	65.5	65.4	65.5	64.9	64.9	65.5	64.5	65.5	64.4	64.0	65.6	64.5

*The specific controls used are as follow:

S-1: Urban < 85%, Rural farm < 25%.
S-2: " " " , " " " " , Nonfarm occupation = 35%+.
S-3: " " " , " " " " , Unpaid family workers < 10%.
S-4: " " " , " " " " , Farm level of living = 85-124.
S-5: " " " , " " " " , " " " " " " , Nonwhite < 15%.
S-6: " " " , " " " " , Manufacturing employment < 35%.

S-7: Manufacturing employment < 35%.
S-8: " " " " , Nonfarm occupation = 35%+.
S-9: " " " " , Unpaid family workers < 10%.
S-10: " " " " , Farm level living = 85-124.
S-11: Wholesale and retail trade employment = 20%+.
S-12: Nonwhite < 15%.

TABLE 21

DISTRIBUTIONS OF REGIONAL STANDARD DEVIATIONS FOR PROPORTION OF ELIGIBLE VOTERS VOTING BY THE TOTAL GROUP OF METROPOLITAN STATE ECONOMIC AREAS AND TWELVE SELECTED CONTROL GROUPS

Regions	Total Metrop. Areas	Control Groups*											
		S-1	S-2	S-3	S-4	S-5	S-6	S-7	S-8	S-9	S-10	S-11	S-12
Northeast	8.3%	8.8%	9.1%	8.9%	8.8%	8.5%	11.5%	10.6%	10.6%	10.6%	10.1%	7.5%	8.2%
Southeast	15.0	15.2	16.0	15.6	15.9	18.8	15.5	15.3	16.5	16.5	15.8	14.6	19.0
Middle States	5.9	5.8	5.2	5.9	6.0	6.0	4.5	4.7	3.7	4.7	4.2	4.3	5.9
Northwest	4.5	4.9	4.0	4.0	5.2	5.2	4.9	4.5	4.0	3.4	4.7	4.5	4.5
Southwest	13.2	11.8	12.7	13.8	12.3	12.8	11.8	13.2	14.3	14.6	13.4	13.5	15.1
Far West.	4.1	2.5	1.6	2.2	2.0	2.0	2.5	4.1	2.1	4.1	4.0	2.5	4.1
Total U.S.A.. . . .	21.2%	21.8%	20.6%	19.3%	18.5%	11.8%	24.1%	23.5%	22.4%	20.9%	19.8%	19.7%	12.6%

*The specific controls used are as follow:

S-1: Urban < 85%, Rural farm < 25%.
S-2: " " " , " " " " , Nonfarm occupation = 35%+.
S-3: " " " , " " " " , Unpaid family workers < 10%.
S-4: " " " , " " " " , Farm level of living = 85-124.
S-5: " " " , " " " " , " " " " " " , Nonwhite < 15%.
S-6: " " " , " " " " , Manufacturing employment < 35%.

S-7: Manufacturing employment < 35%.
S-8: " " " " , Nonfarm occupation = 35%+.
S-9: " " " " , Unpaid family workers < 10%.
S-10: " " " " , Farm level living = 85-124.
S-11: Wholesale and retail trade employment = 20%+.
S-12: Nonwhite < 15%.

which exhibit reductions in their degrees of internal heterogeneity, the data in Table 21 reveal that 45 per cent of the comparisons between control groups and their respective total metropolitan groupings show reductions in sigma value under controlled conditions. It can be seen, however, that the regions differ considerably in their susceptibility to reductions when controls are applied. The two Eastern regions tend not to respond to controls in this way. However, the Middle States and the three Western regions do, with 63 per cent of the comparisons in these four regions showing smaller sigma values for the control groups.

There are differences among the regions with respect to which of the control groups are effective in reducing internal variation. The Middle States are most susceptible to the base controls on manufacturing employment, while the Southwest and Far West respond more to the urban-rural farm base controls. It is also interesting to note that the control group for wholesale and retail trade employment tends to have a reductive effect even in the Northeast and Southeast. It will be recalled that it is also this control which effects a dramatic reduction in the sigma values for fertility in the two Eastern regions.

Turning to the controls on the nonwhite proportion of the total population (S-5 and S-12), it is clearly evident that we have here the most sensitive independent variable in regard to voter turnout. Referring to Table 19, the rho value of .514 when only the proportion nonwhite is controlled (S-12) represents a 34 per cent reduction from the total, uncontrolled, metropolitan rho of .774. When this nonwhite control is applied in conjunction with controls for urban-rural farm and farm level of living (S-5), the

resulting rho of .378 represents a 51 per cent reduction. The reason for these great changes can be seen in the performances of the interregional and intraregional variance components, both of which undergo rather startling reductions.

The reduction in the interregional variance is, of course, due in large measure to the larger mean values in the two Southern regions (see Table 20, S-5 and S-12). This indicates that in those areas of the South where there are few Negroes, voter turnout is relatively high. On the other hand, the reduction in the intraregional variance for these two nonwhite control groups is deceptive because it is due entirely to a statistical artifact which involves the control procedure and the character of rho.[3]

The dramatic changes which occur when the nonwhite control is applied prompted some further investigation of this variable. Since it is the Southeast which is most sensitive to this control,

---

[3]The difficulty here can be clearly seen when the standard deviations in Table 21 are examined (see S-5 and S-12). For example, where the nonwhite control alone is applied (S-12), only one region (the Northeast) shows a reduction in sigma value, and this is negligible. The sigma values for all the other regions either do not change or else change in the direction of a larger standard deviation for the control group. The obvious question is how to reconcile these facts about each individual region with the fact that all regions taken together exhibit a noticeable reduction in intraregional variance. The answer to this apparent anomaly lies in the character of the intraregional variance for all regions. This value is a weighted average of the internal variations for each region. Consequently, when the nonwhite control is applied, the case loss in the two Southern regions is quite large (see Table 14, S-5 and S-12). Since these are also the regions with large amounts of internal variation (see Table 21) it is clearly apparent that the weighted, average, intraregional variance for all regions will be reduced. It might also be noted here that this weighting factor also influences the value of the interregional variance which is simply a weighted average of the individual mean squared deviations for each region. It is because of these considerations that the analysis in this study cannot stop with a simple examination of the values of rho and its two variance components.

the investigation was focused on this region. The decision was made to classify the 35 metropolitan areas in this region on the basis of their nonwhite population and to compute an intraclass correlation for this classification. The results of this are presented in Table 22.

TABLE 22

MEAN VALUES AND STANDARD DEVIATIONS FOR PROPORTION OF ELIGIBLE VOTERS VOTING IN THE METROPOLITAN STATE ECONOMIC AREAS OF THE SOUTHEAST REGION BY THE PROPORTION OF THE TOTAL POPULATION WHICH IS NONWHITE

Proportion Nonwhite	Mean	Standard Deviation	N
40%+	15.2%	5.3%	10
30-39.	23.2	10.4	7
20-29.	26.2	9.2	8
10-19.	39.7	13.8	6
Under 10	42.8	19.1	4
Total Region.	26.6%	15.0%	35

Although the total N is small, it seems rather clear that among metropolitan areas in the Southeast voter turnout is inversely related to the proportion of the total population which is nonwhite. Furthermore, except for one slight reversal the standard deviation for these class means also varies inversely with the proportion of the total population which is nonwhite. The intraclass correlation for this distribution of values is .354.

The rho value of .354 may be interpreted as a measure of the extent to which a person, knowing the nonwhite proportion of

the population, could improve his prediction of voter turnout among metropolitan areas. Stated in other words, if a person wished to predict voter turnout for the individual metropolitan areas in the Southeast, his best prediction would be the universe value of 26.6 per cent. On the average this would give him his least margin of error. However, if the person had one more item of information--namely, the nonwhite proportion of the population in each metropolitan area--he could on the average improve the accuracy of his predictions of voter turnout by 35 per cent. In any scientific field an improvement of this magnitude in predictive accuracy is not to be dismissed lightly.

To summarize the results of this analysis of voter turnout, it can be said that all but one of the applied controls succeed in reducing the degree of association between this dependent variable and regionalism. Most of these reductions, however, are rather small. The one independent variable which is clearly the most influential as far as voter turnout is concerned is that of the nonwhite proportion of the total population. Understandably, the influence of this variable is most acutely felt in the two Southern regions where voter turnout is inversely related to the nonwhite proportion of the total population. Accordingly, when controls are applied so that only those metropolitan areas with small proportions of nonwhite residents are considered, regional differences in voter turnout are greatly reduced.

Except for the two Southern regions, none of the other regions seems to be much affected in its voter turnout by the application of these controls. The Middle States tend to become a bit more homogeneous internally when manufacturing base controls are

applied, while the Far West tends to respond this way when the urban-rural farm base controls are applied. The Northeast and Southeast regions exhibit very little reduction in their measures of internal heterogeneity under controlled conditions. The only real exception to this appears with the control on wholesale and retail trade employment. It will be recalled that this control is also effective in reducing the internal, regional variations in fertility.

Proportion of Voters Voting Democratic

The correlation between Democratic voting in all metropolitan areas and regional location is fairly high as shown by the rho value of .589, which was reported in Chapter III. The influence of controls on this correlation is revealed by the values of rho and its two variance components which are presented in Table 23. It can be seen at a glance that the intraclass correlations for every control group fall below the correlation value for the total group of metropolitan areas. To be sure, a few of these reductions are rather small, but fully three-quarters of them are greater than 10 per cent, while one-half of them exceed a 20 per cent reduction. And, of course, when the nonwhite control is applied (S-5 and S-12), the resulting rho values represent reductions in the neighborhood of 75 per cent or greater. Indeed, the rho value of .118 for the multiple control group which includes the nonwhite control (S-5) falls below the value necessary for statistical significance at the five per cent level (see Table 15, S-5).

The immediate reason for these changes is seen by an examination of the interregional and intraregional components of

TABLE 23

A LISTING FOR PROPORTION OF VOTERS VOTING DEMOCRATIC OF THE VALUES OF RHO, THE INTERREGIONAL VARIANCE, THE INTRAREGIONAL VARIANCE, AND N FOR THE TOTAL GROUP OF METROPOLITAN STATE ECONOMIC AREAS AND TWELVE SELECTED CONTROL GROUPS

Listed Items	Total Metrop. Areas	Control Groups*											
		S-1	S-2	S-3	S-4	S-5	S-6	S-7	S-8	S-9	S-10	S-11	S-12
Rho	.589	.550	.454	.502	.472	.118	.506	.564	.460	.445	.447	.548	.148
Interregional Variance . . .	117.1	107.8	72.2	65.8	73.0	13.0	115.5	134.0	86.2	67.6	76.0	101.7	11.8
Intraregional Variance . . .	76.3	80.1	73.6	65.3	69.9	51.8	96.8	93.2	81.5	69.6	76.8	70.5	49.4
N	149	116	80	93	82	64	77	99	60	70	66	63	114

*The specific controls used are as follow:

S-1: Urban < 85%, Rural farm < 25%.
S-2: " " " , " " " " , Nonfarm occupation = 35%+.
S-3: " " " , " " " " , Unpaid family workers < 10%.
S-4: " " " , " " " " , Farm level of living = 85-124.
S-5: " " " , " " " " , " " " " " " , Nonwhite < 15%.
S-6: " " " , " " " " , Manufacturing employment < 35%.

S-7: Manufacturing employment < 35%.
S-8: " " " " , Nonfarm occupation = 35%+.
S-9: " " " " , Unpaid family workers < 10%.
S-10: " " " " , Farm level living = 85-124.
S-11: Wholesale and retail trade employment = 20%+.
S-12: Nonwhite < 15%.

variance also presented in Table 23. For every control group except one (S-7) the value of the interregional variance falls below that for the total group of metropolitan areas. With respect to the intraregional variance, seven out of the twelve control groups show smaller amounts of internal variation for all regions. All but one of the five control groups which show larger amounts of internal variation involve a control on manufacturing employment.

Turning to the data on regional means in Table 24, we may investigate further the reasons for the shifts in the values of the interregional variance under controlled conditions. Generally speaking, it is only the two Southern regions which exhibit much variation in regional mean values when controls are applied. This, of course, is the same reaction pattern observed for the voter turnout variable. The one control group (S-7) which has a larger interregional variance than the total group of metropolitan areas shows a negligible increase in the mean value for the Southeast and equally negligible reductions in mean values for the Northeast and Middle States. Since the other regions' mean values do not change, the fact that this control group has a 12 per cent larger interregional variance than that for the total metropolitan group requires some explanation. This is afforded by reference to Table 14, where it can be seen that the distribution of metropolitan State economic areas among the various regions in this control group (S-7) places a much greater weighting on the two Southern regions than is the case for the distribution of all metropolitan areas. In fact, all the control groups involving controls for manufacturing employment tend to overweight the representation of the two Southern regions. This means that some care must be taken

TABLE 24

DISTRIBUTIONS OF REGIONAL MEANS FOR PROPORTION OF VOTERS VOTING DEMOCRATIC BY THE TOTAL GROUP OF METROPOLITAN STATE ECONOMIC AREAS AND TWELVE SELECTED CONTROL GROUPS

Regions	Total Metrop. Areas	Control Groups*											
		S-1	S-2	S-3	S-4	S-5	S-6	S-7	S-8	S-9	S-10	S-11	S-12
Northeast	52.8%	52.4%	52.8%	52.1%	51.4%	51.0%	52.8%	52.7%	52.7%	52.0%	51.9%	48.3%	52.6%
Southeast	78.9	76.4	72.2	71.4	71.1	58.8	76.7	79.1	72.9	71.3	72.8	76.6	61.9
Middle States	54.7	54.5	54.4	54.0	56.0	56.0	55.2	54.0	52.5	52.8	56.5	54.2	54.7
Northwest	53.0	52.5	56.7	56.7	52.9	52.9	52.5	53.0	56.7	56.4	53.4	53.0	53.0
Southwest	72.4	73.7	72.1	71.0	74.1	62.7	73.7	72.4	70.0	69.6	72.5	71.7	65.3
Far West.	58.4	57.9	58.1	57.8	58.0	58.0	57.9	58.4	58.3	58.4	58.5	58.4	58.4

*The specific controls used are as follow:

S-1: Urban < 85%, Rural farm < 25%.
S-2: " " " , " " " " , Nonfarm occupation = 35%+.
S-3: " " " , " " " " , Unpaid family workers < 10%.
S-4: " " " , " " " " , Farm level of living = 85-124.
S-5: " " " , " " " " , " " " " " " , Nonwhite < 15%.
S-6: " " " , " " " " , Manufacturing employment < 35%.

S-7: Manufacturing employment < 35%.
S-8: " " " " , Nonfarm occupation = 35%+.
S-9: " " " " , Unpaid family workers < 10%.
S-10: " " " " , Farm level living = 85-124.
S-11: Wholesale and retail trade employment = 20%+.
S-12: Nonwhite < 15%.

in evaluating differences between any summary measures which are not based on the same distribution of units. In Appendix A, Table 13 presents percentage distributions of metropolitan State economic areas among the six regions for each control group and for the total metropolitan area grouping.

Unquestionably the most striking feature in Table 24 is the convergence of regional means when control is exercised over the nonwhite proportion of the population (S-5 and S-12). This, of course, is due to the dramatic reduction in Democratic voting in the Southern regions when only those areas with relatively few Negroes are considered. Additional data pertaining to this phenomenon will be presented later.

Although the Southwest does not exhibit much variation in its mean values except under nonwhite control conditions, the Southeast appears somewhat responsive to controls on nonfarm occupation, unpaid farm family workers, and farm level of living (see S-2, S-3, S-4 and S-8, S-9, S-10). The mean values for these control groups are all somewhere between eight and ten per cent lower than the mean for all metropolitan areas in the Southeast. These results correspond roughly to those appearing for voter turnout, the main difference being the direction of the change. That is to say, when these three controls are applied, the mean values for voter turnout in the Southeast become larger than the mean for all metropolitan areas in this region.

With respect to the data on regional standard deviations in Table 25, a general comparison of control group values with those for the total metropolitan area grouping reveals that 52 per cent of the control groups have smaller sigma values. These

TABLE 25

DISTRIBUTIONS OF REGIONAL STANDARD DEVIATIONS FOR PROPORTION OF VOTERS VOTING DEMOCRATIC BY THE TOTAL GROUP OF METROPOLITAN STATE ECONOMIC AREAS AND TWELVE SELECTED CONTROL GROUPS

Regions	Total Metrop. Areas	Control Groups*											
		S-1	S-2	S-3	S-4	S-5	S-6	S-7	S-8	S-9	S-10	S-11	S-12
Northeast	6.4%	6.7%	6.6%	6.8%	7.2%	6.8%	6.4%	5.8%	5.8%	6.2%	5.9%	4.9%	6.3%
Southeast	12.4	12.3	11.0	10.5	9.6	2.8	12.5	12.8	11.2	10.6	10.5	10.2	9.3
Middle States	6.6	7.1	7.8	6.7	6.9	6.9	7.7	6.9	8.0	6.4	6.2	6.3	6.6
Northwest	7.2	7.7	6.2	6.2	8.4	8.4	7.7	7.2	6.2	5.5	7.8	7.2	7.2
Southwest	12.4	12.2	13.7	12.6	13.5	15.5	12.2	12.4	13.4	12.4	13.6	12.6	12.3
Far West.	5.7	5.4	3.8	5.4	5.5	5.5	5.4	5.7	3.0	5.6	5.8	5.6	5.7
Total U.S.A.. . . .	13.9%	13.7%	12.0%	11.4%	12.0%	8.2%	14.5%	15.0%	12.9%	11.7%	12.3%	13.1%	7.8%

*The specific controls used are as follow:

S-1: Urban < 85%, Rural farm < 25%.
S-2: " " " , " " " " , Nonfarm occupation = 35%+.
S-3: " " " , " " " " , Unpaid family workers < 10%.
S-4: " " " , " " " " , Farm level of living = 85-124.
S-5: " " " , " " " " , " " " " " " , Nonwhite < 15%.
S-6: " " " , " " " " , Manufacturing employment < 35%.

S-7: Manufacturing employment < 35%.
S-8: " " " " , Nonfarm occupation = 35%+.
S-9: " " " " , Unpaid family workers < 10%.
S-10: " " " " , Farm level living = 85-124.
S-11: Wholesale and retail trade employment = 20%+.
S-12: Nonwhite < 15%.

reductions are concentrated in the two Eastern regions and the Far West. They also tend to occur when additional controls are applied to the manufacturing base control (see S-8, S-9, and S-10). Indeed, considering all regions, two-thirds of the comparisons in these three control groups show reductions in sigma value. It may also be observed that the control for wholesale and retail trade employment reduces the standard deviations exhibited by the two Eastern regions and the Middle States. It will be recalled that this occurs also for the fertility and voter turnout variables.

By far the most striking change in sigma value under controlled conditions occurs in the Southeast when the nonwhite control is added to the urban-rural farm and farm level of living controls (see S-5). This standard deviation of 2.8 per cent is unbelievably small, but a check on the original data reveals it to be an accurate measure. This datum, together with the other findings about Democratic voting in the Southeast, prompted an investigation of the relationship between the nonwhite proportion of the population and the political party preference of Southeastern voters. This investigation parallels that reported earlier for voter turnout.

The metropolitan areas in the Southeast were classified into five groups based on the proportion of the population nonwhite. The mean values and standard deviations for each group were computed, the results of which appear in Table 26. These data clearly show that Democratic Party preference among metropolitan area voters in this region is directly related to the proportion of the total population which is nonwhite. The standard deviations (with the exception of the "Under 10" group) are rather

small, which lends confidence to the reported mean values. The intraclass correlation for this distribution is .662. As explained before, this can be interpreted to mean that knowledge of the nonwhite proportion of the population can improve by 66 per cent one's prediction of Democratic voting (from universe values) in metropolitan areas of the Southeast. Putting these results together with those from the investigation of voter turnout, it appears that among metropolitan areas in the Southeast the more Negroes there are, the less likely the population will vote, but the more likely it is that those who do vote will vote Democratic.

TABLE 26

MEAN VALUES AND STANDARD DEVIATIONS FOR PROPORTION OF VOTERS VOTING DEMOCRATIC IN THE METROPOLITAN STATE ECONOMIC AREAS OF THE SOUTHEAST REGION BY THE PROPORTION OF THE TOTAL POPULATION WHICH IS NONWHITE

Proportion Nonwhite	Mean	Standard Deviation	N
40%+	91.6%	5.5%	10
30-39.	83.0	4.9	7
20-29.	77.4	6.4	8
10-19.	65.6	5.4	6
Under 10	62.8	12.1	4
Total Region.	78.9%	12.4%	35

It should be pointed out here that the findings for voter turnout are to be expected in view of the virtual disenfranchisement of Negroes in the Southeast region. It would be interesting, however, to investigate the extent to which this pattern seems to

hold up when only the white electorate is considered. Unfortunately, such an investigation would be hampered by the lack of voting data classified by race of voter.

To summarize, the application of controls serves in every case to reduce at least somewhat the relationship between Democratic voting in metropolitan areas and the regional location of these areas. The greatest reduction occurs when control for the nonwhite proportion of the total population is exercised. This can be attributed almost entirely to the changes in the proportion of Democratic votes cast in the two Southern regions—particularly the Southeast. In general, it appears that among metropolitan areas in the Southeast the greater the proportion of Negroes in the population, the greater is the proportion of voters who vote Democratic.

With respect to other controls there is a tendency for a convergence of values both within and among regions when controls on nonfarm occupation, unpaid farm family workers, and rural farm level of living are exercised. The Southeast in particular seems sensitive to controls, nearly all of which reduce the internal variation of that region. The control for wholesale and retail trade employment effects a noticeable reduction in the internal variation of the two Eastern regions—a pattern which also appears for the fertility and voter turnout variables.

Educational Achievement

Summary data showing the effects of controls on the relationship between the educational achievement of metropolitan area residents and their regional locations are presented in Table 27.

TABLE 27

A LISTING FOR EDUCATION OF THE VALUES OF RHO, THE INTERREGIONAL VARIANCE, THE INTRAREGIONAL VARIANCE, AND N FOR THE TOTAL GROUP OF METROPOLITAN STATE ECONOMIC AREAS AND TWELVE SELECTED CONTROL GROUPS

Listed Items	Total Metrop. Areas	Control Groups*											
		S-1	S-2	S-3	S-4	S-5	S-6	S-7	S-8	S-9	S-10	S-11	S-12
Rho	.194	.269	.143	.226	.100	.104	.220	.172	.148	.116	.033	.195	.219
Interregional Variance . . .	1.0	1.0	.5	.7	.4	.4	1.1	1.1	.7	.7	.4	1.0	.9
Intraregional Variance . . .	3.3	2.3	1.8	1.9	2.0	1.9	2.8	4.0	2.2	2.9	2.9	2.6	2.6
N	149	116	80	93	82	64	77	99	60	70	66	63	114

*The specific controls used are as follow:

S-1: Urban < 85%, Rural farm < 25%.
S-2: " " " , " " " " , Nonfarm occupation = 35%+.
S-3: " " " , " " " " , Unpaid family workers < 10%.
S-4: " " " , " " " " , Farm level of living = 85-124.
S-5: " " " , " " " " , " " " " " " , Nonwhite < 15%.
S-6: " " " , " " " " , Manufacturing employment < 35%.

S-7: Manufacturing employment < 35%.
S-8: " " " " , Nonfarm occupation = 35%+.
S-9: " " " " , Unpaid family workers < 10%.
S-10: " " " " , Farm level living = 85-124.
S-11: Wholesale and retail trade employment = 20%+.
S-12: Nonwhite < 15%.

The intraclass correlation coefficients show reductions in this relationship for seven control groups. Furthermore, five of these groups (S-4, S-5, S-8, S-9, and S-10) exhibit changes of such magnitude that the value of rho falls below the five per cent level of statistical significance (see Table 15).

Turning to the variance components for an explanation of these correlation changes, we find nine of the control groups exhibiting an interregional variance that is smaller than the value for all metropolitan areas taken together. Several of these reductions are fairly large--i.e., 50 per cent or more. The fact that nine groups show this convergence of regional means but only seven groups exhibit reductions in rho value of course indicates some rather substantial reductions in the internal variations of the several regions.

This expectation is borne out by an examination of the intraregional component of variance. Here it can be seen that all but one of the control groups (S-7) have an intraregional variance which is smaller than that for the total group of metropolitan areas. Generally speaking, the controls involving the urban-rural farm base control (S-1 through S-4) seem to effect greater variance reductions than those involving the manufacturing base control (S-7 through S-10). This is true of both the interregional and intraregional variance components.

Pushing the analysis further, we find in Table 28 data on the regional means for each control group. It is the performance of these mean values under various controls which accounts for shifts in the interregional variance observed above. The two Southern regions are, of course, the most highly deviant regions

TABLE 28

DISTRIBUTIONS OF REGIONAL MEANS FOR EDUCATION BY THE TOTAL GROUP OF METROPOLITAN STATE ECONOMIC AREAS AND TWELVE SELECTED CONTROL GROUPS

Regions	Total Metrop. Areas	Control Groups*											
		S-1	S-2	S-3	S-4	S-5	S-6	S-7	S-8	S-9	S-10	S-11	S-12
Northeast	3.8%	3.5%	3.5%	3.4%	3.5%	3.5%	3.4%	4.0%	4.0%	3.7%	3.9%	4.0%	3.9%
Southeast	4.4	4.0	3.6	3.4	3.2	2.2	4.0	4.4	3.6	3.4	3.4	3.5	2.7
Middle States	2.0	1.8	2.0	1.7	2.2	2.2	1.8	1.9	2.2	1.6	2.4	1.9	2.0
Northwest	1.4	1.4	1.8	1.8	1.5	1.5	1.4	1.4	1.8	1.8	1.6	1.4	1.4
Southwest	4.4	4.8	3.1	4.3	3.8	2.7	4.8	4.4	2.8	3.9	3.5	4.2	4.5
Far West.	3.0	2.5	2.1	2.6	2.6	2.6	2.5	3.0	1.9	3.2	3.1	2.9	3.0

*The specific controls used are as follow:

S-1: Urban < 85%, Rural farm < 25%.
S-2: " " " , " " " " , Nonfarm occupation = 35%+.
S-3: " " " , " " " " , Unpaid family workers < 10%.
S-4: " " " , " " " " , Farm level of living = 85-124.
S-5: " " " , " " " " , " " " " " " , Nonwhite < 15%.
S-6: " " " , " " " " , Manufacturing employment < 35%.

S-7: Manufacturing employment < 35%.
S-8: " " " " , Nonfarm occupation = 35%+.
S-9: " " " " , Unpaid family workers < 10%.
S-10: " " " " , Farm level living = 85-124.
S-11: Wholesale and retail trade employment = 20%+.
S-12: Nonwhite < 15%.

with respect to large mean values when all metropolitan areas are considered. It is, therefore, to be expected that the control groups showing the greatest convergence of mean values (S-2, S-4, S-5, and S-10) according to the magnitude of the interregional variance (Table 27) will also show rather substantial reductions in the mean values of the Southeast and Southwest. The data in Table 28 show this to be true. Indeed, with respect to the Southeast region alone, the effectiveness of all controls in reducing the regional mean is such that only three control groups (S-1, S-6, and S-7) have mean values which are larger than those in the Northeast. It may also be pointed out that these three control groups are also the only ones having a larger interregional variance than that for the total group of metropolitan areas.

Considering all of the regions, it appears that the Southern regions change the most under controls, the Northeast changes the least, and the other three regions react somewhere in between these extremes. The greatest reduction in mean value for the two Southern regions occurs when the nonwhite control is added to other controls involving the urban-rural farm base control (S-5). The Southeast also exhibits a large reduction in mean value when only the nonwhite control alone is applied (S-12). The Southwest does not react this way to the nonwhite control applied alone. This region does, however, exhibit a noticeably large reduction in mean value whenever the nonfarm occupation control is applied (S-2 and S-8). Although the Northeast does not change appreciably in any direction, the smallest means appear in this region when controls are applied to the urban-rural farm base control (S-2 through S-5). Since this also tends to be true for the other two

"high-mean-regions"--the Southeast and Southwest--it is not surprising that these control groups show rather small interregional variance components in Table 27. In other words, the convergence of regional means for these control groups is a function of reduced values in the Northeast, the Southeast, and the Southwest.

It was stated earlier that eleven out of the twelve control groups have an intraregional variance which is smaller than that for the total group of metropolitan areas. In order to see what is bringing about this rather wholesale reduction in the average internal variation of the six regions, it is necessary to examine the data on regional standard deviations presented in Table 29. It becomes quickly evident in looking at these data that the majority of the controls serve to reduce at least somewhat the internal variation of every region except the Middle States. In this central region only the control on unpaid farm family workers effects any reduction in sigma value, and the amount of reduction is negligible. It should be recalled here that the data in Table 28 also show the Middle States' exhibiting very little change in mean value when controls are applied. In short, it appears that the metropolitan areas in this region have very similar and stable proportions of unschooled adult males in their populations. The picture is somewhat different for the other regions.

A general inventory of sigma value comparisons between all metropolitan areas and the twelve control groups shows that for all regions 68 per cent of the control groups have smaller standard deviations. This proportion increases to 78 per cent when the Middle States region is excluded. Only one control group (S-7) fails to effect some reduction in any region. The values for this

TABLE 29

DISTRIBUTIONS OF REGIONAL STANDARD DEVIATIONS FOR EDUCATION BY THE TOTAL GROUP OF METROPOLITAN STATE ECONOMIC AREAS AND TWELVE SELECTED CONTROL GROUPS

Regions	Total Metrop. Areas	Control Groups*											
		S-1	S-2	S-3	S-4	S-5	S-6	S-7	S-8	S-9	S-10	S-11	S-12
Northeast	1.8%	1.5%	1.4%	1.2%	1.6%	1.6%	1.8%	2.0%	2.0%	1.9%	2.2%	1.4%	1.7%
Southeast	2.1	1.6	1.1	1.2	1.2	.9	1.6	2.2	1.1	1.1	1.3	1.5	1.0
Middle States	1.1	1.1	1.3	1.0	1.2	1.2	1.4	1.2	1.5	1.0	1.3	1.1	1.1
Northwest	.6	.5	.1	.1	.5	.5	.5	.6	.1	.2	.5	.6	.6
Southwest	2.8	2.6	1.8	2.8	1.9	1.3	2.6	2.8	1.7	2.9	2.0	2.8	3.2
Far West.	2.0	1.2	1.1	1.3	1.3	1.3	1.2	2.0	1.1	2.0	2.1	1.4	2.0
Total U.S.A.. . . .	2.0%	1.9%	1.4%	1.6%	1.6%	1.5%	2.0%	2.2%	1.7%	1.9%	1.8%	1.8%	1.9%

*The specific controls used are as follow:

S-1: Urban < 85%, Rural farm < 25%.
S-2: " " " , " " " " , Nonfarm occupation = 35%+.
S-3: " " " , " " " " , Unpaid family workers < 10%.
S-4: " " " , " " " " , Farm level of living = 85-124.
S-5: " " " , " " " " , " " " " " " , Nonwhite < 15%.
S-6: " " " , " " " " , Manufacturing employment < 35%.

S-7: Manufacturing employment < 35%.
S-8: " " " " , Nonfarm occupation = 35%+.
S-9: " " " " , Unpaid family workers < 10%.
S-10: " " " " , Farm level living = 85-124.
S-11: Wholesale and retail trade employment = 20%+.
S-12: Nonwhite < 15%.

remain essentially the same as those for all metropolitan areas.

Although nearly every control group seems to bring about a reduction in sigma value for several regions, it is fairly clear that the control groups involving the urban-rural farm base control (S-1 through S-5) are the most effective. Eighty-three per cent of the sigma values for these five control groups are smaller than those for all metropolitan areas. If the Middle States region is excluded, this proportion increases to 96 per cent, with only the unpaid farm family worker control in the Southwest failing to exhibit a reduced sigma value (S-3).

It is interesting to note that the Southeast has quite small sigma values for the control groups involving a control for nonwhite residents (S-5 and S-12). This is especially interesting in view of the fact that these same two control groups also have rather strikingly small mean values for this region (Table 28). In other words, these small standard deviations lend confidence to what might appear to be suspiciously small mean values. The same thing occurs for controls on nonfarm occupation in the Southwest--i.e., small mean values associated with small standard deviations (see S-2 and S-8 in Tables 28 and 29). It may also be noted in Table 29 that the two Eastern regions show reductions in sigma value when control is exercised for wholesale and retail trade employment (S-11). This pattern has appeared for every dependent variable examined so far.

Summarizing the results of this analysis of educational achievement under controlled conditions, it can be said that structural differences among metropolitan areas appear to be rather significantly related to the proportion of illiterate adult males

found in the metropolitan areas of the several regions. In five of the control groups the correlation between illiteracy among metropolitan populations and the regional location of those populations falls below the five per cent level of statistical significance.

Detailed examination of the data reveals that the regional means for this educational achievement variable converge for nine out of the twelve control groups. This convergence is primarily a function of reductions in mean value for the Northeast and the two Southern regions, all of which have relatively large means for the total group of metropolitan areas. These reductions appear most prominent when multiple controls involving the urban-rural farm base control are applied.

With respect to changes within regions when controls are applied, the data reveal that eleven out of the twelve control groups have smaller, average, intraregional variance components than the total group of metropolitan areas. This reduction in the internal variation of regions is exhibited fairly clearly for every region except the Middle States area which seems to change very little when controls are exercised. As was observed for the convergence of regional means, the control groups figuring most prominently in this reduction of internal variation are those involving the urban-rural farm base control. The control on wholesale and retail trade employment reduces the standard deviation for the Northeast and Southeast. This control group has effected this kind of reduction for each dependent variable examined thus far.

Religious Affiliation

The analysis of the data reported in Chapter III revealed that of the three nonpolitical variables religious affiliation is the most highly correlated with regional location. Evidence was also presented which indicates that most of this high correlation can be attributed to the high rates of affiliation in the metropolitan areas of the Northeast, where disproportionate numbers of Roman Catholics reside. The question addressed in this discussion is whether the application of controls on structural differences will alter the relatively high correlation between religious affiliation and regionalism.

Table 30 presents data for the intraclass correlation and the two major components of variance. The values of rho in this table show seven control groups with correlations lower than that for all metropolitan areas. One of these (S-6) is quite low, representing a reduction of over 50 per cent. Three other control groups (S-7, S-9, and S-11) show reductions exceeding 20 per cent, while only one of the remaining groups (S-3) has a rho value less than 10 per cent below that for all metropolitan areas. It may also be observed here that three out of the four control groups with lowest values of rho involve a control on manufacturing employment.

Some clues as to why these intraclass correlations are altered are provided by the data on the variance components. These show that five of the control groups (S-1, S-6, S-7, S-9, and S-11) have smaller interregional variance values than the total group of metropolitan areas. Some of these reductions are quite substantial (see S-6, S-7, and S-11). This indicates, of course, a rather

TABLE 30

A LISTING FOR RELIGIOUS AFFILIATION OF THE VALUES OF RHO, THE INTERREGIONAL VARIANCE, THE INTRAREGIONAL VARIANCE, AND N FOR THE TOTAL GROUP OF METROPOLITAN STATE ECONOMIC AREAS AND TWELVE SELECTED CONTROL GROUPS

Listed Items	Total Metrop. Areas	Control Groups*											
		S-1	S-2	S-3	S-4	S-5	S-6	S-7	S-8	S-9	S-10	S-11	S-12
Rho	.353	.313	.385	.334	.492	.520	.163	.225	.316	.270	.383	.267	.393
Interregional Variance . . .	49.5	42.0	58.1	50.0	65.1	81.5	25.2	31.9	53.5	45.0	51.0	30.8	65.5
Intraregional Variance . . .	81.2	78.6	76.6	82.0	57.9	62.9	85.4	85.3	85.9	90.5	65.2	60.5	88.7
N	149	116	80	93	82	64	77	99	60	70	66	63	114

*The specific controls used are as follow:

S-1: Urban < 85%, Rural farm < 25%.
S-2: " " " , " " " " , Nonfarm occupation = 35%+.
S-3: " " " , " " " " , Unpaid family workers < 10%.
S-4: " " " , " " " " , Farm level of living = 85-124.
S-5: " " " , " " " " , " " " " " " , Nonwhite < 15%.
S-6: " " " , " " " " , Manufacturing employment < 35%.

S-7: Manufacturing employment < 35%.
S-8: " " " " , Nonfarm occupation = 35%+.
S-9: " " " " , Unpaid family workers < 10%.
S-10: " " " " , Farm level living = 85-124.
S-11: Wholesale and retail trade employment = 20%+.
S-12: Nonwhite < 15%.

marked convergence of regional mean values when these particular controls are exercised.

Turning to the intraregional variance values in Table 30, it can be seen that six of the control groups exhibit values smaller than that for all metropolitan areas (S-1, S-2, S-4, S-5, S-10, and S-11). Most of these involve the urban-rural farm base control. Since four of these six control groups showing greater internal homogeneity also have interregional variances showing greater external heterogeneity, it is not surprising to find relatively large intraclass correlations for them (see S-2, S-4, S-5, and S-10). On the other hand, the relatively small rho values found for control groups S-6, S-7, and S-9 are seen to be the result of combining small interregional variance values with large intraregional values. In other words, when internal homogeneity accompanies external heterogeneity, rho tends to be large (as seen in control groups S-2, S-4, S-5, and S-10). However, when internal heterogeneity accompanies external homogeneity, rho tends to be small (as seen in control groups S-6, S-7, and S-9). These summary, average measures are valuable in pointing out the direction for further analysis involving regional means and standard deviations.

In Table 31 data are presented on regional means for each control group. Examination of these data reveals a rather puzzling phenomenon. There is really very little variation of mean values in any region, including the rather deviant Northeast and Far West areas. In the Northeast only one control group (S-6) has a mean value which differs more than six per cent from the mean for all metropolitan areas in that region. In the Far West only two

TABLE 31

DISTRIBUTIONS OF REGIONAL MEANS FOR RELIGIOUS AFFILIATION BY THE TOTAL GROUP OF METROPOLITAN STATE ECONOMIC AREAS AND TWELVE SELECTED CONTROL GROUPS

Regions	Total Metrop. Areas	Control Groups*											
		S-1	S-2	S-3	S-4	S-5	S-6	S-7	S-8	S-9	S-10	S-11	S-12
Northeast	53.4%	51.2%	51.8%	51.4%	53.2%	53.5%	46.9%	50.0%	50.0%	50.4%	52.3%	52.5%	53.6%
Southeast	41.4	41.1	39.8	39.7	39.9	39.2	41.2	42.0	40.4	39.4	40.6	43.5	36.8
Middle States	40.6	39.4	38.8	38.7	40.1	40.1	40.2	41.0	39.8	38.9	42.6	43.2	40.6
Northwest	42.3	41.9	43.4	43.4	42.7	42.7	41.9	42.3	43.4	43.6	43.0	42.3	42.3
Southwest	40.2	41.1	35.9	39.9	37.0	32.6	41.1	40.2	35.1	38.8	36.3	40.6	40.7
Far West.	29.2	27.9	25.4	28.5	28.0	28.0	27.9	29.2	24.7	29.7	29.4	30.7	29.2

*The specific controls used are as follow:

S-1: Urban < 85%, Rural farm < 25%.
S-2: " " " , " " " " , Nonfarm occupation = 35%+.
S-3: " " " , " " " " , Unpaid family workers < 10%.
S-4: " " " , " " " " , Farm level of living = 85-124.
S-5: " " " , " " " " , " " " " " " , Nonwhite < 15%.
S-6: " " " , " " " " , Manufacturing employment < 35%.

S-7: Manufacturing employment < 35%.
S-8: " " " " , Nonfarm occupation = 35%+.
S-9: " " " " , Unpaid family workers < 10%.
S-10: " " " " , Farm level living = 85-124.
S-11: Wholesale and retail trade employment = 20%+.
S-12: Nonwhite < 15%.

control groups (S-2 and S-8) have mean values which differ more than four per cent from the mean for the total group of metropolitan areas. In view of this rather marked stability of mean values despite the application of controls, a question arises about the legitimacy of inferring from a comparison of interregional variances that certain control groups bring about a convergence of regional means. Indeed, from the appearance of the data in Table 31, it is only the control group, S-6, which exhibits any readily apparent convergence of means, and this is due primarily to a reduction in the Northeast's mean value. It will be recalled, however, that the data in Table 30 show that control groups S-7 and S-11 have interregional values over 35 per cent smaller than that for all metropolitan areas. Certainly the mean values in Table 31 for these two control groups do not begin to reflect the amount of convergence presumably indicated by the reduction of the interregional variance values.

The explanation for this apparent anomaly, of course, is found in the differential weightings given each region when the control procedure alters the distribution of State economic areas among the several regions. It can be seen from the data in Table 14 (and even more clearly from the data in Table 13 of the Appendix) that the highly deviant Northeast region is very meagerly represented (and thus very lightly weighted) in control groups S-7 and particularly S-11. This, of course, means that the weighted average interregional variance for these control groups will obviously be relatively small because the "potency" of the highly deviant values in the Northeast is reduced.

These observations do not mean that the data here are not

to be trusted at all, but rather that they must be interpreted with caution. As it has been noted before, this requires an analysis that goes beyond the summary measures such as the intraclass correlation and its two major components of variance. Generally, the performances of the regional means in Table 31 indicate at least some convergence of mean values for control groups S-6, S-7, S-11, and perhaps S-9. It will be noted that three of these groups involve a control for manufacturing employment which consistently seems to reduce at least somewhat the proportion of religious affiliation in the Northeast. Since this control represents a selection of metropolitan areas in the Northeast with relatively small proportions of manufacturing employment for that region and since religious affiliation in the Northeast seems to be related to the Roman Catholic element in the population, the thought may be entertained that there is some concentration of Roman Catholics in the more industrialized areas of this region. The evidence for this, however, is only presumptive.

Turning to Table 32, we have regional standard deviations here which can serve as a check on the intraregional variance analysis pursued in Table 30. It is gratifying to be able to report no glaring contradictions between these two sources of information. An inventory of the entire table reveals that 48 per cent of the control groups have sigma values smaller than those for all metropolitan areas. And once again it appears that controls involving the urban-rural farm base control are the most prominent in reducing the internal variation of most of the regions. For these control groups (S-1 through S-5) 60 per cent of the sigma values fall below those for the total metropolitan area groupings. The

TABLE 32

DISTRIBUTIONS OF REGIONAL STANDARD DEVIATIONS FOR RELIGIOUS AFFILIATION BY THE TOTAL GROUP OF METROPOLITAN STATE ECONOMIC AREAS AND TWELVE SELECTED CONTROL GROUPS

Regions	Total Metrop. Areas	Control Groups*											
		S-1	S-2	S-3	S-4	S-5	S-6	S-7	S-8	S-9	S-10	S-11	S-12
Northeast	10.8%	10.2%	10.2%	9.8%	8.1%	8.4%	11.9%	12.7%	12.7%	12.1%	10.5%	6.6%	11.0%
Southeast	8.7	8.4	8.4	8.4	8.7	11.9	8.7	8.6	8.9	8.7	9.3	8.2	10.9
Middle States	7.3	7.4	7.7	7.6	6.8	6.8	8.0	7.7	7.0	7.8	5.7	5.9	7.3
Northwest	8.9	9.7	12.5	12.5	10.8	10.8	9.7	8.9	12.5	10.9	9.8	8.9	8.9
Southwest	11.0	11.2	2.5	12.7	5.5	2.7	11.2	11.0	2.6	12.2	5.5	11.5	14.2
Far West.	5.8	4.3	2.5	4.0	4.2	4.2	4.3	5.8	2.6	5.7	5.8	5.1	5.8
Total U.S.A.. . . .	11.5%	11.1%	11.7%	11.5%	11.0%	11.9%	10.5%	10.8%	11.8%	11.6%	10.9%	9.6%	12.9%

*The specific controls used are as follow:

S-1: Urban < 85%, Rural farm < 25%.
S-2: " " " , " " " " , Nonfarm occupation = 35%+.
S-3: " " " , " " " " , Unpaid family workers < 10%.
S-4: " " " , " " " " , Farm level of living = 85-124.
S-5: " " " , " " " " , " " " " " " , Nonwhite < 15%.
S-6: " " " , " " " " , Manufacturing employment < 35%.

S-7: Manufacturing employment < 35%.
S-8: " " " " , Nonfarm occupation = 35%+.
S-9: " " " " , Unpaid family workers < 10%.
S-10: " " " " , Farm level living = 85-124.
S-11: Wholesale and retail trade employment = 20%+.
S-12: Nonwhite < 15%.

Northeast and Far West are the most conspicuous in this regard, while the Northwest is conspicuous by its failure to conform here.

The standard deviations for the Southwest fluctuate quite widely, although some consistency appears in the fact that controls for nonfarm occupation (S-2 and S-8) and for farm level of living (S-4 and S-10) exhibit remarkably small sigma values, while controls for unpaid farm family workers (S-3 and S-9) have relatively high sigma values. The control on wholesale and retail trade employment (S-11) also appears fairly effective in reducing the internal variation of most of the regions. It should perhaps be pointed out here that this control group exhibits reduced standard deviations for every one of our five dependent variables in the two Eastern regions and for all but one variable (education) in the Middle States. Furthermore, in the Northeast it is this control group which has the smallest sigma values of all control groups for every variable except education.

To summarize briefly the results of this analysis of religious affiliation, it can be said that most of the controls appear to have only a moderate effect on the regional distinctiveness of religious affiliation in metropolitan areas. Most of the regional distinctiveness of this variable rests on the deviant positions of two regions—the Northeast and Far West. Neither of these regions has its position relative to the other regions much altered by the application of controls.

With respect to the modest changes which do occur, those control groups involving a control on manufacturing employment appear most effective in reducing the high proportion of affiliation in the Northeast. This is not, however, a general, across-the-board

reduction for metropolitan areas in this region because the standard deviations for these control groups tend to be larger.

The control groups displaying greater internal homogeneity for most of the regions are those involving the urban-rural farm base control. This is especially true of the Northeast and Far West. Only the Northwest tends to contradict this pattern, and here, none of the controls effects any reduction in the internal variation of this region. In the Southwest it is not so much the base controls which influence internal variation as it is the controls on nonfarm occupation and farm level of living. Wherever these two controls are applied, the sigma value for the Southwest is conspicuously small. The control on wholesale and retail trade employment seems to increase the internal homogeneity of four out of the five regions whose compositions are altered by this control. The most noticeable change here occurs in the Northeast.

Summary and Evaluation

The material presented in this chapter is focused on an analysis of regional differences in the behavior of metropolitan populations when controls are applied primarily to certain structural differences among metropolitan State economic areas. The hypothesis being tested here is that the more similar the structure of organizational units, the more similar will be the behavior of the populations so organized. The measures employed to test this hypothesis include the intraclass correlation (rho), the interregional variance, the intraregional variance, regional means, and regional standard deviations.

On the basis of the hypothesis it is expected that under

controlled conditions the regional distinctiveness of behavior among populations residing in metropolitan areas will be reduced and that this will be reflected by reductions in the values of the intraclass correlation coefficient and the interregional variance. It is also expected that with controls reductions will appear in the values of the intraregional variance and the regional standard deviations. Since the value of the intraclass correlation varies inversely with the value of the interregional variance and directly with the value of the intraregional variance, it follows that reductions in both of these components will effect a reduction in the value of rho only if the changes are greater for the interregional variance than for the intraregional variance.

Twelve control groups were set up on the basis of an examination of the distributions of twelve characteristics as they appear in each region. In the final construction of the control groups, eight characteristics were selected for controls. These characteristics are used either singly or in various combinations with one another, and should be viewed as purely exploratory efforts to investigate relationships between organizational differences and differences in behavior. A summary of the more general findings of this analysis is presented here in the form of an inventory of the differences which appear when the various measures for each control group are compared with those for the uncontrolled or total group of metropolitan State economic areas.

With respect to the intraclass correlations computed for each of our five dependent, behavioral variables, an accounting reveals that two-thirds of the control groups exhibit reductions in the value of rho. If the fertility variable is omitted because

its uncontrolled or total rho value is so low as to render any further reduction rather unexpected, the proportion of the control groups with reduced rho values rises to 77 per cent. Although some of these reductions are fairly small, three-quarters of them exceed 10 per cent.

The frequency with which the several control groups effect reductions in the value of rho is greatest for the two political variables. One-half of the reductions obtained for the voter turnout variable exceed 10 per cent, while three-quarters of the reductions for Democratic voting exceed 10 per cent. All but one of the reductions for the other variables exceed 10 per cent.

In terms of statistical significance the intraclass correlations for fertility fall below the five per cent level for every control group. This, of course, is probably to be expected since even the uncontrolled rho for this variable is not statistically significant at the five per cent level. For two of the other variables, however, the reductions in some cases are great enough to produce measures that are statistically insignificant. This occurs in one of the control groups for Democratic voting and in five of the control groups for educational achievement.

When the reasons for these changes in intraclass correlation values are sought by examining the two major components of variance, an accounting reveals that (omitting fertility) 49 per cent of the reductions in rho involve reductions in both the interregional and intraregional variances, 19 per cent involve larger variances for each component under controlled conditions, and 32 per cent involve smaller interregional values and larger intraregional values. Ideally, of course, our hypothesis prompts the

expectation that both variance components will be reduced. In only one-half of our cases, however, where rho is reduced, does this appear to be true. Generally speaking, the control groups with the urban-rural farm base control (S-1 through S-4) have smaller variance components than the control groups with the manufacturing employment base control (S-7 through S-10). This is particularly true of the intraregional variance. Considering all five behavioral variables, the most successful control in terms of reducing variance components is that for wholesale and retail trade employment (S-11). The least successful in this respect is the control for manufacturing employment only (S-7). It should also be noted here that the political variables are highly sensitive to controls on the nonwhite proportion of the total population. This, of course, is due to the strategic importance of this characteristic for political behavior in the Southern regions.

The data on regional standard deviations provide an opportunity to see the extent to which controls serve to reduce the internal variation of particular regions and particular variables. With respect to all variables and all regions, 53 per cent of the control groups exhibit sigma values which are smaller than those for the total metropolitan areas in each region. These reductions occur most frequently in the Far West and the Southeast regions. Eighty per cent of the control groups show reductions in the former region, while 63 per cent of the control groups show sigma reductions in the latter. The region where controls effect reductions the least frequently (viz., 40 per cent) is the Northwest.

The educational achievement variable appears to be the most sensitive to the controls exercised here, while voter turnout

seems to be the least sensitive in terms of the frequency with which controls effect reductions in sigma value. The regions seem to vary considerably in their responsiveness to various controls. The Northeast and Southeast display small sigma values for all variables when wholesale and retail trade employment is controlled. The Middle States and Far West also tend to react this way for this control group. On the other hand, this control reduces the standard deviation of none of the variables for the Southwest. This region seems to be most sensitive to controls on nonfarm occupation and farm level of living. It is the manufacturing base control groups that are most effective in reducing the internal variation of the political variables in the industrial regions of the Northeast and Middle States. On the other hand, in the Southwest and Far West it is the urban-rural farm base control groups which tend to display reduced sigma values for these political variables. The urban-rural farm base control groups also display sigma reductions for the educational achievement variable. This holds true for all regions except the Middle States, where none of the control groups seems to influence this education variable very much.

Broadly speaking, the control groups which tend to exert a reductive influence on the internal variation of most of the variables in most of the regions are the following: S-11 wholesale and retail trade employment, and S-2 urban-rural farm population with the additional control on nonfarm occupation. The control groups which are least effective in this regard are the following: S-7 manufacturing employment, and S-12 proportion of population which is nonwhite.

The question of how these findings reflect upon the hypothesis under test here is not to be answered in any categorical manner. Nevertheless, a few definite remarks can be made about the results. First, there is clearly a differential susceptibility to particular controls among the several regions. This is evident from even a casual observation of the way in which certain characteristics are distributed among the different regions. Second, there is clearly a differential susceptibility to particular controls among the several variables examined. None of the controls affects all variables in the same degree. Third, regional differences in some types of behavior are quite definitely related to some regional differences in the organization of metropolitan populations. The extent of this relationship cannot be precisely determined. Furthermore, the evidence indicates that the extent of this relationship will vary with different kinds of behavior. Fourth, the various regions tend not to be equally distinctive with respect to any particular behavior. That is to say, regional distinctiveness in behavior usually rests on the deviant position of one or two or perhaps three regions, while the others exhibit considerable similarity.

One of the results of this analysis is particularly interesting in light of the hypothetical expectations noted earlier. The reference here is to the consistency with which the control on wholesale and retail trade employment produces results in line with the hypothesis. It will be recalled that the control for this characteristic is one which selects metropolitan State economic areas with a high proportion of their labor forces employed in this type of enterprise. It seems to the writer that these units

represent in one sense the epitome of the urban type of organization which mediates exchange relationships among different producers. It is here that one finds the commercial nexus in full bloom, so to speak. In so far as this is true, regional differences which are maintained by conditions of relative isolation should very definitely be reduced among such units. This is precisely what happens, although modified somewhat by the fact that there is a relatively large reduction in the internal variation of the different regions when this control is exercised.

The remark was made above that it is difficult to determine the extent to which regional differences in behavior are related to the particular control characteristics employed in this analysis. The reason for this has been discussed before, but should be mentioned again. The difficulty arises out of the control procedure used in this study. When control is exercised by a selection and rejection of units, the representation of each region is usually modified to some extent. This means that the relative weighting of each region in terms of the total sample is altered. Since the summary measures of regional distinctiveness involve weighted averages, the changes which occur will be contingent (in some undetermined degree) upon this alteration of regional weights. It should be noted here that this problem is not confined to the control procedure. It also influences the summary measures for the entire universe of units. That is to say, even before the controls are applied the regions do not stand in a balanced, vis-à-vis relationship with respect to the summary measures of regional distinctiveness used here.

There are perhaps some statistical techniques which can be

used to redress this regional imbalance. It should be recognized, however, that this imbalance is genuine. This is the way the real world is, and this is the way your data come to you. No amount of statistical manipulation can alter the fact that the Southeast has 129 State economic areas while the Far West has only 34. This fact may also be of consequence for the kinds of social phenomena exhibited by these regions. The central issue here seems to be one that revolves around the question of whether statistical measures are intended to describe the real world or pure relationships. The answer will, of course, depend on the problem at hand. In this study, the analysis of the influence of structural controls on regional differences in behavior represents primarily an interest in relationships. Accordingly, the difficulties involved in the measuring instruments used here cannot be ignored. That is why considerable space has been devoted to their discussion.

CHAPTER V

THE FINDINGS: NONMETROPOLITAN STATE ECONOMIC AREAS WITH APPLIED CONTROLS

The findings in this chapter represent an extension of the analysis of nonmetropolitan State economic areas reported in Chapter III. The organization of the material to be reported here parallels that appearing in the foregoing chapter, which deals with the metropolitan areas under controlled conditions. The purpose of this analysis is to explore the effects of various controls on the behavior of nonmetropolitan populations.

Preliminary Remarks and the Specification of Controls

In Chapter IV there were discussed several of the problems involved in the control procedure used in this study. Most of the points made in that discussion are pertinent for the analysis to be reported here. Care had to be exercised to assure an adequate total sample size for the control groups. Each region had to be assured at least some minimum representation when controls were applied. And to meet these requirements the selection of controls involved a certain amount of ad hoc manipulation based on an examination of the distributions of control characteristics within each region. Tables showing these distributions are presented in Appendix B.

The controls finally selected involve six characteristics

which are used in various combinations with one another. In line with the procedure used for the metropolitan area analysis, an effort was made to establish two different base control groups. These are distinguished from one another by differences in the proportion of the labor force employed in agriculture. It should be pointed out here that the general classification of "nonmetropolitan" is not necessarily a rural or agricultural designation. As can be seen from a glance at the tables in Appendix B, the nonmetropolitan State economic areas exhibit a great deal of variety in their structural characteristics. In other words, nonmetropolitan areas really constitute a residual category which is homogeneous only in a negative sense. It is for this reason that the two base control groups were established—one representing a rather large commitment of the population to rural, agricultural enterprise, and the other representing a relatively small commitment of this kind.

In addition to this control on agricultural employment (see Appendix B, Table 5), several other controls are used which involve the following characteristics: (a) the proportion of the labor force employed in wholesale and retail trade (Appendix B, Table 7), (b) rural farm level of living (Appendix B, Table 10), (c) the proportion of the total population which is nonwhite (Appendix B, Table 4), (d) the proportion of the rural farm labor force in nonfarm occupations (Appendix B, Table 12), (e) the proportion of the rural farm labor force who are unpaid family workers (Appendix B, Table 11). Using these characteristics together with the base control on agricultural employment, the following ten control groups were set up:

N-1: Between 10 and 34.9 per cent of the labor force employed in agriculture (base control).

N-2: Between 10 and 34.9 per cent of the labor force employed in agriculture, 13 per cent or more of the labor force employed in wholesale and retail trade.

N-3: Between 10 and 34.9 per cent of the labor force employed in agriculture, rural farm level of living index between 85 and 124, inclusive.

N-4: Between 10 and 34.9 per cent of the labor force employed in agriculture, less than 15 per cent of the population nonwhite.

N-5: Between 30 and 54.9 per cent of the labor force employed in agriculture (base control).

N-6: Between 30 and 54.9 per cent of the labor force employed in agriculture, less than 25 per cent of the rural farm labor force employed in nonfarm occupations.

N-7: Between 30 and 54.9 per cent of the labor force employed in agriculture, less than 10 per cent of the rural farm labor force who are unpaid family workers.

N-8: Between 30 and 54.9 per cent of the labor force employed in agriculture, rural farm level of living index between 90 and 124, inclusive.

N-9: Between 30 and 54.9 per cent of the labor force employed in agriculture, less than 20 per cent of the population nonwhite.

N-10: Less than 5 per cent of the population nonwhite.

It is perhaps apparent that the first four control groups (N-1 through N-4) represent populations which are not directly involved with agricultural enterprise in any large degree, while the next five control groups (N-5 through N-9) represent much more agriculturally oriented populations. This is especially true of control group N-6, which has a low proportion of its rural labor force in nonfarm occupations.[1] The control on rural farm level of

[1] To assure adequate representation in the Southwest and Far West regions it was found necessary to permit some overlap in

living (N-3 and N-8) is selective of units in the upper, middle range of this characteristic. This means that areas with low levels of living, found primarily in the two Southern regions, are rejected together with the areas of highest living standard appearing in the other four regions. The control on unpaid farm family workers (N-7) represents a selection of agricultural populations which are not primarily subsistence farming people, and thus probably constitutes an indirect control on level of living. And, of course, the control on the nonwhite proportion of the population (N-4, N-9, and N-10) serves primarily to limit the representation of the Southern regions.

In Table 33 data are presented which show the total control sample sizes and the regional distributions of the sampling units for each control group and for the total group of nonmetropolitan State economic areas. Percentage distributions of these data may also be found in Table 26 of Appendix B. It can be seen that none of the control groups has an N smaller than 70, and every region is always represented by at least six nonmetropolitan areas. The fact that these nonmetropolitan control group samples tend to be larger than those set up for the metropolitan area analysis can be attributed to the relatively large number of areas available for control selection here. There are almost exactly twice as many nonmetropolitan areas as there are metropolitan areas in the country as a whole.

One noteworthy feature of the controls exercised in this

---

the base control "breaks" on agricultural employment. Thus, State economic areas with 30 to 34.9 per cent of their labor force personnel so employed are included in both the control groups N-1 through N-4 and N-5 through N-9.

TABLE 33

NUMBER OF NONMETROPOLITAN STATE ECONOMIC AREAS IN EACH REGION FOR THE TOTAL GROUP OF NONMETROPOLITAN STATE ECONOMIC AREAS AND TEN SELECTED CONTROL GROUPS

Regions	Total Nonmet. Areas	Control Groups*									
		N-1	N-2	N-3	N-4	N-5	N-6	N-7	N-8	N-9	N-10
Northeast	43	24	19	22	23	10	4	7	8	8	33
Southeast	94	23	8	9	7	48	37	7	7	14	12
Middle States	68	30	29	15	29	41	32	21	23	41	66
Northwest	39	9	7	9	9	25	25	14	22	25	36
Southwest	29	9	8	6	6	23	21	10	10	19	11
Far West.	21	14	14	12	14	11	8	11	9	11	17
Total U.S.A..	294	109	85	73	88	158	127	70	79	118	175

*The specific controls used are as follow:

N-1: Agricultural employment = 10-34.9%.
N-2: " " " " , Wholesale and retail trade employment = 13%+.
N-3: " " " " , Farm level living = 85-124.
N-4: " " " " , Nonwhite < 15%.

N-5: Agricultural employment = 30-54.9%.
N-6: " " " " , Nonfarm occupation < 25%.
N-7: " " " " , Unpaid family workers < 10%.
N-8: " " " " , Farm level living = 90-124.
N-9: " " " " , Nonwhite < 20%.
N-10: Nonwhite < 5%.

analysis is the opportunity provided by the control groups, N-5 through N-9, for examining the extent to which the essence of regionalism resides in rural, agricultural areas. That is to say, if regionalism is born and nurtured under conditions of isolation, it should be most clearly manifest in rural areas where social contact is presumably rather limited. These five control groups, N-5 through N-9, represent the most rural, agricultural areas analyzed in this study. Accordingly, the isolation hypothesis accounting for cultural diversity among regions prompts the expectation that here one may find the most persuasive supporting evidence.

Before presenting the substantive material of this chapter, some reference should be made to the data in Table 34. This is a general reference table showing the values of rho necessary to achieve specified levels of statistical significance for each of our sample control groups. In addition, the absolute minimum value which rho can possibly have is also presented. This table may be regarded as a crude "yardstick" which can be used for evaluating the findings to be reported.

Following the format of the foregoing chapter, the substantive findings in this chapter will be presented and discussed for each dependent variable separately. A concluding section will then summarize and evaluate the general results of this analysis.

Net Reproduction Rate

The fertility data for nonmetropolitan State economic areas presented in Chapter III revealed that 14 per cent of the total variation in this type of behavior is associated with regional location. While this degree of association is not particularly large,

TABLE 34

MINIMUM VALUES OF RHO FOR SIGNIFICANCE LEVELS OF FIVE PER CENT AND ONE PER CENT; ABSOLUTE MINIMUM VALUES OF RHO; TOTAL N FOR TOTAL GROUP OF NONMETROPOLITAN STATE ECONOMIC AREAS AND TEN SELECTED CONTROL GROUPS

Listed Items	Total Nonmet. Areas	Control Groups*									
		N-1	N-2	N-3	N-4	N-5	N-6	N-7	N-8	N-9	N-10
5% Level	.026	.076	.103	.124	.096	.050	.066	.130	.112	.071	.046
1% Level	.043	.129	.174	.211	.164	.084	.110	.220	.188	.119	.077
Absolute Minimum Rho . .	-.021	-.058	-.076	-.090	-.073	-.039	-.050	-.094	-.082	-.054	-.036
Total N.	294	109	85	73	88	158	127	70	79	118	175

*The specific controls used are as follow:

N-1: Agricultural employment = 10-34.9%.
N-2: " " " " , Wholesale and retail trade employment = 13%+.
N-3: " " " " , Farm level living = 85-124.
N-4: " " " " , Nonwhite < 15%.

N-5: Agricultural employment = 30-54.9%.
N-6: " " " " , Nonfarm occupation < 25%.
N-7: " " " " , Unpaid family workers <10%.
N-8: " " " " , Farm level living = 90-124.
N-9: " " " " , Nonwhite < 20%.
N-10: Nonwhite < 5%.

it is statistically significant at the one per cent level and is sufficiently large to justify further investigation under controlled conditions. In Table 35 data are presented on the intraclass correlation (rho) and the two major variance components for each control group used in this analysis.

It may be seen from these data that all but three control groups (N-1, N-4, and N-10) have rho values which are smaller than that for all nonmetropolitan areas taken together. Some of these reductions are rather substantial. Indeed, reference to Table 34 shows that four of the control groups display intraclass correlations which fall below the five per cent level of significance (see N-3, N-6, N-7, and N-8). It is also interesting to note that all but one of the control groups involving the highly agricultural base control (see N-5 through N-8) exhibit rather large reductions in rho value. This seems to indicate that even in rural, agricultural areas the relationship between regionalism and native white fertility is negligible. The only control groups showing an appreciably larger intraclass correlation are two of the groups with a control for the nonwhite segment of the population (N-4 and N-10). This merits further investigation which will be reported later.

The immediate reasons for the altered rho values among the several control groups can be found from an inspection of the variance components in Table 35. All of the control groups display intraregional variance values which are smaller than that for the total group of nonmetropolitan areas. None of these reductions is less than 18 per cent, and two of them exceed 50 per cent (N-2 and N-3). With respect to the interregional variance, all but three of the control groups (N-1, N-4, and N-10) exhibit reductions of

TABLE 35

A LISTING FOR FERTILITY OF THE VALUES OF RHO, THE INTERREGIONAL VARIANCE, THE INTRAREGIONAL VARIANCE, AND N FOR THE TOTAL GROUP OF NONMETROPOLITAN STATE ECONOMIC AREAS AND TEN SELECTED CONTROL GROUPS

Listed Items	Total Nonmet. Areas	Control Groups*									
		N-1	N-2	N-3	N-4	N-5	N-6	N-7	N-8	N-9	N-10
Rho	.142	.164	.132	.076	.324	.082	.033	.046	.019	.128	.399
Interregional Variance.	7,527	7,632	4,093	3,207	13,548	4,113	2,245	4,328	2,647	6,716	22,464
Intraregional Variance.	39,638	28,794	17,092	17,858	23,096	31,185	26,134	29,548	25,627	32,242	31,029
N	294	109	85	73	88	158	127	70	79	118	175

*The specific controls used are as follow:

N-1: Agricultural employment = 10-34.9%.
N-2: " " " " , Wholesale and retail trade employment = 13%+.
N-3: " " " " , Farm level living = 85-124.
N-4: " " " " , Nonwhite < 15%.

N-5: Agricultural employment = 30-54.9%.
N-6: " " " " , Nonfarm occupation < 25%.
N-7: " " " " , Unpaid family workers < 10%.
N-8: " " " " , Farm level living = 90-124.
N-9: " " " " , Nonwhite < 20%.
N-10: Nonwhite < 5%.

10 per cent or greater. Understandably, it is these three exceptions which also show no reduction in the value of rho. This is especially evident for the two control groups with the largest intraclass correlations (N-4 and N-10). Here the interregional variance values are respectively 80 per cent and 198 per cent larger than that for all nonmetropolitan areas. Quite obviously something rather dramatic has happened either to the regional mean values or to the regional distribution of units for these control groups. In order to see more specifically the reasons for the fluctuating values of the interregional variance, data on regional means for each control group are presented in Table 36.

Generally speaking, all of the regions except for the two Southern ones tend to have higher net reproduction rates among the more highly agricultural populations (N-5 through N-9) than among those populations less agriculturally inclined (N-1 through N-4), although some of the differences are small. This, of course, follows from the effects of the two base control groups on the fertility rates of these four non-Southern regions. Among populations with a low degree of agricultural involvement (N-1 through N-4), 94 per cent of the comparisons between the total nonmetropolitan areas in each region and their respective control group areas show the latter having lower fertility rates. On the other hand, among populations with a high degree of agricultural involvement (N-5 through N-9) 80 per cent of the comparisons between the total nonmetropolitan areas in each region and their respective control group areas show the latter having higher fertility rates. These findings are in line with the traditional gradient pattern which shows fertility to be inversely related to urbanism.

TABLE 36

DISTRIBUTIONS OF REGIONAL MEANS FOR FERTILITY BY THE TOTAL GROUP OF NONMETROPOLITAN STATE ECONOMIC AREAS AND TEN SELECTED CONTROL GROUPS

Regions	Total Nonmet. Areas	Control Groups*									
		N-1	N-2	N-3	N-4	N-5	N-6	N-7	N-8	N-9	N-10
Northeast	1127	1115	1097	1131	1120	1348	1290	1220	1264	1363	1138
Southeast	1322	1276	1114	1144	1494	1283	1247	1156	1102	1364	1710
Middle States	1181	1100	1094	1148	1095	1205	1188	1142	1229	1205	1175
Northwest	1199	1112	1078	1112	1112	1193	1193	1251	1212	1193	1205
Southwest	1278	1292	1281	1261	1258	1276	1257	1263	1161	1286	1317
Far West	1037	1029	1029	1035	1029	1080	1072	1080	1104	1080	1033

*The specific controls used are as follow:

N-1: Agricultural employment = 10-34.9%.
N-2: " " " " , Wholesale and retail trade employment = 13%+.
N-3: " " " " , Farm level living = 85-124.
N-4: " " " " , Nonwhite < 15%.

N-5: Agricultural employment = 30-54.9%.
N-6: " " " " , Nonfarm occupation < 25%.
N-7: " " " " , Unpaid family workers < 10%.
N-8: " " " " , Farm level living = 90-124.
N-9: " " " " , Nonwhite < 20%.
N-10: Nonwhite < 5%.

The performance of fertility rates under controlled conditions in the two Southern regions is rather erratic. In the Southwest very little change tends to occur at all. In the Southeast, however, there is considerable fluctuation. The outstandingly high rates in this region appear when controls are applied on the nonwhite proportion of the population (N-4, N-9, and N-10). Since these controls select areas with relatively small proportions of nonwhite persons, we have the same phenomenon here that was observed for metropolitan areas in the previous chapter--namely, a high, native white fertility rate among Southern populations which have few Negro members. Furthermore, as was reported in the metropolitan analysis, a detailed inspection of these high fertility areas reveals them to be primarily areas with a low standard of living as measured by the rural farm level of living index. They also tend to rank low in terms of per capita dollar retail sales and per capita dollar receipts from services.

One may not properly infer from these findings that Southern areas with few Negroes are <u>invariably</u> depressed areas because, as will be seen shortly, the regional standard deviations for these nonwhite control groups are fairly large. However, there may be a tendency in this direction. If this should prove to be true, it should prompt some interesting speculation and research about the position of the Negro with respect to the material welfare of the South.

The effects of controls on the internal variation of each region are shown in the data on regional standard deviations presented in Table 37. Some of the regions--notably the Northeast, Southeast, Middle States, and Northwest--display considerable

TABLE 37

DISTRIBUTIONS OF REGIONAL STANDARD DEVIATIONS FOR FERTILITY BY THE TOTAL GROUP OF NONMETROPOLITAN STATE ECONOMIC AREAS AND TEN SELECTED CONTROL GROUPS

Regions	Total Nonmet. Areas	Control Groups*									
		N-1	N-2	N-3	N-4	N-5	N-6	N-7	N-8	N-9	N-10
Northeast	262	154	144	147	155	291	290	252	264	290	233
Southeast	226	263	159	156	314	186	156	183	134	237	197
Middle States	151	114	106	112	111	136	140	97	138	136	144
Northwest	179	128	125	128	128	177	177	213	180	177	180
Southwest	172	182	187	147	152	178	171	201	78	186	206
Far West	103	102	102	108	102	118	88	118	118	118	99
Total U.S.A.	217	192	146	147	191	189	167	185	167	198	185

*The specific controls used are as follow:

N-1: Agricultural employment = 10-34.9%.
N-2: " " " " , Wholesale and retail trade employment = 13%+.
N-3: " " " " , Farm level living = 85-124.
N-4: " " " " , Nonwhite < 15%.

N-5: Agricultural employment = 30-54.9%.
N-6: " " " " , Nonfarm occupation < 25%.
N-7: " " " " , Unpaid family workers < 10%.
N-8: " " " " , Farm level living = 90-124.
N-9: " " " " , Nonwhite < 20%.
N-10: Nonwhite < 5%.

fluctuation in sigma value under various controls. This is less evident in the Southwest and Far West. Considering all regions, 65 per cent of the control groups have smaller sigmas than those of each region's total group of nonmetropolitan areas. In terms of the two base controls, it is the 10 to 34 per cent agricultural employment control groups (N-1 through N-4), which, generally speaking, exhibit the most frequent and the largest reductions in sigma value. This is especially true for the Northeast, the Middle States, and the Northwest regions. The greatest reductions in the two Southern regions occur when controls on rural farm level of living are exercised (N-3 and N-8). In the Far West the standard deviations are small and change very little under controls. It should also be noted that in the Southeast the nonwhite control groups (N-4, N-9, and N-10) tend to have rather large standard deviations. It will be recalled that these are the control groups with high mean fertility values.

Summarizing briefly, the relationship between fertility and regionalism appears to be rather small for nonmetropolitan populations. Furthermore, most of the regional differences which do occur seem to be rather definitely attributable to regional differences in social structure and in standard of living. Generally speaking, the more agricultural populations tend to be the more fertile populations. This is true for every region except the Southwest, where a fairly high fertility rate seems to prevail among nonmetropolitan populations regardless of the degree of their agricultural involvement. There is some indication in this region, however, that standard of living is a characteristic to which fertility is sensitive.

Proportion of Eligible Voters Voting

It has been shown in previous chapters that of all our dependent variables, voter turnout is the most highly correlated with regionalism. Furthermore, this correlation tends to be altered the least when controls are applied. The data on voter turnout in nonmetropolitan areas to be presented in this chapter are not exceptional in this regard. It is true that most of the high correlation can be attributed to the deviant values found in the Southern regions, where voter turnout is quite low. This does not necessarily mean, however, that political interests and behavior in the South are neglected. It may only mean that in the South, with its one-party political structure, political issues are settled in the primary elections for which we have no data in this study.

In Table 38 there are presented intraclass correlations and measures of the two variance components for each control group and for the total group of nonmetropolitan areas. It is evident that no dramatic changes in the value of rho occur when controls are applied. The greatest reductions (15 or 16 per cent) appear when controls are applied to the nonwhite proportion of the population (N-4, N-9, and N-10), but even these controls still leave over 60 per cent of the total variance in voter turnout associated with regional location. It may also be noted that the rho values tend to be greater for the more rural control groups (N-5 through N-8). Indeed, for the most rural groups with no other controls involved (N-5 and N-6) the value of rho actually is larger than that for all nonmetropolitan areas considered together.

These mildly fluctuating values of the intraclass

TABLE 38

A LISTING FOR PROPORTION OF ELIGIBLE VOTERS VOTING OF THE VALUES OF RHO, THE INTERREGIONAL VARIANCE, THE INTRAREGIONAL VARIANCE, AND N FOR THE TOTAL GROUP OF NONMETROPOLITAN STATE ECONOMIC AREAS AND TEN SELECTED CONTROL GROUPS

Listed Items	Total Nonmet. Areas	Control Groups*									
		N-1	N-2	N-3	N-4	N-5	N-6	N-7	N-8	N-9	N-10
Rho	.749	.734	.662	.675	.632	.771	.776	.730	.667	.639	.639
Interregional Variance.	440.8	279.2	153.2	145.6	126.5	469.8	463.2	291.9	196.4	213.9	124.0
Intraregional Variance.	143.8	93.8	70.1	61.8	66.1	132.9	125.3	95.5	87.2	111.4	66.5
N	294	109	85	73	88	158	127	70	79	118	175

*The specific controls used are as follow:

N-1: Agricultural employment = 10-34.9%.
N-2: " " " " , Wholesale and retail trade employment = 13%+.
N-3: " " " " , Farm level living = 85-124.
N-4: " " " " , Nonwhite < 15%.

N-5: Agricultural employment = 30-54.9%.
N-6: " " " " , Nonfarm occupation < 25%.
N-7: " " " " , Unpaid family workers < 10%.
N-8: " " " " , Farm level living = 90-124.
N-9: " " " " , Nonwhite < 20%.
N-10: Nonwhite < 5%.

correlation are, of course, due to the fluctuations of the total, average variance components. An inspection of these values in Table 38 reveals some very substantial changes. Eight of the ten control groups have interregional variances which are from 34 to 72 per cent lower than that for all nonmetropolitan areas. As might be expected, the greatest of these reductions occurs when the most rigorous control on the nonwhite proportion of the population is exercised (N-10). As might also be expected, the only two control groups which do not exhibit a smaller interregional variance (N-5 and N-6) are the control groups which have larger intraclass correlations than that for the total group of nonmetropolitan areas.

In view of the rather large reductions in interregional variance for most of the control groups, one might reasonably expect to find larger reductions in the value of rho than those observed in Table 38. It may be seen, however, that the smaller interregional variance values which appear under controlled conditions are also accompanied by rather large reductions in the values of the intraregional variance. Indeed, every control group has an interregional variance which is from 8 to 57 per cent smaller than that for all nonmetropolitan areas. In other words, the application of controls (with two exceptions, N-5 and N-6) brings about fairly substantial reductions in both variance components, which means that as far as rho is concerned, the changes are compensatory, and the values of rho are altered only slightly.

Pursuing further the analysis of the effects of controls on regional differences in voter turnout, attention is directed to the data on regional means presented in Table 39. Except for the

TABLE 39

DISTRIBUTIONS OF REGIONAL MEANS FOR PROPORTION OF ELIGIBLE VOTERS VOTING BY THE TOTAL GROUP OF NONMETROPOLITAN STATE ECONOMIC AREAS AND TEN SELECTED CONTROL GROUPS

Regions	Total Nonmet. Areas	Control Groups*									
		N-1	N-2	N-3	N-4	N-5	N-6	N-7	N-8	N-9	N-10
Northeast	68.6%	69.9%	69.9%	69.4%	70.4%	70.0%	66.8%	67.2%	67.0%	72.1%	69.8%
Southeast	28.2	33.7	38.5	42.2	39.4	27.5	26.6	26.8	36.9	40.0	43.9
Middle States	76.4	77.2	76.9	77.9	76.9	78.0	76.4	82.1	77.7	78.0	76.3
Northwest	76.2	74.6	73.6	74.6	74.6	76.0	76.0	76.1	76.2	76.0	76.5
Southwest	44.1	47.1	46.8	45.4	47.4	44.8	44.9	49.4	45.6	47.0	41.3
Far West.	65.6	65.0	65.0	64.8	65.0	62.6	63.1	62.6	62.2	62.6	64.7

*The specific controls used are as follow:

N-1: Agricultural employment = 10-34.9%.
N-2: " " " " , Wholesale and retail trade employment = 13%+.
N-3: " " " " , Farm level living = 85-124.
N-4: " " " " , Nonwhite $<$ 15%.

N-5: Agricultural employment = 30-54.9%.
N-6: " " " " , Nonfarm occupation $<$ 25%.
N-7: " " " " , Unpaid family workers $<$ 10%.
N-8: " " " " , Farm level living = 90-124.
N-9: " " " " , Nonwhite $<$ 20%.
N-10: Nonwhite $<$ 5%.

Southeast, very little change occurs in the regional mean values of the control groups. In the Northeast and Far West there may be a slight tendency for voter turnout to be higher in the less agricultural areas (N-1 through N-4), while in the Middle States and the Northwest the opposite tendency appears, with the more agricultural populations (N-5 through N-8) displaying higher voter turnout. It is doubtful, however, that these differences would prove to be significant. Among these four regions only one control group (N-7 in the Middle States) has a regional mean that differs as much as seven per cent from the mean of its total parent group.

The stability of these regional means indicates that changes in the correlation between regionalism and voter turnout under various controls are primarily due to changes in the two Southern regions, particularly the Southeast. The importance of these two regions for the overall correlation of .749 (when all nonmetropolitan areas are included) is indicated by the fact that the intraclass correlation drops to .207 when the two Southern regions are excluded. Needless to say, the difference between a classification that "explains" 75 per cent of a variance and one that "explains" 21 per cent is not to be ignored. This evidence provides further support for one of the conclusions stated in the foregoing chapter—namely, that regional distinctiveness in behavior tends to rest on the distinctiveness of only one or two regions while the others appear fairly similar to one another.

Although we have noted that the shifts in mean value of the Southeast (and to some extent the Southwest) are relatively large when controls are exercised, a close inspection of Table 39 does not show differences of such a magnitude as to lead one to expect

the kinds of reductions in the interregional variance which do appear in Table 38. Six of the control groups display interregional variance values in that table which are over 50 per cent smaller than that for all nonmetropolitan areas. Yet only one control group (N-10) affects the Southeast enough to alter its mean value by as much as 50 per cent, and this is the greatest change in mean value that appears in the entire Table 39. As may be surmised, the explanation for this is to be found in Table 33 (or Table 26 of Appendix B) where the regional distributions of nonmetropolitan areas are presented. In only two control groups (N-5 and N-6) does the proportionate representation of the Southeast even approximate the representation of this region when the total group of nonmetropolitan areas is measured. In other words, we are confronted again by the problem of changes in weighted average measures which are attributable in part to changes in the "weighting pattern" itself. To anticipate somewhat, this weighting factor also influences to some extent the rather remarkable reductions already noted for the intraregional variance.

The data on regional standard deviations presented in Table 40 provide an opportunity to investigate the effects of controls on the internal variation of each region. For the table as a whole 65 per cent of the standard deviations for the control groups are smaller than those for the total uncontrolled populations. Among the more agricultural groups (N-5 through N-9) the data show 70 per cent of the sigma values as reductions, while among the less agricultural populations (N-1 through N-4) the proportion of reductions is only 54 per cent. There is considerable variety, however, in the manner in which the various regions

TABLE 40

DISTRIBUTIONS OF REGIONAL STANDARD DEVIATIONS FOR PROPORTION OF ELIGIBLE VOTERS VOTING BY THE TOTAL GROUP OF NONMETROPOLITAN STATE ECONOMIC AREAS AND TEN SELECTED CONTROL GROUPS

Regions	Total Nonmet. Areas	Control Groups*									
		N-1	N-2	N-3	N-4	N-5	N-6	N-7	N-8	N-9	N-10
Northeast	13.5%	8.2%	8.6%	8.4%	8.1%	10.6%	12.1%	10.4%	9.8%	10.7%	8.3%
Southeast	15.2	14.7	11.3	8.5	15.0	14.4	13.6	14.2	16.5	17.3	14.8
Middle States	6.6	7.3	7.2	6.4	7.1	6.3	6.2	5.8	6.9	6.3	6.8
Northwest	3.7	3.9	3.9	3.9	3.9	3.5	3.5	4.0	4.3	3.5	3.3
Southwest	17.5	13.7	14.5	14.6	12.4	18.4	18.2	18.5	16.2	17.7	16.9
Far West.	4.0	4.0	4.0	4.6	4.0	3.8	3.1	3.8	3.4	3.8	3.2
Total U.S.A..	24.2%	19.4%	14.9%	14.4%	13.9%	24.6%	24.3%	19.6%	16.8%	18.1%	13.7%

*The specific controls used are as follow:

N-1: Agricultural employment = 10-34.9%.
N-2: " " " " , Wholesale and retail trade employment = 13%+.
N-3: " " " " , Farm level living = 85-124.
N-4: " " " " , Nonwhite < 15%.

N-5: Agricultural employment = 30-54.9%.
N-6: " " " " , Nonfarm occupation < 25%.
N-7: " " " " , Unpaid family workers < 10%.
N-8: " " " " , Farm level living = 90-124.
N-9: " " " " , Nonwhite < 20%.
N-10: Nonwhite < 5%.

respond to the different controls. The more highly agricultural populations in the Northeast and Southwest tend to be more heterogeneous than the less agricultural populations of these regions, while the opposite pattern appears in the Middle States and the Far West. The Southeast and Northwest are a bit erratic in this regard.

It is perhaps not surprising to observe in Table 40 that the regions displaying the most frequent and the largest (proportionately speaking) reductions when controls are applied are those with the largest sigma values to begin with-namely, the Northeast and the two Southern regions. This means, of course, that at least part (and perhaps a good deal) of the reductions observed in the total intraregional variance (Table 38) are genuine. That is to say, the problem of shifts in the weighting pattern does not appear to assume overwhelming importance in accounting for the reductions in internal variation when controls are applied. Undoubtedly, there is some influence in this respect, but it is largely confined to a redistribution of the Southeast's weighting factor among several of the other regions (see Table 26, Appendix B).

In view of the fact that controls on the nonwhite proportion of the population seem to yield the greatest reductions in rho value, some additional data can be presented that permit further investigation of this phenomenon in the Southeast, which is the region most sensitive to this control. These data parallel in form those presented in Chapter IV for metropolitan areas. All of the nonmetropolitan areas in the Southeast have been classified by the proportions of their populations which are nonwhite. Five

categories have been set up for each of which means and standard deviations have been computed. These data are presented in Table 41.

TABLE 41

MEAN VALUES AND STANDARD DEVIATIONS FOR PROPORTION OF ELIGIBLE VOTERS VOTING IN THE NONMETROPOLITAN STATE ECONOMIC AREAS OF THE SOUTHEAST REGION BY THE PROPORTION OF THE TOTAL POPULATION WHICH IS NONWHITE

Proportion Nonwhite	Mean	Standard Deviation	N
40%+	16.3%	7.8%	23
30-39.	24.9	9.2	25
20-29.	28.3	11.7	12
10-19.	34.1	16.8	13
Under 10	41.3	16.1	21
Total Region.	28.2%	15.2%	94

It is immediately apparent that the proportion nonwhite and voter turnout are variables which are inversely related to one another in the Southeast. This is true of both the mean and standard deviation values. These values also correspond rather closely to those presented in Table 22 for metropolitan areas. The intraclass correlation for this distribution of nonmetropolitan areas is .311. As discussed earlier, this can be interpreted as a measure of the improvement in prediction which knowledge of the nonwhite proportion of the population can provide. It should be noted that these findings for nonmetropolitan areas, which are based on a rather substantial N, lend confidence to those reported for metropolitan areas. The importance of the Negro for political behavior

in the Southeast can hardly be disputed.

To summarize, the data reveal that nonmetropolitan voter turnout is rather highly correlated with regionalism. None of the controls applied in this analysis alter this correlation very much. In the Middle States, the Northwest, and the Far West voter turnout is remarkably stable among the various population subgroups in these regions. To a somewhat less extent this is also true of the Northeast. These four regions are also fairly similar in the proportions of their eligible voters who voted in 1940.

It is in the two Southern regions, particularly the Southeast, where regional distinctiveness in voter turnout is most evident. Indeed, most of the overall correlation between voter turnout and regionalism can be attributed to the "deviant behavior" of the residents in these regions. Within the Southeast and Southwest areas there is considerable variation among the various population subgroups—a variation that is most clearly related to the incidence of nonwhite residents, although some other factors also appear to be influential: viz., standard of living, degree of agricultural involvement, etc. Generally speaking, there are slight indications that the regional distinctiveness of the Southern regions (especially the Southeast) is more pronounced among the more agricultural segments of the population.

Proportion of Voters Voting Democratic

It has been shown in previous chapters that next to the variable on voter turnout, the voter preference variable is most highly correlated with regionalism. Nearly two-thirds of the variance in Democratic voting among nonmetropolitan areas is associated

with the regional residence of the voter. This is not particularly surprising in view of the well-known political subdivisions of the country often referred to by such expressions as "Solid South," "New England Republican," etc. The effects of controls on this regional political structure for metropolitan areas are rather varied, as we have seen. The question remains, however, of what effects different controls have for nonmetropolitan areas. Summary data bearing on this question are presented in Table 42, which contains intraclass correlations and variance components for each control group.

It may be observed that the values of rho in this table resemble in pattern the correlations discussed in the foregoing section on voter turnout. Eight of the ten control groups exhibit correlations which fall below that for the total uncontrolled group of nonmetropolitan areas. The only two groups which fail to show this are the highly agricultural populations with no additional controls (N-5 and N-10). This is in line with the expectation that the essence of regionalism is apt to be found in the more isolated, rural areas. As might be anticipated, the greatest reductions in the value of rho occur when controls are exercised on the nonwhite proportion of the population (N-4, N-9, and N-10). This, of course, suggests that it is the Southern regions to which a good deal of the high, uncontrolled correlation can be attributed. Evidence in support of this idea is provided by an intraclass correlation which excludes the two Southern regions. The value of this measure is .126 when all nonmetropolitan areas in the four non-Southern regions are included. In other words, over 80 per cent of the high, uncontrolled correlation (viz., .652) for all

TABLE 42

A LISTING FOR PROPORTION OF VOTERS VOTING DEMOCRATIC OF THE VALUES OF RHO, THE INTERREGIONAL VARIANCE, THE INTRAREGIONAL VARIANCE, AND N FOR THE TOTAL GROUP OF NONMETROPOLITAN STATE ECONOMIC AREAS AND TEN SELECTED CONTROL GROUPS

Listed Items	Total Nonmet. Areas	Control Groups*									
		N-1	N-2	N-3	N-4	N-5	N-6	N-7	N-8	N-9	N-10
Rho	.652	.633	.578	.572	.502	.690	.769	.545	.572	.526	.435
Interregional Variance.	236.3	161.9	104.1	83.9	69.7	251.2	301.2	108.2	117.9	108.4	55.4
Intraregional Variance.	122.4	85.8	67.3	54.2	60.2	106.6	85.1	77.0	77.2	88.7	66.7
N	294	109	85	73	88	158	127	70	79	118	175

*The specific controls used are as follow:

N-1: Agricultural employment = 10-34.9%.
N-2: " " " " , Wholesale and retail trade employment = 13%+.
N-3: " " " " , Farm level living = 85-124.
N-4: " " " " , Nonwhite < 15%.

N-5: Agricultural employment = 30-54.9%.
N-6: " " " " , Nonfarm occupation < 25%.
N-7: " " " " , Unpaid family workers < 10%.
N-8: " " " " , Farm level living = 90-124.
N-9: " " " " , Nonwhite < 20%.
N-10: Nonwhite < 5%.

regions can be attributed to the distinctiveness of the Southeast and Southwest.

Turning to the variance components which provide the bases for the intraclass correlation ratios, it can be seen in Table 42 that all control groups except two (N-5 and N-6) have an interregional variance which is between 32 and 76 per cent smaller than that for all nonmetropolitan areas. The two exceptional control groups (N-5 and N-6) display larger interregional variances which, of course, contribute substantially to the larger rho values shown for these groups. With respect to the average internal variation of the regions, every control group without exception has an intraregional variance which is from 13 to 56 per cent smaller than that exhibited by the total group of nonmetropolitan areas. Generally speaking, the less agricultural, subgroup populations (N-1 through N-4) have smaller intraregional variance values than the more agricultural subgroups (N-5 through N-9). This pattern parallels that found for the voter turnout variable (see Table 38).

To understand the reasons for the fluctuating values of these variance components, data on regional means and standard deviations are needed. In Table 43, regional means for each control group are presented. These data show relative stability for every region except the Southeast. There are some indications in the Northwest of greater Democratic voting among the less agricultural population subgroups (N-1 through N-4), while in the Southeast it is the more agricultural groups (N-5 through N-8) which show greater preference for the Democrats. There are no real indications of either of these voter preference patterns in the other regions. It appears rather obvious that the major effect of controls is felt

TABLE 43

DISTRIBUTIONS OF REGIONAL MEANS FOR PROPORTION OF VOTERS VOTING DEMOCRATIC BY THE TOTAL GROUP OF NONMETROPOLITAN STATE ECONOMIC AREAS AND TEN SELECTED CONTROL GROUPS

Regions	Total Nonmet. Areas	Control Groups*									
		N-1	N-2	N-3	N-4	N-5	N-6	N-7	N-8	N-9	N-10
Northeast	44.2%	43.7%	43.3%	44.5%	43.2%	44.3%	36.8%	40.9%	42.2%	42.1%	43.9%
Southeast	79.1	75.4	71.9	67.1	62.6	80.2	83.0	74.2	73.1	68.0	56.6
Middle States	45.7	44.5	44.4	46.8	44.5	44.6	43.4	45.1	46.6	44.6	45.5
Northwest	46.4	53.8	53.6	53.8	53.8	45.5	45.5	47.6	45.7	45.5	46.2
Southwest	70.1	69.1	69.8	71.4	71.0	70.3	70.2	65.9	71.6	68.7	73.0
Far West.	55.9	55.1	55.1	55.4	55.1	53.5	54.2	53.5	54.0	53.5	56.4

*The specific controls used are as follow:

N-1: Agricultural employment = 10-34.9%.
N-2: " " " " , Wholesale and retail trade employment = 13%+.
N-3: " " " " , Farm level living = 85-124.
N-4: " " " " , Nonwhite < 15%.

N-5: Agricultural employment = 30-54.9%.
N-6: " " " " , Nonfarm occupation < 25%.
N-7: " " " " , Unpaid family workers < 10%.
N-8: " " " " , Farm level living = 90-124.
N-9: " " " " , Nonwhite < 20%.
N-10: Nonwhite < 5%.

in the Southeast and that the most sensitive control here is that for the nonwhite proportion of the population. Furthermore, it can be seen that the only two control groups which do not exhibit a reduction in their interregional variances (N-5 and N-6) have quite high regional mean values in the Southeast. When it is recalled that these are also the only two control groups which have a proportionately large representation of nonmetropolitan areas in the Southeast (see Table 13, Appendix B), it becomes obvious that this combination of heavy weighting and extreme mean values will produce a relatively large interregional variance. The consequence for rho, of course, is to inflate its value—which is precisely what happens (see Table 42).

With respect to the internal variation of the regions (which on the average exhibit reductions for every control group), opportunity for further analysis is provided by data on standard deviations which appear in Table 44. Considering all regions and all control groups, an accounting reveals that 75 per cent of the standard deviations are smaller than those of the total parental groups. While some of these reductions are not very large, over two-thirds of them show differences which exceed 10 per cent of the uncontrolled value. The regions which display the least frequent and the smallest percentage reductions are the Middle States and the Southwest. It is also clear that in the Northeast the more agricultural population subgroups (N-5 through N-9) are not so homogeneous as the less agricultural populations (N-1 through N-4). In the Northwest and Far West the very opposite pattern tends to prevail. The one control group in which all regions display reductions in sigma value is N-3, where a control for farm

TABLE 44

DISTRIBUTIONS OF REGIONAL STANDARD DEVIATIONS FOR PROPORTION OF VOTERS VOTING DEMOCRATIC BY THE TOTAL GROUP OF NONMETROPOLITAN STATE ECONOMIC AREAS AND TEN SELECTED CONTROL GROUPS

Regions	Total Nonmet. Areas	Control Groups*									
		N-1	N-2	N-3	N-4	N-5	N-6	N-7	N-8	N-9	N-10
Northeast	11.5%	8.4%	8.3%	7.9%	8.2%	11.4%	10.9%	11.7%	11.6%	11.6%	8.9%
Southeast	14.1	13.4	10.6	7.0	9.4	12.7	10.3	9.2	9.1	13.8	9.6
Middle States	6.6	6.6	6.8	4.7	6.7	6.4	5.9	6.4	5.8	6.5	6.6
Northwest	8.6	8.1	9.2	8.1	8.1	7.1	7.1	7.6	7.0	7.1	8.6
Southwest	13.2	11.9	12.3	11.8	12.0	14.2	13.7	14.0	15.5	13.7	12.9
Far West.	5.8	5.2	5.2	5.5	5.2	4.3	3.5	4.3	4.4	4.3	6.0
Total U.S.A..	18.9%	15.8%	13.0%	11.7%	11.5%	18.9%	19.7%	13.6%	14.1%	14.1%	11.1%

*The specific controls used are as follow:

N-1: Agricultural employment = 10-34.9%.
N-2: " " " " , Wholesale and retail trade employment = 13%+.
N-3: " " " " , Farm level living = 85-124.
N-4: " " " " , Nonwhite < 15%.

N-5: Agricultural employment = 30-54.9%.
N-6: " " " " , Nonfarm occupation < 25%.
N-7: " " " " , Unpaid family workers < 10%.
N-8: " " " " , Farm level living = 90-124.
N-9: " " " " , Nonwhite < 20%.
N-10: Nonwhite < 5%.

standard of living is applied to the less agricultural base control. It may also be observed that the reductions in sigma value for this control group are the largest that appear anywhere in the table for the Northeast, the Southeast, the Middle States, and the Southwest regions. In view of this it is not surprising that this control group (N-3) exhibits the smallest intraregional variance of all control groups (see Table 42).

As has been seen throughout this study, the relationship between the two political variables and regionalism is a rather strong one. Most of this relationship, however, is attributable to the deviant positions of the Southern regions, especially the Southeast. This is true regardless of whether one focuses on metropolitan or nonmetropolitan areas. It is also true that the most strategic factor presumably influencing political behavior in the Southeast is the incidence of nonwhite members of the population. This has been demonstrated for voter turnout in metropolitan and nonmetropolitan areas and for voter preference in metropolitan areas only. To complete the pattern of this analysis, data have been gathered on the relationship between voter preference and the nonwhite population proportion in nonmetropolitan areas of the Southeast. These data appear in Table 45.

Inspection of these mean values shows the pattern to be clear and consistent with the other investigations referred to above. Democratic voting and the proportion of nonwhite persons in Southeast population groups are variables which are directly related to one another. The intraclass correlation for this distribution is .532 which represents considerable improvement in predictive accuracy. It may be concluded, therefore, that in the

population groups of the Southeast, the greater the proportion of Negroes, the less likely the general population will vote, but the more likely the actual voters will vote Democratic. This holds true for both metropolitan and nonmetropolitan populations, although the pattern appears a bit stronger for the former. This suggests, of course, that in the arsenal of techniques used by Southern whites to control the subjugated, nonwhite, minority group, politics may assume rather strategic importance.

TABLE 45

MEAN VALUES AND STANDARD DEVIATIONS FOR PROPORTION OF VOTERS VOTING DEMOCRATIC IN THE NONMETROPOLITAN STATE ECONOMIC AREAS OF THE SOUTHEAST REGION BY THE PROPORTION OF THE TOTAL POPULATION WHICH IS NONWHITE

Proportion Nonwhite	Mean	Standard Deviation	N
40%+	89.3%	8.0%	23
30-39.	84.6	8.8	25
20-29.	83.2	8.5	12
10-19.	76.5	11.7	13
Under 10	60.4	10.6	21
Total Region.	79.1%	14.1%	94

To summarize generally the results of this analysis of voter preference, it can be said that the regional distinctiveness of this type of behavior rests primarily on the peculiarities of the two Southern regions. Among the other four regions differences in Democratic voting are not very large, nor are they much affected by controls. Furthermore, none of the controls applied in this

study seems to affect the Southwest's voting pattern very much. In the Southeast, however, controls on the standard of living and especially on the nonwhite proportion of the population seem to effect substantial shifts in the voting pattern which noticeably reduce this region's political distinctiveness.

Educational Achievement

It will be recalled from the analysis of metropolitan and nonmetropolitan areas in Chapter III that these two types of population groups differ greatly with respect to their correlations of educational achievement and regionalism. In fact, the intraclass correlation of .194 for metropolitan areas is less than one-half that for nonmetropolitan areas, which is .439. The over-all effects of controls on this high, nonmetropolitan correlation are revealed in Table 46, which contains intraclass correlations and variance components for each control group.

An inspection of the rho values in this table shows eight of the control groups with smaller values than that for the entire group of nonmetropolitan areas. None of these reductions is very large, although most of them do exceed 10 per cent. There do not seem to be any real patterns emerging from these data. The base controls taken alone (N-1 and N-5) are not at all successful in reducing the correlation. It is only when these are supplemented by additional controls that any noticeable reductions occur.

While it might be inferred that the modest changes in rho under controlled conditions indicate only moderate educational differences among the various population subgroups, an inspection of the variance components reveals some very substantial changes.

TABLE 46

A LISTING FOR EDUCATION OF THE VALUES OF RHO, THE INTERREGIONAL VARIANCE, THE INTRAREGIONAL VARIANCE, AND N FOR THE TOTAL GROUP OF NONMETROPOLITAN STATE ECONOMIC AREAS AND TEN SELECTED CONTROL GROUPS

Listed Items	Total Nonmet. Areas	Control Groups*									
		N-1	N-2	N-3	N-4	N-5	N-6	N-7	N-8	N-9	N-10
Rho	.439	.493	.379	.380	.361	.447	.427	.398	.370	.343	.386
Interregional Variance.	7.4	4.6	2.5	1.8	1.6	8.6	9.2	4.2	2.0	3.8	2.7
Intraregional Variance.	9.0	4.2	3.4	2.4	2.4	9.8	11.1	5.2	2.7	6.3	3.9
N	294	109	85	73	88	158	127	70	79	118	175

*The specific controls used are as follow:

N-1: Agricultural employment = 10-34.9%.
N-2: " " " " , Wholesale and retail trade employment = 13%+.
N-3: " " " " , Farm level living = 85-124.
N-4: " " " " , Nonwhite < 15%.

N-5: Agricultural employment = 30-54.9%.
N-6: " " " " , Nonfarm occupation < 25%.
N-7: " " " " , Unpaid family workers <10%.
N-8: " " " " , Farm level living = 90-124.
N-9: " " " " , Nonwhite < 20%.
N-10: Nonwhite <5%.

Eight of the control groups have interregional variances which are from 38 to 78 per cent smaller than that for all nonmetropolitan areas. These reductions tend to be greatest among the less agricultural populations (N-1 through N-4), although a noticeably large reduction occurs when a control for living standard is applied to the more agricultural population group (N-8).

Since the value of rho exhibits only moderate changes despite the substantial shifts in the value of the interregional variance, one must assume that the intraregional variance also undergoes considerable alteration when controls are applied. Inspection of these values reveals this to be true. Eight of the ten control groups show intraregional variances which fall between 30 and 74 per cent below the corresponding value for all nonmetropolitan areas. In other words, the failure of rho to exhibit more than small changes in value can be attributed to the compensating nature of the substantial changes in the two components of variance.

The convergence of regional means suggested by the reductions in interregional variance values may be observed from the data presented in Table 47. It becomes quickly apparent that the deviant regions for this education characteristic are the Southern ones. This holds true for the entire table. The importance of these deviant regions for the overall correlation of .439 can be demonstrated by the fact that the corresponding intraclass correlation for the four non-Southern regions is .185. It may also be observed that among the four control groups whose populations are less agriculturally oriented (N-1 through N-4) the convergence of regional means for the non-Southern regions is rather marked. Indeed, the intraclass correlations of the non-Southern regions for

TABLE 47

DISTRIBUTIONS OF REGIONAL MEANS FOR EDUCATION BY THE TOTAL GROUP OF NONMETROPOLITAN STATE ECONOMIC AREAS AND TEN SELECTED CONTROL GROUPS

Regions	Total Nonmet. Areas	Control Groups*									
		N-1	N-2	N-3	N-4	N-5	N-6	N-7	N-8	N-9	N-10
Northeast	3.2%	2.6%	2.7%	2.7%	2.6%	4.1%	3.6%	3.6%	4.2%	3.6%	2.9%
Southeast	7.8	6.5	5.6	5.1	5.8	8.2	8.4	7.5	6.0	5.2	6.0
Middle States	1.8	1.8	1.8	2.3	1.8	1.6	1.5	1.4	1.9	1.6	1.7
Northwest	1.5	2.0	2.1	2.0	2.0	1.4	1.4	1.5	1.4	1.4	1.5
Southwest	6.7	7.8	6.7	6.4	5.4	6.4	6.0	5.6	3.6	6.5	7.2
Far West.	2.4	2.6	2.6	2.5	2.6	2.8	2.3	2.8	3.1	2.8	2.1

*The specific controls used are as follow:

N-1: Agricultural employment = 10-34.9%.
N-2: " " " " , Wholesale and retail trade employment = 13%+.
N-3: " " " " , Farm level living = 85-124.
N-4: " " " " , Nonwhite < 15%.

N-5: Agricultural employment = 30-54.9%.
N-6: " " " " , Nonfarm occupation < 25%.
N-7: " " " " , Unpaid family workers < 10%.
N-8: " " " " , Farm level living = 90-124.
N-9: " " " " , Nonwhite < 20%.
N-10: Nonwhite < 5%.

these control groups are (listed in order) .011, .013, -.027, and .058. These, of course, represent reductions from the uncontrolled correlation of .185 referred to above for these regions, and are values which fall far below the five per cent level of statistical significance. This means that the correlations reported for all regions in these control groups (see Table 46) are almost entirely due to the deviant positions of the Southern regions.

There is no regular pattern of control effects which is common to all regions. In the Northeast and Southeast the more agricultural populations (N-5 through N-8) tend to have larger proportions of adult male illiteracy than the less agricultural groups (N-1 through N-4). On the other hand, the opposite pattern appears in the Northwest and to some extent in the Middle States and Southwest. No pattern at all seems to obtain in the Far West. Understandably, it is the regions with the highest mean values which change the most under applied controls: viz., the Northeast and the two Southern regions. Since most of these changes take the form of reductions in mean value, one would expect some reduction in the total, average, interregional variance. When this tendency is supplemented by the great reductions in the weighting factor of the high value Southeast (brought about by control selection; see Table 13 in Appendix B), it becomes obvious that major, downward shifts in the interregional variance must occur. And this is precisely what happens, as we have seen in the data reported in Table 46.

Turning to the data on regional standard deviations which are presented in Table 48, it can be seen that there are great differences among the various regions with respect to their

TABLE 48

DISTRIBUTIONS OF REGIONAL STANDARD DEVIATIONS FOR EDUCATION BY THE TOTAL GROUP OF NONMETROPOLITAN STATE ECONOMIC AREAS AND TEN SELECTED CONTROL GROUPS

Regions	Total Nonmet. Areas	Control Groups*									
		N-1	N-2	N-3	N-4	N-5	N-6	N-7	N-8	N-9	N-10
Northeast	1.7%	1.1%	1.0%	1.1%	1.1%	2.3%	2.8%	2.1%	2.5%	2.0%	1.5%
Southeast	4.1	2.1	1.8	1.0	1.3	4.1	4.6	1.5	2.2	1.6	1.6
Middle States	1.2	1.4	1.4	1.7	1.3	1.0	.9	.7	.7	1.0	1.1
Northwest	.9	1.0	1.1	1.0	1.0	.7	.7	.8	.7	.7	.9
Southwest	5.1	5.1	4.2	3.2	3.7	5.1	5.2	5.3	2.9	5.6	6.3
Far West.	1.4	1.4	1.4	1.4	1.4	1.4	1.3	1.4	1.4	1.4	1.4
Total U.S.A..	4.1%	3.0%	2.4%	2.0%	2.0%	4.3%	4.5%	3.1%	2.2%	3.2%	2.5%

*The specific controls used are as follow:

N-1: Agricultural employment = 10-34.9%.
N-2: " " " " , Wholesale and retail trade employment = 13%+.
N-3: " " " " , Farm level living = 85-124.
N-4: " " " " , Nonwhite < 15%.

N-5: Agricultural employment = 30-54.9%.
N-6: " " " " , Nonfarm occupation < 25%.
N-7: " " " " , Unpaid family workers < 10%.
N-8: " " " " , Farm level living = 90-124.
N-9: " " " " , Nonwhite < 20%.
N-10: Nonwhite < 5%.

reactions to controls. The Far West and Northwest show very little change in any control group, while the other regions (especially the Southeast) have their sigma values noticeably altered by various controls. For the table as a whole, 48 per cent of the control groups' standard deviations are smaller than those for the total, uncontrolled groups. However, no particular control group seems generally prominent in this regard. In the Southwest the greatest reductions appear when standard of living is controlled (N-3 and N-8). In the Northeast the less agricultural, population control groups (N-1 through N-4) seem to be the most homogeneous, while in the Southeast all but two of the control groups (N-5 and N-6) exhibit marked reductions in sigma value.

It is the findings here for the Southeast which are particularly interesting in view of the strategic importance of this region for the over-all correlation of education and regionalism. The small sigma values of the Southeast under controlled conditions lend confidence to the mean values reported for this region. There can be little doubt that the Southeast is a genuinely distinctive area as far as nonmetropolitan, adult, male illiteracy is concerned. By the same token, however, the small standard deviations displayed by the less agricultural populations of the four non-Southern regions (N-1 through N-4) lend confidence to the findings reported before, which show negligible differences in adult, male illiteracy.

To summarize, there is a fairly high correlation between regionalism and the incidence of adult, male illiteracy among nonmetropolitan populations. Most of this correlation can be attributed to the distinctively high illiteracy rates which obtain in

the Southern regions. These Southern regional differences are moderately associated with some structural and level of living characteristics. However, none of the controls applied in this study are sufficient to render the Southern regions indistinguishable from the rest of the country.

On the other hand, some controls are quite successful in leveling differences among the non-Southern regions. This is especially evident where controls are selective of the less agriculturally oriented populations (N-1 through N-4). Negligible educational differences among non-Southern regions are found for these population groups.

Religious Affiliation

As in the case of the education variable just discussed, religious affiliation is a type of behavior which is quite different for metropolitan and nonmetropolitan areas in its degree of association with regionalism. However, where education shows a relatively high degree of association in nonmetropolitan areas, religious affiliation shows a relatively low degree of association for these population groups. Summary data showing the effects of controls on this low correlation between regionalism and religious affiliation are presented in Table 49.

It may be surprising to observe that only three control groups (N-5, N-7, and N-8) have intraclass correlations which are smaller than that for all nonmetropolitan areas. It should also be noted that everyone of the more rural agricultural, control group populations (N-5 through N-8) has a rho value falling considerably below those exhibited by the population groups with a

TABLE 49

A LISTING FOR RELIGIOUS AFFILIATION OF THE VALUES OF RHO, THE INTERREGIONAL VARIANCE, THE INTRAREGIONAL VARIANCE, AND N FOR THE TOTAL GROUP OF NONMETROPOLITAN STATE ECONOMIC AREAS AND TEN SELECTED CONTROL GROUPS

Listed Items	Total Nonmet. Areas	Control Groups*									
		N-1	N-2	N-3	N-4	N-5	N-6	N-7	N-8	N-9	N-10
Rho	.133	.193	.202	.218	.242	.103	.159	.100	.083	.144	.282
Interregional Variance.	22.4	28.7	31.1	34.8	39.7	18.5	26.9	27.0	23.1	26.9	52.3
Intraregional Variance.	126.3	92.2	89.4	88.5	95.4	116.0	108.4	125.5	128.2	116.5	118.5
N	294	109	85	73	88	158	127	70	79	118	175

*The specific controls used are as follow:

N-1: Agricultural employment = 10-34.9%.
N-2: " " " " , Wholesale and retail trade employment = 13%+.
N-3: " " " " , Farm level living = 85-124.
N-4: " " " " , Nonwhite < 15%.

N-5: Agricultural employment = 30-54.9%.
N-6: " " " " , Nonfarm occupation < 25%.
N-7: " " " " , Unpaid family workers < 10%.
N-8: " " " " , Farm level living = 90-124.
N-9: " " " " , Nonwhite < 20%.
N-10: Nonwhite < 5%.

lower degree of agricultural involvement. Indeed, with one exception (N-6) the correlations for the more agricultural groups are about one-half the size of the correlations for the less agricultural groups. This is certainly not the pattern one would probably predict from hypotheses about the growth and development of regionalism. Furthermore, the improbability of this is increased by the fact that two of the control groups with reduced correlations (N-7 and N-8) display values which fall below the five per cent level of statistical significance (see Table 34).

The immediate reasons for these unusual results may be found in the data on the variance components presented in Table 49. Here it can be seen that all of the more agricultural control groups have smaller interregional variances and larger intraregional variances than the less agricultural groups. With this pattern of differences in variance components between the two base control groupings, the observed pattern of rho values is understandable. There are some other features of this table, however, which require further investigation.

It will be noted that <u>all but one</u> control group (N-8) have intraregional variances which are smaller than the total, uncontrolled, internal variance. At the same time, <u>only one</u> control group (N-5) displays a reduced interregional variance. When these findings are related to the performance of rho discussed above, a problem arises—namely, how can rho be reduced for <u>three</u> control groups? This problem is valid as an inspection of the data will reveal. The reduction of rho for control group N-5 appears justified because there is a greater decrease in the interregional variance than in the intraregional variance. By the same reasoning,

the reduction of rho for control group N-8 also appears justified because the increase in the interregional variance is accompanied by an even larger increase in the value of the intraregional variance. However, it is the reduction of rho for control group N-7 which seems to pose a contradiction. Here we observe an increase in the interregional variance being accompanied by a decrease in the intraregional variance. In other words, these control conditions show greater heterogeneity among the regions and greater homogeneity within the regions. How, then, is it possible for rho to be decreased?

The answer here is both simple and depressing. The reduction of rho in control group N-7 is a statistical artifact arising out of the relatively small size of the sample N. It will be recalled from the discussion of control procedure at the beginning of Chapter IV that the importance of the corrective factor in the formula for rho is directly related to the size of the sample. Since this corrective factor is a negative term, it reduces the value of rho. In short, what we have here is a situation in which the reduction in the sample size is so great that the corrective factor for chance fluctuation actually reduces the correlation, despite changes in the variance components of the basic rho ratio which tend to increase the correlation. It should perhaps be pointed out here that the reductive effect of the corrective factor on the value of rho for small samples becomes particularly acute when the correlation is small to begin with. This, of course, is the situation which obtains for religious affiliation in nonmetropolitan areas. It is also the kind of situation which produces the negative values of rho which we have found before for

the fertility variable.

To allay some of the uncertainties about what happens to religious affiliation when controls are applied, an investigation of regional means and standard deviations seems to be called for. Data on regional means are presented in Table 50. The most obvious feature of these data at first glance is the consistently low values for religious affiliation in the Far West. Since the differences among the other five regions are relatively small, one may infer with some confidence that a good deal of the correlation for all regions is a product of the deviant behavior of Far West residents. This inference is supported by evidence in the form of an intraclass correlation of .038 which appears when the Far West region is excluded. This means that over 70 per cent of the total nonmetropolitan correlation for all regions (viz., .133) is contributed by the Far West.

With respect to the specific effects of controls on religious affiliation, each region shows some response, but the changes are not very large. Ninety per cent of the control group mean values do not differ more than 15 per cent from their corresponding total regional means. There appears to be some tendency for the Northwest and to a less extent for the Far West to have a consistently greater degree of affiliation in the more agricultural areas (N-5 through N-8), while in the Southwest the opposite tendency appears—that is, greater affiliation in the less agricultural areas (N-1 through N-4). For the other regions, however, no such patterns seem to emerge. There is one interesting feature in the mean values of the Southeast where controls for nonwhite residents seem to depress the incidence of religious affiliation (N-4, N-9,

TABLE 50

DISTRIBUTIONS OF REGIONAL MEANS FOR RELIGIOUS AFFILIATION BY THE TOTAL GROUP OF NONMETROPOLITAN STATE ECONOMIC AREAS AND TEN SELECTED CONTROL GROUPS

Regions	Total Nonmet. Areas	Control Groups*									
		N-1	N-2	N-3	N-4	N-5	N-6	N-7	N-8	N-9	N-10
Northeast	42.3%	41.8%	40.5%	41.7%	42.5%	40.6%	48.6%	41.1%	44.1%	39.3%	44.8%
Southeast	36.6	36.2	36.5	38.1	30.7	38.7	39.2	41.3	37.8	32.4	23.3
Middle States	41.7	40.2	40.1	36.3	40.4	41.8	43.6	39.5	39.8	41.8	42.1
Northwest	40.0	33.4	38.1	33.4	33.4	41.7	41.7	42.4	41.0	41.7	40.4
Southwest	34.8	40.0	40.8	39.6	37.3	33.1	32.2	35.0	33.2	33.5	37.3
Far West.	23.9	25.0	25.0	24.3	25.0	27.0	26.6	27.0	27.3	27.0	22.8

*The specific controls used are as follow:

N-1: Agricultural employment = 10-34.9%.
N-2: " " " " , Wholesale and retail trade employment = 13%+.
N-3: " " " " , Farm level living = 85-124.
N-4: " " " " , Nonwhite < 15%.

N-5: Agricultural employment = 30-54.9%.
N-6: " " " " , Nonfarm occupation < 25%.
N-7: " " " " , Unpaid family workers < 10%.
N-8: " " " " , Farm level living = 90-124.
N-9: " " " " , Nonwhite < 20%.
N-10: Nonwhite < 5%.

and N-10). This is particularly interesting because, for these control groups, one of the sources of a downward bias in the data due to under-reporting is minimized. Reference here is to the fact that these control groups are selective of areas with relatively few Negroes. According to the Bureau of the Census, certain Negro Baptist groups were conspicuous in their failure to report information for the 1936 Census of Religious Bodies. In other words, the low values of religious affiliation for these control groups cannot be attributed to data which are poor because of a Negro nonreporting bias. And, as it will be seen shortly, two of these control groups (N-9 and N-10) have rather small standard deviations. This indicates that their mean values are fairly trustworthy.

Turning to the data on standard deviations in Table 51, some very marked changes in value appear when controls are applied. For the table as a whole, 77 per cent of the control groups' standard deviations are smaller than the corresponding sigma values for all nonmetropolitan areas. The frequency of these sigma reductions is high for every region except the Northwest. Here, most of the changes take the form of rather small increases in value. It should be noted here that the standard deviations for the Northwest are in every case the highest values observed for all regions. Reference to Table 13 in Appendix B will also show that the Northwest is under-represented (proportionately speaking) in the less agricultural control groups (N-1 through N-4) and over-represented in the more agricultural groups (N-5 through N-9). This differential weighting of this highly heterogeneous region is probably an important factor contributing to the results reported for the total intraregional variance in Table 49—namely, the

TABLE 51

DISTRIBUTIONS OF REGIONAL STANDARD DEVIATIONS FOR RELIGIOUS AFFILIATION BY THE TOTAL GROUP OF NONMETROPOLITAN STATE ECONOMIC AREAS AND TEN SELECTED CONTROL GROUPS

Regions	Total Nonmet. Areas	Control Groups*									
		N-1	N-2	N-3	N-4	N-5	N-6	N-7	N-8	N-9	N-10
Northeast	13.4%	9.5%	9.3%	9.9%	9.1%	14.0%	8.7%	11.3%	13.3%	11.8%	11.1%
Southeast	10.2	8.2	5.7	4.9	10.0	9.2	8.6	9.8	7.0	6.4	5.6
Middle States	10.1	8.9	9.2	7.0	9.0	10.0	10.3	8.3	10.1	10.0	9.9
Northwest	14.4	15.9	15.0	15.9	15.9	14.2	14.2	17.4	14.9	14.2	15.0
Southwest	11.1	11.3	11.7	9.5	10.1	10.9	9.5	10.4	8.4	11.4	8.4
Far West	7.1	6.4	6.4	6.7	6.4	6.4	6.8	6.4	6.9	6.4	7.1
Total U.S.A.	12.2%	10.9%	10.9%	11.1%	11.6%	11.5%	11.7%	12.4%	12.4%	11.9%	13.1%

*The specific controls used are as follow:

N-1: Agricultural employment = 10-34.9%.
N-2: " " " " , Wholesale and retail trade employment = 13%+.
N-3: " " " " , Farm level living = 85-124.
N-4: " " " " , Nonwhite < 15%.

N-5: Agricultural employment = 30-54.9%.
N-6: " " " " , Nonfarm occupation < 25%.
N-7: " " " " , Unpaid family workers < 10%.
N-8: " " " " , Farm level living = 90-124.
N-9: " " " " , Nonwhite < 20%.
N-10: Nonwhite < 5%.

difference between the more and less agricultural base control groups.

The regions exhibiting the largest and most frequent reductions in sigma value are the Northeast and Southeast, followed by the Middle States. Ninety per cent of the measures for these regions take the form of reductions which, in the Northeast and Middle States, appear most prominent for the less agricultural areas. In the Southeast the pattern is a little erratic. Generally speaking, every control group succeeds in reducing at least somewhat the internal heterogeneity of most of the regions.

To summarize briefly, most of the controls applied to non-metropolitan areas do not serve to reduce further the rather small correlation which exists between regionalism and religious affiliation. Most of this correlation can be attributed to the deviant position of the Far West, where there is a very low degree of religious affiliation. The general failure of controls to effect reductions in correlation can be attributed primarily to the fact that controls tend to render the regions more homogeneous internally without an accompanying, commensurate convergence of regional mean values. There is, however, some tendency for convergence in this regard.

With respect to the effects of controls in specific regions, there are indications that in the Northwest and (to some extent) the Far West religious affiliation is higher among the more agricultural segments of the population. In the Southwest it is the less agricultural segment of the population which has higher degrees of religious affiliation. In the other three regions there does not seem to be such consistency of pattern for these two base

control groups.

In general, the similarity among regions in religious affiliation seems to be greater for the more agriculturally oriented populations. However, within regions similarity in religious affiliation appears greater for the less agriculturally oriented populations. This, of course, indicates that the correlation between regionalism and religious affiliation is relatively low for the more agricultural segments of regional populations. This finding tends to contradict expectations based on the idea that regional distinctiveness tends to be developed and maintained in relatively isolated, rural type situations.

Summary and Evaluation

This analysis has been focused on the effects of controls for regional differences in behavior among nonmetropolitan populations. The form of this analysis parallels that followed for metropolitan areas in Chapter IV. The hypothetical expectation prompting this investigation is that controls on regional differences in social structure and socio-economic characteristics will be accompanied by reductions in behavioral differences among regional nonmetropolitan populations. The measures of behavior used in this analysis are as follow: the intraclass correlation, the interregional variance, the intraregional variance, regional means, and regional standard deviations.

It is believed that when controls are applied, there will be a leveling of differences in behavior both within and among the several regions. This leveling of behavioral differences should be reflected by a reduction in the value of the following measures:

the intraclass correlation, the interregional variance, the intraregional variance, and the regional standard deviations.

After an examination of the distributions of twelve characteristics within each region, ten control groups were finally set up. These groups, which are really samples of the regional populations, were selected on the basis of six characteristics, some of which are used in conjunction with one another. Since the entire analysis here is focused on nonmetropolitan areas--a rather broad, residual category--two base control groups were set up which are distinguished from one another by the degrees of agricultural involvement exhibited by their populations. The results of this analysis, which takes the form of comparisons between control group measures and the measures of the total or uncontrolled group of nonmetropolitan areas, may be summarized in inventory form.

With respect to the intraclass correlation, 68 per cent of the control groups have values falling below those for all nonmetropolitan areas. This proportion of reductions in value varies somewhat with the different types of behavior which are correlated with regionalism. Thus, 70 per cent of the control groups show reductions for fertility, 80 per cent for voters voting, 80 per cent for Democratic voting, 80 per cent for education, and 30 per cent for religious affiliation. It can be seen, therefore, that except for religious affiliation the frequency with which controls succeed in reducing the correlation between behavior and regionalism is rather high. Although some of these reductions are small, about 80 per cent of them represent differences which fall at least ten per cent below the values exhibited for all nonmetropolitan

areas. The control groups most successful in this regard are the more agricultural populations with additional controls on unpaid farm family workers (N-7) and on level of living standards (N-8). With respect to statistical significance, six control groups have correlation values which fall below the five per cent level. Four of these involve fertility and two of them involve religious affiliation. However, one of these low value measures for religious affiliation is suspect because of a statistical artifact arising out of a small sample size.

Turning to the two variance components, an accounting reveals that about two-thirds of the control groups have smaller interregional variance values than those for all nonmetropolitan areas. Only one of these reductions appears for religious affiliation, which means that for the other four types of behavior 78 per cent of the control groups exhibit reduced interregional variances. In terms of the intraregional variance, the reductions in value under controlled conditions are almost 100 per cent complete. To be precise, 94 per cent of the control groups exhibit reduced internal variations for the regions. All of the interregional reductions and 90 per cent of the intraregional reductions fall at least 10 per cent below the variance values for all nonmetropolitan areas. It may also be reported that over 90 per cent of the reductions in rho value are achieved by reductions in both components of variance. This, of course, conforms to the ideal pattern hypothetically expected. It is also interesting to note that the control group which tends to be least successful in reducing the interregional variance is the most agriculturally oriented population in the entire study (N-6). This tends to support the idea that

regionalism is born and nurtured under conditions of rural isolation. The fact, however, should not be overlooked that this rather extremely rural segment of the population contains substantial differences in level of living standards among regions. When these living standards are controlled, reductions in interregional variance appear for most of the dependent variables.

Although a very high proportion of control groups display reductions in the value of the total, average, intraregional variance, the reactions of the specific regions to controls cannot safely be inferred from such summary data. Accordingly, an analysis of regional standard deviations is required. The over-all results of this show two-thirds of the regional sigma values falling below the corresponding values for all nonmetropolitan areas. This reduction in the internal variation of the individual regions varies both with the specific variables measured and the specific regions. For fertility, the frequency of reduction under controls is 65 per cent; for voters voting, 65 per cent; for Democratic voting, 75 per cent; for education, 48 per cent; and for religious affiliation, 77 per cent. A rank ordering of the regions with respect to the frequency of sigma reductions when controls are applied shows the Southeast at the top with 86 per cent reductions (all variables considered) followed by the Northeast with 74 per cent, the Middle States with 70 per cent, the Far West with 60 per cent, the Northwest with 54 per cent, and the Southwest with 52 per cent. These patterns of reduction both for the variables and the regions differ somewhat from the patterns observed in metropolitan areas.

It is difficult to specify the control groups which are

most effective in reducing the internal variation of the several regions because when all dependent variables are considered, the number of regions exhibiting such reductions are about the same for each control group. Thus, for example, the least effective control groups have 60 per cent of their sigma values reduced, while the most effective control groups have 77 per cent of their sigmas reduced. The differences in frequency here do not seem to be very large. However, when the criterion of effectiveness includes not only the frequency but also the magnitude of the reductions, it appears that the two control groups involving a control on level of living (N-3 and N-8) together with the wholesale and retail trade control group (N-2) are the most effective controls used in this analysis. By the same token the more rural base control group (N-5) and the more rural group with an additional control on nonwhite residents (N-9)—these two control groups appear to be the least effective in reducing internal variation.

The relationship of the findings summarized here to the hypothesis guiding this research is positive in the sense that the results tend to support hypothetical expectations for most of the variables. This, of course, is not to imply that there are no ambiguities or contradictions. These do appear, particularly for the religious behavior variable. Furthermore, with a few exceptions none of the controls for structure and living standard serves to level regional differences in behavior completely. With these reservations noted, however, the fact remains that the evidence tends more to favor the hypothesis than to contradict it. For most of the dependent variables, regional differences are reduced by most of the control groups. This holds true regardless of whether

one examines the intraclass correlations, the interregional variance, the intraregional variance, or the regional standard deviations. The consistency of these results is rather impressive, especially when it is remembered that the controls used here are rough and exploratory.

As was found in the case of metropolitan areas, both the regions and the behavior variables exhibit a differential susceptibility to particular controls. Moreover, regional distinctiveness among nonmetropolitan areas (like their metropolitan counterparts) tends to rest on the distinctiveness or deviant positions of only one or two regions, while the other regions are markedly similar. Generally speaking, it appears that the frequency of reductions in behavior differences both internal and external with respect to regions is greater for nonmetropolitan populations. However, the magnitude of the reductions in the correlation of behavior with regionalism is greater for metropolitan populations.

Unfortunately, no precise conclusions can be drawn from these findings about the presumed effects of regional differences in social structure and standard of living for regional differences in behavior. That relationships exist between these types of phenomena cannot reasonably be denied. However, the extent of these relationships is clouded with uncertainty. This, of course, arises from the problems surrounding the control procedure which is used in this study. That is to say, when by the application of controls the distributions of regional weights are altered, there is no way of determining accurately how much of the subsequent change in correlation measures is influenced by these alterations and how much of it can be attributed to intrinsic changes in behavior.

CHAPTER VI

SUMMARY AND CONCLUSIONS

The individual summaries at the close of each preceding chapter in this report preclude the necessity for a detailed summary in this chapter. It would seem advisable, however, to present some sort of overview which summarizes and places in perspective the entire research problem and the findings derived from its investigation. Accordingly, the material in this chapter consists of a general but brief abstract of the study together with an evaluation of its procedures and results.

The Problem, Research Design, and General Findings

The basic questions addressed in this research may be stated briefly as follow: (a) To what extent does regionalism represent an independent variable for human behavior? (b) To what extent do structural and socio-economic characteristics constitute features of social life which give to regionalism its "independent" character? These questions are derived from the work of regional scholars and other investigators who have made extensive use of regional classifications for both theoretical and research purposes. A good deal of empirical research data has shown noticeable differences among regions with respect to a variety of social phenomena. This kind of demonstration of heterogeneity among regions has usually not been accompanied by very much concern for

the question of homogeneity within regions. This latter condition has tended to be assumed rather than demonstrated. Since the efficacy of the regional concept as a research tool rests fundamentally on conditions of similarity within regions and of differences among regions, neglect of either of these only serves to cast doubt or at least to raise questions about results which appear in such regional studies.

The "independent" character of regionalism, alluded to above, figures rather prominently in the literature of this field of study. Very often this idea is not directly stated but rather appears as an implicit assumption in the work of regional scholars. It is felt by these investigators that the region is an area of distinctive, human group living. The roots of this distinctiveness are embedded in the peculiarities of the natural environment and in the idiosyncrasies of the inhabitants' cultural heritage. Thus, the region is a force which shapes or at least influences in important ways the ideas, interests, and desires of its residents. It is in this sense that regionalism may be regarded as an independent variable for human behavior.

The research undertaken in this study represents an attempt to measure the extent to which different regional populations in the United States have distinctive unitary characters with respect to behavior. This, of course, requires a consideration not only of behavioral differences among the several regions but also of behavioral similarities within each region. Accordingly, statistical treatment calls for an analysis of variance technique. The measure used throughout the study is the intraclass correlation (rho). Basically, this measure indicates what proportion of the

total variation in a particular type of behavior is associated with the regional locations of that behavior. Although such a measure does not demonstrate a cause-and-effect relationship between regional location and behavior, it does provide an estimate of the extent to which regions can be regarded as "explanatory" variables for human behavior.

Of the many regional delineations available for investigation, the one finally selected is that proposed by the late regional scholar, Professor Howard W. Odum. This delineation divides the country into six regions, each of which consists of four or more entire states. Geographically, these may be identified as follow: Northeast, Southeast, Middle States, Northwest, Southwest, and Far West.

The types of behavior investigated are five in number:

1. Net reproduction rate, 1936-1940.
2. Proportion of eligible voters voting, 1940.
3. Proportion of voters voting Democratic, 1940.
4. Proportion of adult males with no formal schooling, 1940.
5. Proportion of population with religious affiliation, 1936.

These types of behavior are regarded as dependent variables in the context of this study. Each of them represents a behavioral product of motivations which presumably are rooted in culture patterns. If regions are genuinely distinctive cultural subdivisions of the larger society, then one should expect to find some behavioral diversity among regions. The intraclass correlation for each dependent variable shows the extent to which this expectation is met.

The basic measurement units selected for this study are

State economic areas, each of which consists of one or more entire counties. In effect, therefore, each region represents a population of State economic areas. The total number of these areas for the entire country is 443.

In addition to measuring the extent to which the behavior of State economic area populations is regionally distinctive, this research also attempts to investigate how much of this regional distinctiveness appears to be associated with regional differences in structural and socio-economic characteristics. This is done by controlling for such characteristics and then observing what changes appear in the value of the intraclass correlations as a consequence of these controls. It is felt that such controls should tend to level regional differences in behavior because, by controlling for these characteristics, the basic conditions of life among the several regions are rendered more similar. If behavior patterns are viewed as forms of adaptation to these objective conditions of life, it follows that some convergence of behavior patterns should accompany the exercise of controls.

The control technique employed consists of selecting for measurement those population units (i.e., State economic areas) with specified characteristics and of rejecting all others. Thus, for example, controlling for metropolitan status involves selecting from each region only those State economic areas which are metropolitan in character. An intraclass correlation is then computed for these metropolitan units, and the magnitude of that correlation indicates the extent to which metropolitan populations are regionally distinctive with respect to the particular kind of behavior being measured. By comparing the value of this metropolitan

correlation with that of the correlation for all units (i.e., metropolitan and nonmetropolitan) one may infer at least roughly how much of the total regional differences in behavior is associated with regional differences in metropolitan status. The particular characteristics for which controls were exercised are as follow:

1. Metropolitan status.
2. Proportion of total population which is urban.
3. Proportion of total population which is rural nonfarm.
4. Proportion of total population which is rural farm.
5. Proportion of employed workers in manufacturing.
6. Proportion of employed workers in agriculture.
7. Proportion of employed workers in wholesale and retail trade.
8. Rural farm level of living index.
9. Per capita dollar retail sales.
10. Per capita dollar receipts from services.
11. Proportion of rural farm labor force employed in nonfarm occupations.
12. Proportion of rural farm labor force employed as unpaid family workers.
13. Proportion of total population which is nonwhite.

All of these data are for 1940 and are viewed as independent factors in the context of this study.

The general results of this investigation of regionalism can be summarized broadly as follow: First, the significance of regional location for behavior varies greatly with the different types of behavior examined. For example, viewing the country as a whole the relationship between regionalism and fertility is almost negligible. On the other hand, the relationship between regionalism and political behavior (as exhibited by voting) is quite pronounced. Second, regional differences in behavior do not appear in the form of each region's assuming a distinctive behavioral pattern. Rather, regional distinctiveness seems to rest on "peculiar" behavior patterns displayed by only one or two regions

while the remaining regional populations differ little from one another. For example, with respect to political behavior it is the two Southern regions which are distinctive, while for religious affiliation it is the Far West and to some extent the Northeast which display notably different patterns of behavior. Although the extent of the relationship between regionalism and fertility is small, most of the association appears to be due to the deviant positions of the high-fertility Southeast and the low-fertility Far West. For educational achievement it is the two Southern regions which contribute the most to the correlation with regionalism. Third, differences in the social structure and the socio-economic level of the various regional populations rather definitely appear to be related to regional differences in behavior. For example, the amount of regional distinctiveness in fertility, educational achievement, and religious affiliation is dramatically altered when one distinguished between metropolitan and nonmetropolitan types of social organization.

These general findings are not particularly surprising if it is kept in mind that they have reference to one point in a transitional period which seems to be ushering in a full-scale urban-industrial social order. A society in transition does not exhibit uniform and simultaneous changes in all its cultural features. In this regard the so-called principle of "culture lag" has long been observed by investigators of social change. Accordingly, if one may assume that the relative isolation of a former day gave rise to culturally distinct regions, the breaking-down of this isolation can be expected to have differential effects on both the cultural items and the particular regions involved in the

change.

In this same culture lag context it should also be noted that the behavior variables which exhibit the greatest amount of regional convergence in the metropolitan type social structure are fertility and educational achievement. It seems to the writer that in comparison with the political and religious types of activity, the fertility and educational variables represent behavior that is much more intimately involved in the routine, day-to-day functioning of a social system. That is to say, they might be viewed as part of the "social reflexes" of a population which is adjusting to new conditions of life. The consequences of inappropriate fertility patterns or inadequate educational equipment are felt rather immediately in the work-a-day world. On the other hand, the social and philosophical beliefs which hold sway in the political and religious arenas of life are less directly involved in the immediate, daily adjustments of people. This phenomenon has been observed by many students of social change including Professor Ogburn, who has remarked on the fact that social beliefs and philosophies are usually among the last features of culture to exhibit noticeable change in transition periods.

The general conclusion that regional differences in social structure and socio-economic level are associated with regional differences in behavior deserves perhaps a little more comment. It has already been noted that for three of the five dependent behavioral variables (viz., fertility, educational achievement, and religious affiliation) measures of regional distinctiveness are greatly altered when metropolitan and nonmetropolitan populations are compared with one another. Furthermore, a rather detailed

examination of differences in the regional distinctiveness of metropolitan and nonmetropolitan populations reveals that metropolitan areas tend to level regional differences in behavior. Since these metropolitan areas are the least isolated and the most urbanized and industrialized units of social organization, the fact that regional differences in behavior tend to be reduced in these areas supports the main hypothetical expectation of this study--namely, that regionalism is a product of isolation, and that with the emergence of urban industrialism, this condition of isolation is broken down, and behavior becomes adapted to urban conditions which tend to become similar for all regions.

This investigation of the behavioral effects of regional differences in structure and socio-economic characteristics was pursued further by controlling the metropolitan and nonmetropolitan populations for a variety of other characteristics. In over two-thirds of the cases where additional controls were applied, there was a measurable reduction in the amount of regionally distinctive behavior exhibited. That is to say, regional differences in behavior tended to be leveled. Furthermore, this occurred <u>despite</u> the fact that the application of additional controls tended to render the individual regions more homogeneous internally by reducing behavioral diversity within each region. For metropolitan areas the additional control which seemed to be most effective in reducing behavior differences was that for wholesale and retail trade employment. For nonmetropolitan areas the greatest regional differences in behavior tended to appear among the most rural agricultural populations. However, when standard of living was also controlled for these rural groups, regional differences in behavior

were greatly reduced.

It should be noted, however, that the success with which structural and socio-economic controls served to level regional differences in behavior was not very great for the two political voting variables. The importance of regional location for this type of behavior appears indisputable and of course rests fundamentally on the rather deviant position of the one-party Southern regions. The only control which seemed to alter substantially the voting pattern in the South was that which selected out sub-areas with low proportions of nonwhite residents in the population. It may be argued that such a control is "unfair" because of the strategic importance of the race issue in the South. While the validity of the assertion that race is an important issue in the South cannot reasonably be disputed, it is interesting to note that the cultural distinctiveness of the South with respect to voting can be quantitatively determined within rough limits by simply observing the incidence of Negro residents in the population. The voting pattern in sub-areas of the South is most distinctive where the proportion of Negroes is high, while in those sub-areas where the proportion is low, the voting pattern tends to resemble that for the rest of the country. In short, the essence of Southern regionalism as far as voting behavior is concerned seems to be strongly related to the factor of population composition.

It has already been reported that over two-thirds of the detailed structural and socio-economic controls succeeded in bringing about reductions in regional behavior differences. This success, however, was only a partial one. With but a few exceptions most of these reductions left a significant amount of regional

diversity in behavior still unaccounted for. In this regard it is instructive to note that the most dramatic alterations in regional behavior patterns were effected by the big structural control represented by the metropolitan area. This suggests perhaps that at our present stage of knowledge social structure is best represented by syndromes of characteristics rather than individual index items.

Evaluation and Suggestions for Further Research

An evaluation of this study should be placed in perspective by keeping in mind the exploratory nature of the work. That is to say, most of the critical comments on the study's inadequacies arise from wisdom born of hindsight. Certain shortcomings in the research design have already been pointed out in Chapter II. Most of these involve problems of procuring appropriate data and of tailoring the project to dimensions manageable for one researcher with very limited financing. There was one decision, however, made at the research design stage of the study, which in retrospect is now regretted. This was the decision to use State economic areas exclusively as the basic measurement unit.

Although the selection of State economic areas as the basic measurement unit provides a total sample of 443 cases which encompass the entire country, the distribution of these cases among the six regions is very uneven. Some regions have four times as many representative units as other regions. This means that the application of controls is considerably restricted if some representation of all six regions is to be maintained. It is, of course, highly probable that no unit could have been found which would

obviate this problem of differential representation among the several regions. However, the stringency of the restriction on controls posed by this problem could have been eased by a larger sample size. It now appears that the best solution to this problem would have been to retain the metropolitan State economic areas as representative of the highly urban units but then to sample counties for the nonmetropolitan representation. In this way the total sample could have been increased to perhaps 1,000 cases. This would not improve the control procedure for the metropolitan segment of the population, but it would expand greatly the opportunities for analysis of the nonmetropolitan segment.

Another shortcoming in the research design which could have been avoided (although not without a considerable expenditure of time and effort) was the inadequate index of urbanism. The characteristic-proportion of the total population which is urban-is not very satisfactory because it indicates nothing about the magnitude of urban concentration. Thus, for example, two areas may have the same size population and the same proportion urban, but in one area the urban population is concentrated in one or two cities while in the other area the urban population is distributed among a large number of small towns. It would have been desirable and possible to classify the basic measurement units by size of largest city. Although this would not be completely satisfactory, it would be a considerable improvement over the simple proportion-of-total-population-urban index.

Unfortunately, neither of the inadequacies cited above can be viewed with much concern when attention is directed to the major difficulty encountered in the study. This is a problem rooted in

certain basic features of the research design--namely, the control procedure and the measuring instrument used.

It will be recalled that controls are applied in this study by selecting from the universe of cases those units (State economic areas) which have certain characteristics. Thus, in a sense controlling constitutes a sampling type of operation in which the selected units comprise a control group that is homogeneous with respect to one or more characteristics. Since the control characteristics differ in the way they are distributed among the several regions, this control procedure via selection serves to establish control groups which may differ considerably in size and in the amount of representation accorded the various regions. Thus, for example, one control group may consist of 80 State economic areas 30 per cent of which represent the Northeast region while another control group may have 100 State economic areas only 15 per cent of which represent the Northeast region. This sort of phenomenon in and of itself is not particularly troublesome, and indeed, provides interesting information about how the real world is arranged in terms of regional identification. Difficulties, however, enter when measurements are applied and comparisons are made.

The intraclass correlation (rho) which is used throughout this study is a summary measure which indicates what proportion of the total variance in behavior is associated with the regional location of that behavior. Computation of this measure involves the use of a weighted average for each region, and the weighting factor is the number of units (State economic areas) representing each region. It can be clearly seen, therefore, that the value of the

intraclass correlation is affected to some undetermined degree by a control procedure which alters regional weighting factors. This, of course, means that the effect of particular characteristics on regional differences in the types of behavior measured cannot be specified with any precision.

This problem bears a close resemblance to the one that demographers run into when they must compare the vital rates of two populations whose age compositions differ. Some standardization technique must be employed. With respect to this study it would probably be possible to establish some standard regional composition of units which would represent a stable weighting pattern for the various regional averages derived from different controls. In this way, differences in the summary rho measures could be interpreted as fairly precise indexes of the influence of particular characteristics on regional differences in behavior. The question of what kind of standard weighting pattern should be used introduces a subject for endless debate. Three possible alternatives can be mentioned which might be considered--namely, patterns based on (a) equal weights for all regions, (b) weights corresponding to the regional representation of basic measurement units, and (c) weights corresponding to the population size of each region.

The question of what further research is prompted by the results of this study can be discussed rather expeditiously. Perhaps the most obvious, unanswered question raised in this study is that of whether the distinctiveness of regional culture patterns is being reduced over time as our society becomes more urbanized and industrialized. Evidence presented in this study prompts this expectation. However, definitive proof requires a duplication of

this study at other points in time. If this expectation should prove to be true, duplication of this study design (with appropriate modifications) would provide information as to whether such cultural convergence is a product of decreasing homogeneity within regions and/or decreasing heterogeneity among regions.

Another question for further research relates to a methodological problem of statistical import. Throughout this study the computation of the intraclass correlation coefficient has involved treating the basic measurement units (State economic areas) as standard comparable units. In point of fact, however, the population size of these units does exhibit considerable variation. It would be of more than passing interest therefore to see what effect a weighting of these units according to population size would have on the value of the correlation.

A third area for further research involves focusing attention on regional differences in the characteristics of Standard Metropolitan Areas. Various research reports have shown a number of noticeable regional differences in metropolitan characteristics. However, little or no application of analysis of variance techniques has been employed in these studies. The findings of this study which show a leveling of regional differences in metropolitan areas for <u>certain</u> kinds of behavior cannot be safely interpreted as indicating that metropolitan areas are similar in all characteristics. Indeed, to the extent that our society is developing a regional division of labor one might reasonably expect to find some regional selectivity with respect to population types and functions in metropolitan areas.

The writer's interest in regional comparisons of

metropolitan centers, however, extends beyond a mere desire to supplement the descriptive content of previous studies. Rather, it is believed that at least some metropolitan centers occupy strategic positions with respect to interregional contact. This is supported in part by the fact that the nineteen largest metropolitan centers in the country are located on navigable water routes. Such routes provide for a large-bulk, long-haul type of transportation which facilitates interregional exchange. Indeed, in one sense it may be argued that metropolitan centers are essential for the establishment of a genuine, regionally organized social system. At first glance this may seem to contradict the hypothesis that metropolitan centers serve to level regional differences. Actually, it is not so much a contradiction as it is a modification of the hypothesis.

Regions which are born of and exist in conditions of isolation are really not regions in a sociological sense if by the term "region" we mean an area which is a functional subpart of a larger social system. A so-called region which is completely isolated represents a self-sustaining entity in and of itself. It is only when such an area establishes regular contact with other areas and enters into exchange relationships with these other areas that interregional dependencies arise which serve to identify the emergence of a more comprehensive unit of social organization. The question of just how much interaction must occur between areas before they take on the character of true regions cannot be answered with any precision except arbitrarily because it is basically a definition-type question. Nevertheless, some degree of interaction is necessary if an area is to be assigned a regional character, in

the sociological sense of the term. The writer feels that this rather fundamental sociological viewpoint has often been overlooked or at least neglected in much of the regionalists' work.

Accepting this sociological viewpoint one may still reflect on the question of how interregional exchange is affected. Quite obviously, all the individuals of one region do not interact directly with all the individuals of another region. Rather, it would seem that a territorial division of labor among regions presupposes some measure of internal organization of regional populations. Furthermore, it is probable that the more dependent a regional population is on interregional exchange, the more likely it is to exhibit an internal organization which is oriented toward the mediating agencies that dominate this interregional exchange relationship. Since dominance, ecologically speaking, tends to vary with community size, and since we have already noted that the water-site locations of the largest metropolitan centers are strategic for mediating interregional exchange, it would seem reasonable to infer that such large centers probably serve as dominating nuclei for the internal organization of regional populations which are caught up in such exchange patterns. At least this is a potentially testable hypothesis.

There are a number of interesting and sociologically pertinent phenomena which seem to accompany the development of exchange relationships among interdependent units of a social system. For example, it can hardly be disputed that dependency relationships which involve the procurement and distribution of sustenance materials generate strong compulsions among the participating units to maintain and strengthen the system. This is simply another way

of noting the tendency for social systems to become stabilized and to resist changes which appear to interrupt the sustenance flow. By the same reasoning one would expect the participating units to exhibit some standardization of cultural equipment in order to facilitate the smooth functioning of the system. This means that as a social system expands and becomes more finely articulated (and hence complex), a compelling need arises to level any differences which might impede the operation of sustenance functions. Indeed, once a unit is caught up in such a pattern of interdependencies, the rewards for its effective participation in the system are likely to prompt an increasingly intensive commitment. Conversely, attempts to withdraw from the system are likely to involve substantial penalties because of the sustenance consequences. Thus, the process of involvement tends to become a one-way street.

It should be noted, however, that a complete convergence of regional culture patterns is hardly to be expected in even the most highly integrated social system. In point of fact the possibility of complete convergence is precluded by the regional differences which constitute the foundation for the whole system's territorial division of labor. Furthermore, it is probable that there are quite a variety of regionally distinctive lifeways which are of significance only for the local populations and which do not materially influence the interregional exchange of sustenance materials. Such regional differences may persevere more or less indefinitely providing their consequences remain purely local. However, even this is not assured because the establishment and elaboration of an interregional exchange system is almost inevitably

accompanied by increased population mobility. To the extent that this occurs, local culture patterns become subject to change through modification and supplementation. In the light of these remarks the writer sees the future as holding little promise for the development of a social system in which "cultural pluralism" is conspicuous on a regional basis.

The picture which has been sketched here of an emerging urban-industrial social system with a regionally organized division of labor dominated by large metropolitan centers is frankly hypothetical. It is a picture, however, which is partially validated by some of the findings of this research study. In this regard, one may cite the convergence of group behavior patterns in metropolitan areas. This convergence is evident for metropolitan areas both within and among the several regions studied. It is also instructive to note that the areas most committed to rural agricultural activity are among those which show the least tendency toward an interregional convergence of the behavior patterns examined. In the light of these findings it is difficult to dispute the tentative conclusion that regionally distinctive behavior is contingent upon the structural conditions of collective life. Furthermore, it is difficult to dispute the inference that the dependency of metropolitan populations on interregional exchange relationships is greater than that of populations in rural agricultural areas. In short, although there is nothing in these research findings which illuminates an interregional division of labor, there are findings which support derivative expectations based on the assumption of a developing interregional division of labor.

In conclusion, it is the writer's belief that the most promising area for future research on the subject of regionalism is that of functional integration with specific attention paid to the role of metropolitan centers. Regional scholars, it is felt, have placed too much emphasis on a cultural diversity which is the fading heritage of a former pattern of living conditions. They have tended to neglect changes which seem to be binding the nation's various territorial areas into a relatively close-knit social system. It is these changes and their cultural ramifications which provide rich opportunities for fruitful research today and in the future.

APPENDIX A

TABLES

TABLE 1

FREQUENCY DISTRIBUTIONS OF METROPOLITAN STATE ECONOMIC AREAS FOR EACH REGION BY THE PROPORTION OF THE TOTAL POPULATION WHICH IS URBAN

Percentage	Regions					
	North-east	South-east	Middle States	North-west	South-west	Far West
90+	6	1	5			
85-89	2		6	1	1	1
80-84	6	6	7	3	2	2
75-79	4	2	2	1	4	1
70-74	9	8	7	1	2	2
65-69	3	5	6			
60-64	5	4	3			4
55-59	1	3	2		1	
50-54	1	1	1			1
45-49	2	1	1			1
40-44	1	2			1	1
35-39	2					
30-34			1			
25-29						
20-24		1				
15-19	1	1				
Under 15. . .						
Total . . .	43	35	41	6	11	13

TABLE 2

FREQUENCY DISTRIBUTIONS OF METROPOLITAN STATE ECONOMIC AREAS FOR EACH REGION BY THE PROPORTION OF THE TOTAL POPULATION WHICH IS RURAL NONFARM

Percentage	Regions					
	North-east	South-east	Middle States	North-west	South-west	Far West
50-69	2					
45-49	1					
40-44	1					
35-39		2	1		1	
30-34	5	2	2			1
25-29	5	6				1
20-24	8	9	13	1	1	6
15-19	8	6	3	1	5	1
Under 15. . .	13	10	22	4	4	4
Total . . .	43	35	41	6	11	13

TABLE 3

FREQUENCY DISTRIBUTIONS OF METROPOLITAN STATE ECONOMIC AREAS FOR EACH REGION BY THE PROPORTION OF THE TOTAL POPULATION WHICH IS RURAL FARM

Percentage	Regions					
	North-east	South-east	Middle States	North-west	South-west	Far West
50+						
45-49		1				
40-44		2				
35-39						
30-34		2	1			2
25-29		1				1
20-24	3	2	4		2	
15-19	3	4	1			
10-14	8	4	14	1	1	6
5-9	12	15	8	5	8	1
Under 5 . . .	17	4	13			3
Total . . .	43	35	41	6	11	13

TABLE 4

FREQUENCY DISTRIBUTIONS OF METROPOLITAN STATE ECONOMIC AREAS FOR EACH REGION BY THE PROPORTION OF THE TOTAL POPULATION WHICH IS NONWHITE

Percentage	Regions					
	North-east	South-east	Middle States	North-west	South-west	Far West
45+		5				
40-44		5				
35-39		2				
30-34		5				
25-29		5				
20-24		3			2	
15-19	2	3			3	
10-14	4	3	4	1	3	1
5-9	6	2	11	1	2	2
Under 5 . . .	31	2	26	4	1	10
Total . . .	43	35	41	6	11	13

TABLE 5

FREQUENCY DISTRIBUTIONS OF METROPOLITAN STATE ECONOMIC AREAS FOR EACH REGION BY THE PROPORTION OF EMPLOYED PERSONS WORKING IN AGRICULTURE

Percentage	Regions					
	North-east	South-east	Middle States	North-west	South-west	Far West
65+						
60-64						
55-59						
50-54						
45-49						
40-44		1				
35-39		1				1
30-34						
25-29		1	1			1
20-24		2	1		2	
15-19	3	2	3			3
10-14	4	6	4	1	1	1
5-9	13	10	17	3	4	4
Under 5 . . .	23	12	15	2	4	3
Total . . .	43	35	41	6	11	13

TABLE 6

FREQUENCY DISTRIBUTIONS OF METROPOLITAN STATE ECONOMIC AREAS FOR EACH REGION BY THE PROPORTION OF THE EMPLOYED LABOR FORCE IN MANUFACTURING

Percentage	Regions					
	North-east	South-east	Middle States	North-west	South-west	Far West
50-59	3		4			
45-49	7	1	7			
40-44	9	2	2			
35-39	8		7			
30-34	5	2	8		1	
25-29	2	9	2	1		2
20-24	4	7	4		1	3
15-19	2	5	3	3	3	1
10-14	2	6	2	1	4	6
Under 10. . .	1	3	2	1	2	1
Total . . .	43	35	41	6	11	13

TABLE 7

FREQUENCY DISTRIBUTIONS OF METROPOLITAN STATE ECONOMIC AREAS FOR EACH REGION BY THE PROPORTION OF THE EMPLOYED LABOR FORCE IN WHOLESALE AND RETAIL TRADE

Percentage	Regions					
	North-east	South-east	Middle States	North-west	South-west	Far West
25+	1	1	3	2	3	
23-24		4	1	3	4	5
20-22	4	8	15	1	3	5
18-19	9	7	5		1	2
15-17	18	8	16			
13-14	10	3	1			1
10-12	1	1				
8-9		3				
Under 8 . . .						
Total . . .	43	35	41	6	11	13

TABLE 8

FREQUENCY DISTRIBUTIONS OF METROPOLITAN STATE ECONOMIC AREAS FOR EACH REGION BY PER CAPITA DOLLAR RETAIL SALES

Per Capita Dollars	Regions					
	North-east	South-east	Middle States	North-west	South-west	Far West
475+	1	1	2	1		2
450-474 . . .	2				1	4
425-449 . . .	6	1	6	1	1	2
400-424 . . .	4		8	3	1	2
375-399 . . .	7	2	9		3	
350-374 . . .	2	5	5		2	2
325-349 . . .	9	6	6		1	
300-324 . . .	4	3			1	
275-299 . . .	3	6	3		1	1
250-274 . . .	5	8				
225-249 . . .				1		
200-224 . . .			1			
175-199 . . .			1			
150-174 . . .						
125-149 . . .		1				
100-124 . . .		1				
Under 100 . .		1				
Total . . .	43	35	41	6	11	13

TABLE 9

FREQUENCY DISTRIBUTIONS OF METROPOLITAN STATE ECONOMIC AREAS FOR EACH REGION BY PER CAPITA DOLLAR RECEIPTS FROM SERVICES

Per Capita Dollars	Regions					
	North-east	South-east	Middle States	North-west	South-west	Far West
55+	1					4
45-54		1	2		1	
40-44			5	2	1	1
35-39	1	4	3	1	2	1
30-34	11	8	6	2	3	4
25-29	6	6	7		1	1
20-24	10	6	7		2	1
15-19	10	6	8	1	1	1
10-14	3	1	2			
5-9	1	2	1			
Under 5 . . .		1				
Total . . .	43	35	41	6	11	13

TABLE 10

FREQUENCY DISTRIBUTIONS OF METROPOLITAN STATE ECONOMIC AREAS FOR EACH REGION BY THE RURAL FARM LEVEL OF LIVING INDEX

Index	Regions					
	North-east	South-east	Middle States	North-west	South-west	Far West
135+						
130-134 . . .	1		6			
125-129 . . .	8		8	1		1
120-124 . . .	15		14	1		4
115-119 . . .	8		4	3		4
110-114 . . .	6	1	3	1		4
105-109 . . .	3	1	2		2	
100-104 . . .		5	2		1	
95-99		5	1		5	
90-94		3			1	
85-89		5	1			
80-84	1	4			2	
75-79	1	3				
70-74		1				
65-69		1				
60-64		2				
55-59		2				
Under 55. . .		2				
Total . . .	43	35	41	6	11	13

TABLE 11

FREQUENCY DISTRIBUTIONS OF METROPOLITAN STATE ECONOMIC AREAS FOR EACH REGION BY THE PROPORTION OF THE RURAL FARM LABOR FORCE WHO ARE UNPAID FAMILY WORKERS

Percentage	Regions					
	North-east	South-east	Middle States	North-west	South-west	Far West
28+		2				
25-27						
22-24						
19-21		2				
16-18		1	2		1	
13-15		5	2		1	
10-12	2	5	6	2	1	1
7-9	15	9	11	2	2	2
4-6	18	6	14	1	5	8
Under 4 . . .	8	5	6	1	1	2
Total . . .	43	35	41	6	11	13

TABLE 12

FREQUENCY DISTRIBUTIONS OF METROPOLITAN STATE ECONOMIC AREAS FOR EACH REGION BY THE PROPORTION OF THE RURAL FARM LABOR FORCE IN NONFARM OCCUPATIONS

Percentage	Regions					
	North-east	South-east	Middle States	North-west	South-west	Far West
55+	7	6	5	1		
50-54	12	6	5		2	1
45-49	6	4	7			
40-44	12	5	3	1	1	2
35-39	2	2	5	1	3	3
30-34	3	2	9		1	4
25-29	1	4	3	1	1	2
20-24		3		1	2	
15-19		1	2	1	1	1
Under 15. . .		2	2			
Total . . .	43	35	41	6	11	13

TABLE 13

PERCENTAGE DISTRIBUTIONS OF METROPOLITAN STATE ECONOMIC AREAS AMONG REGIONS FOR THE TOTAL GROUP OF METROPOLITAN STATE ECONOMIC AREAS AND TWELVE SELECTED CONTROL GROUPS

Regions	Total Metrop. Areas	Control Groups*											
		S-1	S-2	S-3	S-4	S-5	S-6	S-7	S-8	S-9	S-10	S-11	S-12
Northeast	28.8%	30.2%	38.8%	35.5%	34.1%	40.6%	16.9%	16.2%	26.7%	20.0%	18.2%	7.9%	36.0%
Southeast	23.5	24.1	25.0	20.4	19.5	7.8	33.8	32.3	33.3	25.7	25.8	20.6	6.1
Middle States	27.5	25.0	20.0	24.7	22.0	28.1	18.2	21.2	15.0	20.0	16.7	30.2	36.0
Northwest	4.0	4.3	3.8	3.2	4.9	6.2	6.5	6.1	5.0	5.7	7.6	9.5	5.3
Southwest	7.4	8.6	6.2	7.5	9.8	4.7	13.0	11.1	10.0	11.4	13.6	15.9	5.3
Far West.	8.7	7.8	6.2	8.6	9.8	12.5	11.7	13.1	10.0	17.1	18.2	15.9	11.4
Total	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
N	149	116	80	93	82	64	77	99	60	70	66	63	114

*The specific controls used are as follow:

S-1: Urban < 85%, Rural farm < 25%.
S-2: " " " , " " " " , Nonfarm occupation = 35%+.
S-3: " " " , " " " " , Unpaid family workers < 10%.
S-4: " " " , " " " " , Farm level of living = 85-124.
S-5: " " " , " " " " , " " " " " " , Nonwhite < 15%.
S-6: " " " , " " " " , Manufacturing employment < 35%.

S-7: Manufacturing employment < 35%.
S-8: " " " " , Nonfarm occupation = 35%+.
S-9: " " " " , Unpaid family workers < 10%.
S-10: " " " " , Farm level living = 85-124.
S-11: Wholesale and retail trade employment = 20%+.
S-12: Nonwhite < 15%

APPENDIX B

TABLES

TABLE 1

FREQUENCY DISTRIBUTIONS OF NONMETROPOLITAN STATE ECONOMIC AREAS FOR EACH REGION BY THE PROPORTION OF THE TOTAL POPULATION WHICH IS URBAN

Percentage	Regions					
	North-east	South-east	Middle States	North-west	South-west	Far West
90+	1					
85-89						
80-84						
75-79			1			
70-74	1					
65-69	1			1		
60-64		1	2			
55-59	1	2	2			1
50-54	6	1	4	2		
45-49	3	1	4	2	3	1
40-44	5	2	9	2	1	2
35-39	6	1	7	7	8	4
30-34	5	7	10	4	4	6
25-29	2	15	11	3	5	5
20-24	5	17	8	8	4	
15-19	3	15	6	3	3	1
Under 15. . .	4	32	4	7	1	1
Total . . .	43	94	68	39	29	21

TABLE 2

FREQUENCY DISTRIBUTIONS OF NONMETROPOLITAN STATE ECONOMIC AREAS FOR EACH REGION BY THE PROPORTION OF THE TOTAL POPULATION WHICH IS RURAL NONFARM

Percentage	Regions					
	North-east	South-east	Middle States	North-west	South-west	Far West
50-69	8	3		3		1
45-49	2	3			1	2
40-44	10	3	1	2	1	4
35-39	9	7	2	2	4	7
30-34	9	14	14	12	2	4
25-29	1	23	28	12	7	3
20-24	2	16	18	7	10	
15-19	1	19	4	1	4	
Under 15. . .	1	6	1			
Total . . .	43	94	68	39	29	21

TABLE 3

FREQUENCY DISTRIBUTIONS OF NONMETROPOLITAN STATE ECONOMIC AREAS FOR EACH REGION BY THE PROPORTION OF THE TOTAL POPULATION WHICH IS RURAL FARM

Percentage	Regions					
	North-east	South-east	Middle States	North-west	South-west	Far West
70+		11				
65-69		10				
60-64	1	17	4	1	2	
55-59	1	13	3	3	1	
50-54		8	7	4	7	
45-49	1	11	12	6	2	
40-44	2	7	12	6	4	2
35-39	4	8	9	10	5	3
30-34	3	2	8	2	3	6
25-29	8		7	3	2	2
20-24	4	2	3	2	2	4
15-19	8	1	3	2	1	3
10-14	7	2				1
5-9	3	1				
Under 5 . . .	1	1				
Total . . .	43	94	68	39	29	21

TABLE 4

FREQUENCY DISTRIBUTIONS OF NONMETROPOLITAN STATE ECONOMIC AREAS FOR EACH REGION BY THE PROPORTION OF THE TOTAL POPULATION WHICH IS NONWHITE

Percentage	Regions					
	North-east	South-east	Middle States	North-west	South-west	Far West
45+		17				
40-44		6				
35-39	1	17			1	
30-34		8			1	
25-29	2	4			2	
20-24		8			3	
15-19	1	7			2	
10-14	1	6	2	1	3	
5-9	5	9		2	6	4
Under 5 . . .	33	12	66	36	11	17
Total . . .	43	94	68	39	29	21

TABLE 5

FREQUENCY DISTRIBUTIONS OF NONMETROPOLITAN STATE ECONOMIC AREAS FOR EACH REGION BY THE PROPORTION OF EMPLOYED PERSONS WORKING IN AGRICULTURE

Percentage	Regions					
	North-east	South-east	Middle States	North-west	South-west	Far West
65+		6				
60-64		12	1	1	1	
55-59		11	5	5	1	
50-54		11	5	4	3	
45-49	1	7	8	4	5	
40-44	1	12	6	9	5	2
35-39	5	10	13	7	5	3
30-34	3	8	9	1	5	6
25-29	2	7	9	4	2	2
20-24	5	6	3	3	1	2
15-19	7	1	5	1	1	3
10-14	7	1	4			1
5-9	11	1				2
Under 5 . . .	1	1				
Total . . .	43	94	68	39	29	21

TABLE 6

FREQUENCY DISTRIBUTIONS OF NONMETROPOLITAN STATE ECONOMIC AREAS FOR EACH REGION BY THE PROPORTION OF THE EMPLOYED LABOR FORCE IN MANUFACTURING

Percentage	Regions					
	North-east	South-east	Middle States	North-west	South-west	Far West
50+	1					
45-49	2	1	1			
40-44	2	2	2			1
35-39	5	3				2
30-34	5	2	5			
25-29	5	6	7			3
20-24	6	8	8		1	2
15-19	8	18	11	2		1
10-14	5	22	15	3	4	2
Under 10. . .	4	32	19	34	24	10
Total . . .	43	94	68	39	29	21

TABLE 7

FREQUENCY DISTRIBUTIONS OF NONMETROPOLITAN STATE ECONOMIC AREAS FOR EACH REGION BY THE PROPORTION OF THE EMPLOYED LABOR FORCE IN WHOLESALE AND RETAIL TRADE

Percentage	Regions					
	North-east	South-east	Middle States	North-west	South-west	Far West
25+						
23-24						
20-22		3			3	3
18-19	1		2	4	1	2
15-17	8	3	25	20	14	10
13-14	25	8	29	8	5	6
10-12	7	24	12	7	6	
8-9	1	36				
Under 8 . . .	1	20				
Total . . .	43	94	68	39	29	21

TABLE 8

FREQUENCY DISTRIBUTIONS OF NONMETROPOLITAN STATE ECONOMIC AREAS FOR EACH REGION BY PER CAPITA DOLLAR RETAIL SALES

Per Capita Dollars	Regions					
	North-east	South-east	Middle States	North-west	South-west	Far West
475+	2					2
450-474 . . .	1			1		1
425-449 . . .						2
400-424 . . .	2			1		3
375-399 . . .	3			1	1	4
350-374 . . .	8	1	5	5	1	1
325-349 . . .	6	1	3	2	1	6
300-324 . . .	2	2	13	3		2
275-299 . . .	4	1	8	7	5	
250-274 . . .	3	1	14	5	3	
225-249 . . .	4	1	7	9		
200-224 . . .	4	4	10	3	3	
175-199 . . .	1	10	5	1	9	
150-174 . . .	1	22	1	1	4	
125-149 . . .	1	23	2		1	
100-124 . . .	1	20			1	
Under 100 . .		8				
Total . . .	43	94	68	39	29	21

TABLE 9

FREQUENCY DISTRIBUTIONS OF NONMETROPOLITAN STATE ECONOMIC AREAS FOR EACH REGION BY PER CAPITA DOLLAR RECEIPTS FROM SERVICES

Per Capita Dollars	Regions					
	North-east	South-east	Middle States	North-west	South-west	Far West
55+	1					
45-54						
40-44						
35-39	1					
30-34						3
25-29		2	1		2	3
20-24	5	1	3	2	1	8
15-19	11	5	17	16	4	7
10-14	21	23	38	17	16	
5-9	3	57	9	4	6	
Under 5 . . .	1	6				
Total . . .	43	94	68	39	29	21

TABLE 10

FREQUENCY DISTRIBUTIONS OF NONMETROPOLITAN STATE ECONOMIC AREAS FOR EACH REGION BY THE RURAL FARM LEVEL OF LIVING INDEX

Index	Regions					
	North-east	South-east	Middle States	North-west	South-west	Far West
135+.			6			
130-134 . . .			7			3
125-129 . . .	2		14	3	1	
120-124 . . .	9		6	7	1	5
115-119 . . .	11		7	10	2	4
110-114 . . .	9	1	6	12	1	6
105-109 . . .	4		6	2	2	3
100-104 . . .		2	3	4	2	
95-99	1	5	5	1	2	
90-94	3	6	4		2	
85-89	1	10	3		2	
80-84	1	9			2	
75-79		18	1		4	
70-74		16			3	
65-69	1	11			2	
60-64		11			1	
55-59		2				
Under 55. . .		3			2	
No measure. .	1					
Total . . .	43	94	68	39	29	21

TABLE 11

FREQUENCY DISTRIBUTIONS OF NONMETROPOLITAN STATE ECONOMIC AREAS FOR EACH REGION BY THE PROPORTION OF THE RURAL FARM LABOR FORCE WHO ARE UNPAID FAMILY WORKERS

Percentage	Regions					
	North-east	South-east	Middle States	North-west	South-west	Far West
28+		1			1	
25-27		2				
22-24		11				
19-21		7		1	1	
16-18		18	7	1	1	
13-15	2	22	10	6	5	
10-12	3	17	14	8	8	
7-9	16	7	26	16	9	
4-6	16	5	11	7	4	13
Under 4 . . .	6	4				8
Total . . .	43	94	68	39	29	21

TABLE 12

FREQUENCY DISTRIBUTIONS OF NONMETROPOLITAN STATE ECONOMIC AREAS FOR EACH REGION BY THE PROPORTION OF THE RURAL FARM LABOR FORCE IN NONFARM OCCUPATIONS

Percentage	Regions					
	North-east	South-east	Middle States	North-west	South-west	Far West
55+	3	2				1
50-54	2					2
45-49	6	1				
40-44	6	4	2			2
35-39	7	7	4			1
30-34	8	4	6		1	4
25-29	6	9	13	2	3	2
20-24	4	18	15	4	4	7
15-19		21	10	11	10	2
Under 15. . .	1	28	18	22	11	
Total . . .	43	94	68	39	29	21

TABLE 13

PERCENTAGE DISTRIBUTIONS OF NONMETROPOLITAN STATE ECONOMIC AREAS AMONG REGIONS FOR THE TOTAL GROUP OF NONMETROPOLITAN STATE ECONOMIC AREAS AND TEN SELECTED CONTROL GROUPS

Regions	Total Nonmet. Areas	Control Groups*									
		N-1	N-2	N-3	N-4	N-5	N-6	N-7	N-8	N-9	N-10
Northeast	14.6%	22.0%	22.4%	30.1%	26.1%	6.3%	3.1%	10.0%	10.1%	6.8%	18.8%
Southeast	32.0	21.1	9.4	12.3	8.0	30.4	29.1	10.0	8.9	11.9	6.8
Middle States	23.1	27.5	34.1	20.5	33.0	25.9	25.2	30.0	29.1	34.7	37.7
Northwest	13.3	8.3	8.2	12.3	10.2	15.8	19.7	20.0	27.8	21.2	20.6
Southwest	9.9	8.3	9.4	8.2	6.8	14.6	16.5	14.3	12.6	16.1	6.3
Far West.	7.1	12.8	16.5	16.4	15.9	7.0	6.3	15.7	11.4	9.3	9.7
Total	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
N	294	109	85	73	88	158	127	70	79	118	175

*The specific controls used are as follow:

N-1: Agricultural employment = 10-34.9%.
N-2: " " " " , Wholesale and retail trade employment = 13%+.
N-3: " " " " , Farm level living = 85-124.
N-4: " " " " , Nonwhite < 15%.

N-5: Agricultural employment = 30-54.9%.
N-6: " " " " , Nonfarm occupation < 25%.
N-7: " " " " , Unpaid family workers < 10%.
N-8: " " " " , Farm level living = 90-124.
N-9: " " " " , Nonwhite < 20%.
N-10: Nonwhite < 5%.

BIBLIOGRAPHY

Bogue, D. J. "An Outline of a Complete System of Economic Areas," American Journal of Sociology, LX, No. 2 (September 1954).

_____. State Economic Areas, Washington: Government Printing Office, 1951.

_____. The Structure of the Metropolitan Community. Ann Arbor: University of Michigan, 1950.

Dickinson, R. E. City, Region and Regionalism. London: K. Paul, Trench, Trubner and Co., 1947.

Duncan, O. D. "Regional Comparisons Standardized for Urbanization," Social Forces, 26, No. 4 (May 1948).

Forde, C. D. Habitat, Economy, and Society. New York: Dutton, 1950.

Freedman, R., et al. Principles of Sociology. New York: Holt, 1952.

Garnsey, M. E. "The Dimensions of Regional Science," Paper presented before the annual meeting of the Regional Science Association, New York City, December 1955.

Hawley, A. H. Human Ecology. New York: Ronald Press, 1950.

Isard, W. "Interregional and Regional Input-Output Analysis: A Model of a Space Economy," Review of Economics and Statistics, XXXIII (November 1951).

_____. "Regional Science, The Concept of Region, and Regional Structure," Paper presented before the annual meeting of the Regional Science Association, New York City, December 1955.

Isard, W. and Kavesh, R. "Economic Structural Interrelations of Metropolitan Regions," American Journal of Sociology, LX (September 1954).

James, P. E. "Toward a Further Understanding of the Regional Concept," Annals of the Association of American Geographers, XLII (September 1952).

Jensen, M. (ed.). Regionalism in America. Madison: University of Wisconsin Press, 1951.

Kish, L. "Differentiation in Metropolitan Areas," American Sociological Review, XIX, No. 4 (1954).

_____. "On the Differentiation of Ecological Units," Unpublished Ph.D. dissertation, University of Michigan, 1952.

Knapp, R. H. American Regionalism and Social Education. Cambridge: Harvard University Press, 1947.

Losch, A. The Economics of Location. New Haven: Yale University Press, 1954.

Mangus, A. R. Rural Regions of the United States. Washington: Government Printing Office, 1940.

McKenzie, R. D. The Metropolitan Community. New York: McGraw-Hill, 1933.

_____ (ed.). Readings in Human Ecology. Ann Arbor: University of Michigan, 1934.

Odum, H. W. Southern Regions of the United States. Chapel Hill: University of North Carolina Press, 1936.

Odum, H. W. and Jocker, K. C. (eds.). In Search of the Regional Balance. Chapel Hill: University of North Carolina Press, 1945.

Odum, H. W. and Moore, H. E. American Regionalism. New York: Holt, 1938.

Ogburn, W. F. "Regions," Social Forces, 15, No. 1 (October 1936).

Strong. H. "Regionalism: Its Cultural Significance," Economic Geography, XII (1936).

U. S. Bureau of the Census. Religious Bodies, 1936. Washington: Government Printing Office, 1941.

_____. Sixteenth Census of the United States. Vol. II, Population. Washington: Government Printing Office, 1943.

_____. U. S. Census of Population, 1950. Vol. II, Characteristics of the Population. Washington: Government Printing Office, 1952.

_____. Vote Cast in Presidential and Congressional Elections, 1928-1944. Washington: Government Printing Office, 1946.

U. S. National Resources Committee. Regional Factors in National Planning and Development. Washington: Government Printing Office, 1935.

Wardwell, Charles A. R. Regional Trends in the United States Economy. Washington: Government Printing Office, 1951.

Zimmerman, C. C. and DuWors, R. E., Graphic Regional Sociology. Cambridge: The Phillips Book Store, 1952.