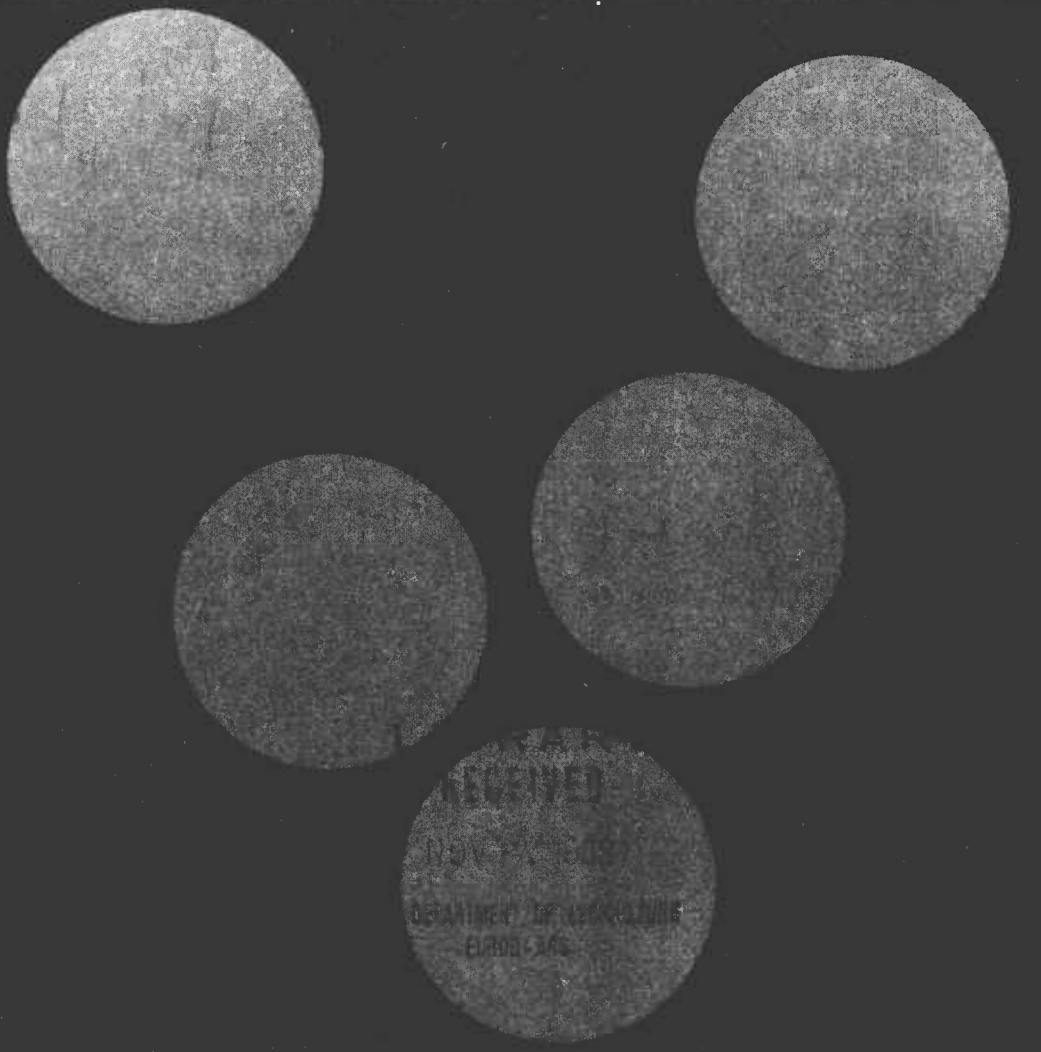

A STUDY OF INDUSTRY LOCATION USING MULTIPLE REGRESSION TECHNIQUES

AGRICULTURAL ECONOMIC REPORT NO. 140
ECONOMIC RESEARCH SERVICE
U.S. DEPARTMENT OF AGRICULTURE

CONTENTS

TABLES

FIGURES

Washington, D.C. 20250 August 1968

HIGHLIGHTS

This research was undertaken using a multiple regression model that would serve one or more of the following purposes: (1) determine the factors associated with location of individual manufacturing industries, (2) project the spatial distribution of manufacturing industries, and (3) provide assistance to local development planners in evaluating the prospects for their areas to acquire additional employment in specific manufacturing industries.

The 506 State economic areas were used as the grid system for location analysis. The four-digit (Standard Industrial Classification) manufacturing category was chosen for industry analysis. Fifty-three industries were analyzed in the study.

The dependent variable in the analysis is the change between 1947 and 1958 of employment in each of the 53 four-digit manufacturing industries. The independent variables are composed of 74 area characteristics.

A number of agricultural and rural attributes were included to determine whether they were associated with location of individual manufacturing industries. The spatial distribution of manufacturing industries was also estimated. Examples of the analysis follow:

1. Employment in the radio tubes industry (SIC 3662) in a given area was positively associated with the percent of the rural nonfarm population and the percent of class I (Census of Agriculture) farms in the area. This indicates a tendency for firms in this industry to locate in well-to-do rural areas.

2. Areas attracting growth in the men's, youths' and boys' underwear industry (SIC 2322) were relatively nonunionized and had high proportions of employment in bituminous coal mining and quarrying. The attraction to mining areas indicates a tendency to locate in rural areas.

3. Employment in tobacco manufactures was associated with growth in the children's and infants' dresses industry (SIC 2361). This variable, tobacco, stands as a proxy for a set of area characteristics common to both the tobacco and the children's dress industry. The children's and infants' dresses industry has a tendency to locate in Florida, Kentucky, New Jersey, North Carolina, Pennsylvania, and Virginia.

4. New employment in the wood household furniture, upholstered, industry (SIC 2512) tended to be in rural areas with poor farms. In general, the industry exhibited a strong tendency to concentrate in areas with a high proportion of total employment in the industry in 1947.

Local development workers can refer to table 2 to identify attributes that would be favorable or unfavorable for industry location in a particular area. Table 6 can then be used to identify industries that match the attributes of the area. Local policymakers could list the attributes of their area and scan table 6 for appropriate industries. Some attributes can be changed or modified by local policy.

Certain independent variables when considered separately were related significantly to growth in many industries. One variable, the percent distribution of prime military contracts in the United States in fiscal 1960, was positively associated with 16 industries and negatively associated with two. This variable was significantly related to more industries than any other. The variable, median income of females with income in 1960, was negatively associated with the five industries with which it was significantly related.

Results of this analysis should be used with care since high intercorrelations among the independent variables were found to exist. Some variables were dropped on this account. A variable that appears signiticant could in fact represent a number of variables.

A STUDY OF INDUSTRY LOCATION USING MULTIPLE REGRESSION TECHNIQUES

By

Robert G. Spiegelman 1/ 2/

I. INTRODUCTION

Many areas of the United States are failing to share in the national prosperity. They suffer chronic unemployment, low per capita incomes, and low rates of economic growth. To understand the reasons for differences among regions in economic performance, and to develop policies that may reduce the differences, it is necessary to understand how the market economy is operating to create these differences. In essence, national growth has a spatial dimension--economic activity exhibits location preference. These spatial preferences create differences among regions in per capita income and rates of growth. To a considerable extent, these differences reflect immobilities and other imperfections in the operation of the market economy.

The theory of regional economic development has long held that a region grows by expansion of so-called basic activities, those that essentially serve markets outside the region. One of the most important forms of basic activity is manufacturing. Knowledge of the factors determining the location of manufacturing establishments is essential if regional growth is to be understood or planned.

The objective of this study is to determine whether the econometric techniques of multiple regression could be successfully used to analyze the forces of location for manufacturing activity. A multiple regression model was developed that would serve one or more of the following purposes: (1) determine the forces associated with location of individual manufacturing industries, (2) project the spatial distribution of manufacturing industries, and (3) provide assistance to local development planners in evaluating the prospects for their areas to acquire additional employment in specific manufacturing industries.

The vast amount of regional data collected by the Bureau of the Census, the development of facilities for rapid computation of equations with a large number of variables, and the increased ability to interpret results of such equations through the advances in econometrics--all lead to the expectation that multiple regression can be used to increase understanding of the location process.

1/ This report was prepared by the author under contract with the Economic Research Service, U.S. Department of Agriculture. He is a Senior Economist with the Stanford Research Institute.

2/ Bernard R. Hoffnar, Area Analysis Branch, Economic Development Division, ERS, assisted in the planning and writing of the manuscript.

A few other investigators have attempted to apply this technique to highly aggregated data, but in this study the technique was applied to disaggregated data. The data used consisted of employment in manufacturing industries that were classified by four digits (1945 Standard Industrial Classification, or SIC numbers). Fifty-three industries thus classified were studied on a geographic grid to determine the significant area characteristics or variables influencing the area preferences of the various industries. The geographic grid consisted of the 506 State economic areas (SEA's) as designated by the Bureau of the Census.

To carry out the objectives of this study, it was necessary to contract with the Industry Division of the Bureau of the Census to provide data on employment in the manufacturing industries in each of the areas. Since these data are confidential (involving problems of disclosure), it was also necessary for the Bureau of the Census to prepare the moment matrices. 3/

The results of the research are presented in this report. They consist of a statement of the econometric model and the method of approach. The specific regressions computed for the 53 manufacturing industries are not included in this report but are available to interested persons. 4/ These specific industry results should be regarded as examples of how the model functions.

3/ The moment matrices were then used to compute the regression equations. This procedure, unfortunately, precluded analysis of the residuals.

4/ Write the Economic Development Division, ERS, U.S. Department of Agriculture, Washington, D.C. 20250, for copies of the regressions.

II. THE USE OF MULTIPLE REGRESSION TO ANALYZE CHANGES IN INDUSTRY LOCATION

The Contribution of Multiple Regression to the Analysis

Regression analysis is a statistical technique to estimate from empirical data the relationships between two or more economic variables. Multiple regression implies that more than two variables are involved. This technique has been employed by others to analyze the changes in the location of manufacturing industries, and to determine the variables associated with these changes. 5/ In past applications of multiple regression, the two-digit (SIC) manufacturing industries have been the ultimate level of disaggregation, and States have been used as the area grid. The present study departs from both criteria, using instead a four-digit SIC designation for manufacturing industries, and State economic areas as the grid. Justification for this departure is given in the following summary of our analysis of these criteria.

For the regression technique to be analytically useful, the population in which measurements of the dependent variable are made must be reasonably homogeneous--i.e., the different members of the population generate different values of the dependent variable only because there are variations in the specified set of independent variables and not because the members actually belong to different populations.

The two-digit classification of industry leaves much to be desired. This classification combines activities that are actually widely divergent in location requirements. Further, almost every State contains a mixture of area characteristics similar to that of its neighbors. Many differences will be averaged out. Thus, in terms of location characteristics (other than the spatial grid itself) the States do not form sufficiently distinguishable elements.

The assumptions underlying this study were as follows:

1. The four-digit classification of manufacturing industry comprises a reasonably homogeneous economic unit, with specific and identifiable location requirements.

2. The 506 State economic areas (SEA's) comprise sets of counties that are reasonably homogeneous in terms of economic and physical characteristics relevant to location decisions (other than those pertaining directly to site characteristics).

5/ W. Thompson and J. Mattila, An Econometric Model of Postwar State Industrial Development, Wayne State University Press, 1959; and Victor Fuchs, Changes in the Location of Manufacturing in the United States since 1929, Yale University Press, New Haven, 1962.

3. The economic and physical characteristics of an area are important in the location decisions of manufacturing establishments.

4. Linear multiple regression analysis can explain that part of the spatial distribution of growth by area which is in fact related to area characteristics--i.e., we can correctly specify the relevant independent variables and the form of the relationship.

Multiple regression can explain location patterns that result from the location decisions of individual owners and managers when these decisions are economically "rational" and are based upon past experience and knowledge of existing area characteristics. Regression can also explain location patterns that are created by a process of differential economic success. For example, if economic success is awarded to electronics plants that locate near universities, a close correlation of growth in electronics employment with distribution of universities may result either from the actual decisions made by entrepreneurs to locate their plants near universities, or by a process of differential success in which plants so located expand while plants located elsewhere fail to expand.

Some limitations of multiple regression are: First, it attempts to explain location patterns on the basis of past experience, and on the basis of a pattern of area characteristics that exist in the period covered by the analyses. To some extent, location decisions may be the result of expectations regarding future patterns of area characteristics for which we have no information and which cannot be reflected in our estimates. Second, location decisions are made by individual owners and managers with personal, social, or political motivations that may be "irrational" in terms of economic motivation.

Area Delineation

The delineation of SEA's as the geographic units for analysis was arrived at as follows: First, the decision was made to use geographic units smaller than States since one of the goals of the research was to investigate the validity of subregionalization as an analytic concept for location analysis. Although the county is a convenient substate regional grid, it is too small for purposes of industrial development analysis. It is a purely political creation with little relationship to the economic structure of the area. Thus, a metropolitan area or a unified agricultural region may comprise several counties, and so might areas of common geographic characteristics such as alluvial plains or mountain foothills. 6/

6/ The concept of the county as the unit of analysis runs counter to the current development thinking that stresses growth centers; these urban centers of new activity are likely to dominate and support an area containing several counties.

An attempt to properly structure development regions containing several counties has been made by Litton Industries for Appalachia. 7/ The Litton study listed several criteria for subregional delineation: (1) similarity of terrain and natural vegetation, (2) adequacy of existing and proposed transportation routes, (3) employment of a significant part of the labor force in a single industry or group of related industries, (4) trading and commuting patterns, (5) similarity in projected population trends from 1960 to 1990, (6) similarity in agricultural land use, (7) existing and potential recreational facilities, and (8) subregional delineations previously established by State planning agencies. The effort to compose such a list of criteria, establish a consistent weighting set (which was not done in the Litton study for Appalachia), and to apply these criteria throughout the Nation, was well beyond that of this project. Therefore, one of the existing national attempts at subregional structuring was chosen--the State economic areas, as constructed for the Bureau of the Census by Donald Bogue and Calvin Beale. 8/ According to these authors,

> The State Economic Areas . . . consist of groups of counties within a State (in some instances they are only a single county) which have similar economic and social characteristics. The boundaries of these areas have been drawn in such a way that each State is subdivided into a few parts, with each part having certain significant characteristics which distinguish it from other areas which it adjoins.

The criteria for combining counties to form SEA's are rather generally stated, except that each Standard Metropolitan Statistical Area (SMSA) is an SEA. The SEA delineation used in this study is that contained in the Census of Population for 1960. It differs from the original delineation by Bogue and Beale in that several counties designated as metropolitan areas or parts of metropolitan areas for the first time in the 1960 census, changed SEA designation.

Thus, we believe that the 506 SEA's form a reasonable grid system for location analysis for three reasons: (1) The set mitigates the undue influence of size on the relationships that would exist if counties were the unit, since the tendency is for metropolitan areas to be single counties, while the less populated nonmetropolitan counties tend to be combined to form SEA's. In other words, the size distribution of SEA's will be flatter than the size distribution of counties. (2) The SEA concept can retain most of the homogeneity of the county system while not fragmenting meaningfully integrated economic areas comprising more than one county. (3) This set provides a far larger number of observations than the individual 48 States, thus increasing

---

7/ Litton Industries, Preliminary Analysis for Economic Development Plan for the Appalachian Region, May 1965, Washington, D.C.

8/ Donald Bogue and Calvin Beale, Economic Areas of the United States, The Free Press of Glencoe, Ill., page xliii.

the likelihood of having statistically significant regressions.

Figure 1 (pages 32-33) is a map of the United States with SEA's outlined. SEA's that are also SMSA's are designated by letters and economic subregions and nonmetropolitan SEA's are designated by numbers. Table 1 shows the population distribution among SEA's. SEA's vary greatly in size--from rural, sparsely populated areas of less than 50,000 to the densely populated New York metropolitan area, with a population of over 10 million. Nevertheless, a distribution table shows that counties are grouped in SEA's in such a way as to result in most SEA's falling into a rather narrow size bank; more than half have populations between 100,000 and 300,000 with the median SEA having a population of about 210,000.

Table 1.--506 State economic areas grouped by 1960 population

SEA's	Population
Number	Thousands
3	5,000 and over
16	1,000-4,999
2	900-999
4	800-899
6	700-799
19	600-699
16	500-599
22	400-499
62	300-399
119	200-299
180	100-199
52	50-99
5	0-49
506	

Industry Selection

The four-digit manufacturing level was chosen for industry classification. The two-digit level aggregates many establishments that produce very different products, and that could be expected to have different location requirements. For example, electrical machinery (SIC 36) includes establishments producing complex telecommunications equipment as well as those producing large electric generators and those producing insulated wire and cable for use in electrical equipment. Establishments in most four-digit groups, however, are reasonably homogeneous in terms of the factors likely to influence their location. There is likely to be a sufficient number of establishments in each four-digit industry to provide enough observations for regression analysis. The seven-digit product breakdown was not available. The 53 industries used in the analysis are identified and described in Appendix A.

An "observation" of an industry exists if an area contains an establishment of the industry at either the beginning or the end of the period. For example, 100 observations means that 100 SEA's had an establishment of a given industry in 1947 and/or in 1958. Even at the four-digit level, the number of nonzero observations per industry is considerably less than the 506 SEA's for most industries. Numbers of nonzero observations for the 53 four-digit industries examined in this study are shown in the tabulation below:

Number of nonzero observations (SEA's containing establishments)	Number of four-digit industries
1 to 99	15
100 to 199	18
200 to 299	14
300 to 399	2
400 to 506	4
	53

Thus, 33 of the 53 industries examined were observed in fewer than 200 SEA's. A more refined industry classification would undoubtedly create serious restrictions on the analysis because of a paucity of observations.

Sixty industries were selected initially for study. The industries selected were largely those which had the highest proportionate growth in 1947-58, and which had total employment of 10,000 or more in 1958. Fast-growing industries with significant total employment were selected because these would be most likely to have experienced locational shifts lending themselves to explanation and because these are the most interesting industries to explore for purposes of predicting industrial growth in an area. One of the industries included (SIC 2025) experienced no total growth. This industry was included to determine whether special problems are encountered in analyzing such an industry.

The number of industries finally analyzed was 53 (Appendix A), as data problems or computational errors necessitated dropping seven of the industries from the analysis. The 53 industries analyzed in this study provide a cross-section of manufacturing industries with 16 of the 20 two-digit industries being represented (only SIC 21, 29, 30, and 31 are not represented in the group). The present sample is too small, however, to draw conclusions about which two-digit industry groups are most amenable to this kind of analysis. The results of the analysis are presented in chapter IV after description of the econometric model in chapter III.

III. AN ECONOMETRIC MODEL FOR ESTIMATING CHANGE IN LOCATION FOR 53 MANUFACTURING INDUSTRIES

The Basic Linear Multiple Regression Model

The variable to be predicted is called the "dependent variable," and the variables used for prediction are called "independent variables." (See table 2 for a list of variables.) The relationship between the variables may take any one of several forms--linear, quadratic, exponential, etc. It is assumed in this study that linear relationships are reasonable approximations of the form of the true relationships.

The method of computation is least squares, which minimizes the variance of the error terms; i.e., the method maximizes the portion of the total variance in the dependent variable that is explained by the independent variables.

The form of the multiple regression equation is:

$$Y_j = a + \sum_{i=1}^{k} b_i X_{ij} + u_j \qquad j = 1, 2, \ldots, N$$

where **a** is the Y-axis intercept; b_i 's are the regression coefficients for each of the k independent variables; u_j is the error term; and N is the number of observations.

If the least squares estimate is to be unbiased, the error term must be drawn from a distribution with the following characteristics:

$$E(u_j) = 0 \qquad \text{for all } j$$

$$E(u_j u_1) = \begin{cases} 0 & \text{for } j = 1 \\ s_u^2 & \text{for } j = 1 \end{cases}$$

That is, the expected value, or mean, of the distribution of the error terms is zero, and the variance of the distribution is independent of the values of the X's. The error term must be random--i.e., not attributable to missing variables, incorrectly specified relationships, or econometric problems such as serial correlation or heteroscedasticity (discussed later). 9/

The Structure of the Dependent Variable

The dependent variable is the change between 1947 and 1958 of employment in each of the 53 four-digit manufacturing industries. Change, rather than existing structure, is selected as the variable to measure, because change encompasses the factors currently influencing location decisions. The

---

9/ The inability to examine the residuals makes it difficult to determine the nature of the biases that exist in the regressions. However, some can be determined and will be discussed later.

Table 2.--Variables for multiple regression analysis

1/	1	Dependent variable: $\left(E_{iu}^{58} - E_{iu}^{47}\right) / E_i^{47}$
	2	Percent of urban population, 1960
	3	Percent of rural nonfarm population, 1960
	4	Median school years completed, persons 25 years old and over, 1960
	5	Percent of male professional, technical and kindred workers, managers, officials, and proprietors (except farm) in total employment (both male and female), 1960
	6	Percent of male craftsmen, farmers and kindred workers, and operatives and kindred workers in total employment (both male and female), 1960
	7	Median income of males with income, 1960
	8	Median income of females with income, 1960
	9	Percent of total farmland acreage in Class I commercial farm category, Census of Agriculture, 1954
	10	Percent in Class II
	11	Percent in Class III
	12	Percent in Class IV
	13	Percent in Class V
	14	Percent in Class VI
	15	Percent of total farmland acreage by type of farm in cash-grain category, 1954
	16	Percent in cotton
	17	Percent in vegetables
	18	Percent in fruit and nuts
	19	Percent in dairying

See footnotes on page 13.

20 Percent in poultry

21 Percent in livestock other than 19 and 20

22 Average value of land and buildings per farm, total for commercial farms, 1954

23 Percent of all farms owning a tractor, other than garden, 1954

24 Mean January temperature (°F), 1931-60

25 Mean July temperature

26 Population of largest city in SEA (in thousands), 1960. If an SEA had no city with population of 10,000, the SEA was assigned a value of 5

27 Population per square mile (land area, 1950), 1960

28 Percent change in population, 1950-60

29 (Dummy) 2/ SEA includes, or is a metropolitan area contiguous to, a county which includes a city with population of 500,000 or more in 1960. Yes = 1, No =0

30 (Dummy) 2/ Index of water availability. SEA has rivers with observed minimum daily flow (10-year period) not less than 100 cubic feet per second or located on Great Lakes. Yes = 1, No = 0

31 (Dummy) 2/ SEA has port on ocean, gulf, or commercially navigable inland waterway (9 feet or more controlling depth), or is contiguous to county with a major port and has good access to that port. Yes = 1, No = 0

32 (Dummy) 2/ State has inducements for industry location. Yes = 1, No = 0

Possible inducements in the form of legislation: (1) property tax exempted, (2) and (3) city and county bonds for facilities, (4) State financial assistance, (5) State-chartered private development corporation

All States qualified that had a check in at least 3 of the 5 possibilities

See footnotes on page 13.

33 Miles of surfaced roads per 1,000 population (by State), 1950

34 (Dummy) 2/ SEA has college or university with 1958 enrollment of 5,000 or more students. Yes = 1, No = 0

35 Percent distribution of prime military contracts in the United States, fiscal 1960

36 Relative extent of unionization (by State), 1953

Index of actual union membership in State. Number of workers in each industry category in State times national rate of union membership in each industry category

37 Power cost in mills per kwh for industry (State averages), taken as average cost to the chemical industry in 1958

38 Index of building construction cost, 1954, New York City = 100

Indexes for 160 cities. Judgment used to assign index to SEA for which no index is computed, based on proximity to cities with index.

1/ 39 Employment in each industry in 1947

40	Var. 3 X Var. 29	urbanization - big city
41	Var. 28 X Var. 29	percent change in population times big city
42	Var. 2 X Var. 31	urbanization - port
43	Total employment in thousands	
44	Var. 29 X Var. 32	State inducement - big city
45	Var. 7 X Var. 32	State inducement - income of males
46	Var. 34 X Var. 35	university - percent defense contracts
47	Var. 2 X Var. 24	percent urban - January temperature
48	Var. 2 X Var. 24 X Var. 34	percent urban population - January temperature - university enrollment

See footnotes on page 13.

Table 2 (continued)

49 Average wage rate in manufacturing in area

50 - 54

Percentage of employment (times 10) in two-digit mining industries, 1954

50 SIC 10, metal mining
51 SIC 11, anthracite mining
52 SIC 12, bituminous coal and lignite mining
53 SIC 13, crude petroleum and natural gas extraction
54 SIC 14, mining and quarrying of nonmetallic minerals, except fuels

1/ 55 Employment in each industry in 1947 divided by total employment in 1947 in thousands

56-75 Percentage of employment (times 10) in two-digit manufacturing industries, 1947

56 SIC 20, food and kindred products
57 SIC 21, tobacco manufactures
58 SIC 22, textile and mill products
59 SIC 23, apparel and other finished products made from fabrics and similar materials
60 SIC 24, lumber and wood products, except furniture
61 SIC 25, furniture and fixtures
62 SIC 26, paper and allied products
63 SIC 27, printing, publishing, and allied industries
64 SIC 28, chemicals and allied products
65 SIC 29, products of petroleum and coal
66 SIC 30, rubber products
67 SIC 31, leather and leather products
68 SIC 32, stone, clay, and glass products
69 SIC 33, primary metal industries
70 SIC 34, fabricated metal products, including ordnance
71 SIC 35, machinery, except electrical
72 SIC 36, electrical machinery, equipment and supplies
73 SIC 37, transportation equipment
74 SIC 38, professional, scientific and controlling instruments; photographic and optical goods; watches and clocks
75 SIC 39, miscellaneous manufacturing industries

1/ 76 Dependent variable: $E_{iu}^{58} - E_{iu}^{47}$

1/ Industry specific variables.
2/ Appendix B contains a discussion of the use of dummy variables.

Sources for Multiple Regression Analysis Variables

Variable No.	
2 to 8	U.S. Bureau of the Census, Census of Population, 1960; State Economic Areas, PC (3) 1A. Var. 1 and 2 - Table 5; Var. 3 - Table 3; Var. 4-7 - Table 4.
9 to 21	U.S. Bureau of the Census, Census of Agriculture, 1954, Vol. 1 (32 parts). Var. 8-13 - Ch. C, Economic Area Table 1; Var. 14-20 Ch. C, Economic Area Table 4.
22 and 23	U.S. Bureau of the Census, Census of Agriculture, 1954, Vol. 1. Var. 21 - Ch. C, Table 1; Var. 22 - Ch. D, Table 2.
24	U.S. Bureau of the Census, County and City Data Book for Metropolitan Areas, 1956.
25	U.S. Dept. of Commerce, Weather Bureau, Decennial Census of United States Climate, Monthly Averages for State Climatic Division, 1931-1960.
	Variables 24 and 25: For nonmetropolitan areas--temperature of State climatic division most closely approximating the SEA; in cases where more than one climatic division is averaged
26	Bogue, Donald and Calvin Beale, Economic Areas of the United States, The Free Press of Glencoe, Ill., 1961.
27 and 28	Bogue and Beale, op. cit., when there was no change in SEA; County and City Data Book for Metropolitan Areas, 1956 and 1962, when SEA changed.
29	U.S. Bureau of the Census, Census of Population, 1960. List of 21 cities with population of 500,000 or more.

Table 2 (continued)

30 U.S. Geological Survey, Surface Water Supply of the United States, 1959, Water Supply Papers 1621 to 1638, Government Printing Office, 1960.

31 Ibid., "United States Water Resource Development, Jan. 1963" (map). (Judgment introduced to eliminate reservoir areas used for irrigation, and river areas used for recreation. Minimum required flow was based on statement in "The Big Decision," Chemical Week, July 20, 1963, p. 78.)

32 Business Week, Dec. 16, 1961, p. 127.

33 U.S. Bureau of Public Roads, Highway Statistics, 1955, "Mileage of Public Roads and Streets."

34 College Placement Directory, O. Zimmerman and I. Lavine, eds., 3rd ed., Industrial Research Service, Dover, N.H., 1958.

35 Isard, W. and J. Gauschow, Awards of Prime Military Contracts by County, State and Metropolitan Area of the United States, Fiscal 1960, Regional Science Institute, Philadelphia, Pa., 1962.

36 Fuchs, V. R., Changes in the Location of Manufacturing in the United States Since 1929, pp. 91-2, 354-5.

37 Chemical Week, July 20, 1963, p. 77, citing Census of Manufactures.

38 New Building Cost Calculation Index, April 1963, F. W. Dodge Corp., New York City.

49 Unpublished data provided by The Industry Division, Bureau of the Census.

existing structure accumulates a history of location decisions, many of which were based on conditions and requirements no longer valid--for example, the location at railheads of industries now shipping by truck, the location near coal mines of industries now using natural gas, the location near metal mining regions now unproductive, etc. There is, however, a danger in analyzing change rather than structure; change in a short period of time usually affects only a small portion of the entire industry. Random influences which are minor in terms of the whole industry are likely to be large factors in terms of the portion of the industry affected by change. In this situation, the proportion of explained variation will be lower when analyzing change than when analyzing existing structure. Nevertheless, the variables that are found to be significant when analyzing change are apt to be more meaningful for predicting present and future location decisions.

Structure of Equations for Prediction of Location Changes

The prediction of change in location starts with the premise that if the spatial allocation of an industry was optimal in the base year, and if no forces acted to alter the conditions for optimality, then there will be balanced growth in each area (each area will have the same growth rate as the average growth rate for the Nation), and the spatial distribution of employment in the given industry will not change.

If growth is not balanced, then equilibrium is achieved by a constant entering the regression, and/or by the introduction of other variables to "explain" the divergence. The basic regression model, termed the "absolute" form of the regression in this study, is as follows:

$$(1) \qquad E_{iu} = E_{iu}^{58} - E_{iu}^{47} = a_o + a_1 E_{iu}^{47} + a_2 E_i^{47} + \sum_{j=3}^{N} a_j X_{ij}$$

$u = 1 \ldots 53$ (industries)

$i = 1 \ldots 506$ (areas)

$j = 3 \ldots 77$ (area characteristics)

E_{iu}^{47} = base year (1947) employment in industry u in area i

E_{iu}^{58} = terminal year (1958) employment in industry u in area i

E_i^{47} = base year total employment in area i

X_{ij} = characteristic j of area i

There are 53 regressions, one for each industry, run as a cross-sectional analysis over all areas in which there is some employment in the industry in either the base or terminal year. For each regression, the number of observations varies depending upon the number of areas containing the industry. Areas without any employment in the specific industry were omitted as observations for two reasons: (1) Most industries were located in fewer than half of the State economic areas (see chapter II). For these industries, the zero

observations would be so large as to dominate the regression making it difficult to discern characteristics of areas with positive employment, and (2) in terms of change, there are two kinds of areas that would record a zero value for the dependent variables--those with the same number of employees in the base and terminal years, and those with no employees in either year. If these two kinds of areas have significantly different characteristics, the ability of the regression to predict change would be reduced.

The constant (a_o) (if any) and the coefficient of the base year distribution of industry u's employment (E^{47}_{iu}) represent the internal forces acting to cause concentration, dispersion, or spatially balanced growth. Spatial concentration occurs if geographic change in employment in a industry over time results in the areas with the highest employment in the base year increasing their share of employment. Concentration may be measured in terms of the change in the entire size distribution of employment by area.

THEORETICAL CUMULATIVE DISTRIBUTION OF EMPLOYMENT BY SIZE OF AREA

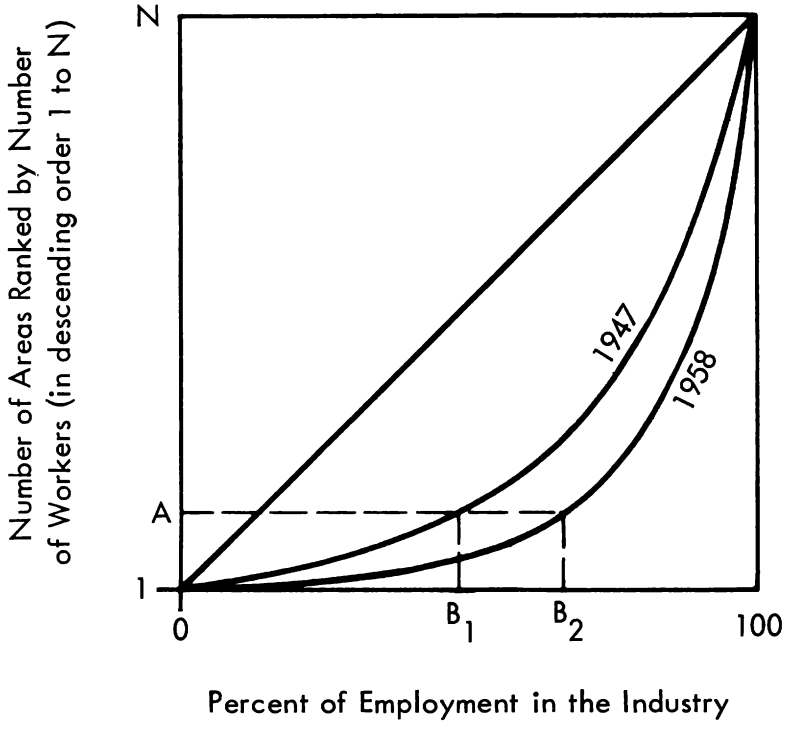

Figure 2

Thus, concentration may be measured by comparing cumulative distributions of industry employment in two time periods. Areas are ranked according to the number of workers in an industry, with the area having the largest number being first. A curve may then be drawn, showing at each point in this ranking the percent of employment accounted for by this and higher ranking areas. An example, demonstrating concentration, is shown in figure 2. In this figure, concentration is evidenced because a given number of the largest areas (A) had a higher percentage of industry employment in 1958 (B_2) than in 1947 (B_1). Dispersion would be shown by the reverse situation. Concentration is the end result of the location process and can be measured by direct investigation of employment statistics. Regression analysis, however, can be used analytically to discover some of the conditions underlying concentration.

The second variable in the regression (E_i^{47}) attempts to explain the tendency for growth of an industry to occur in areas that are large in total employment. This tendency is termed the area "size" effect and is to be distinguished from the "concentration" effect mentioned above. The latter explains how industry growth is related to its own base year distribution, while the former explains how industry growth is related to the base year distribution of total employment or population. If the area size effect dominates, an area with total employment 10 times that of a smaller area will experience growth in industry u 10 times that of the smaller area. Size effect must be isolated to find other relationships that are relevant for smaller areas. The presence of concentration or size effects creates a serious statistical difficulty in that the regression may suffer from heteroscedasticity. 10/

Besides internal forces of concentration and area size effects, there are various external forces acting to change industry location patterns (in the case of balanced growth, the net effect of such forces is zero). These external forces are captured by variables that represent area characteristics such as weather, water flow, port facilities, wage rates, and the presence of other industries.

Finally, there are a variety of forces that affect location, but that are not captured in the regression. Hopefully, these represent random deviation from the regression and are represented as an error term with a mean of zero and normally distributed about the regression line. Unfortunately, many of the forces that are included in the regression do not operate in simple linear fashion.

10/ One of the assumptions underlying the regression is that the deviations from the regression are homoscedastic--i.e., the unexplained variance is uncorrelated with the independent variables. We expect, however, that where size effects are present, the regression will be heteroscedastic--i.e., the unexplained variance is positively correlated with size. In this case, the absolute variance would be greater for large values of total employment than for small values if the regression were partitioned by size of area. See J. Johnston, Econometric Methods, McGraw-Hill, New York, 1963, pages 207-11.

If size effects are strong, they are apt to dominate this form of the regression so that other forces relevant to location are "swamped." To alleviate this condition and to correct for heteroscedasticity, a second form of the regression was also used, in which the dependent variable and independent variables representing concentration and area size effects were normalized for total employment. In the first form of the regression, the dependent variable is the absolute change in employment in industry u in area i; in the second form of the regression, the dependent variable is the change in employment in industry u in area i relative to total employment in area i in the base year. This regression, called in this study the "normalized" form, is as follows:

$$(2) \qquad \frac{\Delta E_{iu}}{E_i^{47}} = \frac{E_{iu}^{58} \quad E_{iu}^{47}}{E_i^{47}} = \hat{a}_o + \hat{a}_1 \frac{E_{iu}^{47}}{E_i^{47}} + \sum_{j=2}^{N} \hat{a}_j X_{ij} \; ; \quad \begin{matrix} u = 1 \ldots 53 \\ i = 1 \ldots 506 \end{matrix}$$

The following sections contain a description of each of the three sets of forces acting on location change, and a discussion of how multiple regression can be used to isolate, in an approximate way, the internal forces and size effects.

Measuring Internal Forces of Concentration

The forces determining the spatial allocation of industry may be divided into internal and external forces. The external forces are those of area attraction and repulsion, represented in the regression by area characteristic variables. The internal forces are economies of scale and other conditions specific to the industry, rather than the area, that cause the industry to cluster or disperse. The internal forces are isolated best in the normalized version of the regression. The constant in this form of the regression is the average increase in industry u's employment as a proportion of total area employment in the base year. Since a high value of the constant means that on the average both large and small areas have the same ratio of growth in the industry to base-year total employment, a high value of the constant reflects the tendency for absolute employment in the industry to concentrate in large areas. In addition to the constant, forces of concentration are measured by the coefficient of base-year employment.

To determine whether the values of the coefficients of the regression are suggesting internal forces of concentration, the regression must be transformed so that the dependent variable becomes the national or average growth rate of the industry over the period.

The absolute form of the regression may be directly transformed as follows. Substitute the mean values for the size and concentration variables in equation (1):

(1A) $$\Delta\bar{E}_u = a_o + a_1 \bar{E}^{47} + a_2 \bar{E}^{47} + \sum_{j=3}^{N} a_j \bar{X}_j$$

Divide through by the mean base-year employment in industry u,

(1B) $$\frac{\Delta\bar{E}_u}{\bar{E}^{47}_u} = \frac{a_o}{\bar{E}^{47}_u} + a_1 + a_2 \frac{\bar{E}^{47}}{\bar{E}^{47}_u} + \sum_{j=3}^{N} a_j \bar{X}_j$$

On the left side is the mean change in employment in industry u relative to the mean base-year employment. This is equal to the change in employment in industry u in the United States, divided by the total U.S. employment in industry u in the base year, which is the growth rate of the industry as demonstrated below:

$$\frac{\Delta\bar{E}_u}{E^{47}_u} = \frac{\Sigma E_{iu} / N}{\Sigma E^{47}_{iu} / N} = \frac{\Sigma \Delta E_{iu}}{\Sigma E^{47}_{iu}} = \frac{\Delta E_u}{E^{47}_u}$$

Theoretically, the constant and the first term of the regression may be interpreted as representing the internal forces of concentration; after transformation, their sum will show the strength of these internal forces. If the sum exceeds the actual growth rate of the industry, this shows that internal forces are tending to cause concentration, whereas if their sum is less than the actual growth rate, this indicates a tendency toward dispersion since base year employment is not attracting its proportionate share of new growth.

Unfortunately, the constant in the absolute form cannot be interpreted as representing only the internal or exogenous forces since it also captures other attributes of the regression such as nonlinearity. To confirm that the tendencies implied in the transformed absolute equation exist, a similar transformation is performed on the normalized form of the regression. The results of the first transformation are considered to be valid only if verified in the transformation of the normalized form of the regression.

To make this transformation, however, requires that strong assumptions be made concerning the nature of the error term. If the normalized regression is as stated above in equation (2), then multiplying through by base-year total employment (E^{47}_i) yields the following equation, with the error term (u_i) included:

$$\Delta E_{iu} = \hat{a}_o E^{47}_i + \hat{a}_1 E^{47}_{iu} + \left(E^{47}_i \sum_{j=3}^{N} \hat{a}_j X_{ij} \right) + u_i E^{47}_i$$

The assumptions regarding the error term are that the mean of the distribution of the error term is zero and that there is no correlation between the error term and base-year total employment. With these assumptions, the mean values may be substituted, yielding the following:

$$\Delta\bar{E}_u = \hat{a}_o \bar{E}^{47} + a_1 \bar{E}_u^{47} + \left(\bar{E}^{47} \sum_{j=3}^{N} a_j \bar{X}_j \right) + 0$$

Dividing through by $\bar{E}_u^{47}$ yields the normalized regression transformed so that the industry growth rate is on the left-hand side of the equation:

$$\frac{\Delta\bar{E}_u}{\bar{E}_u^{47}} = a_o \frac{\bar{E}^{47}}{\bar{E}_u^{47}} + \hat{a}_1 + \left(\frac{\bar{E}^{47}}{\bar{E}_u^{47}} \sum_{j=3}^{N} a_j X_j \right) + 0 \quad .$$

The sum of the first two terms may be interpreted as representing the internal forces of concentration. Since there are grave difficulties in interpreting the results of transforming either the absolute or normalized versions of the regression, the results of the transformations are relied upon only when they are consistent. 11/ Whether or not the industry actually concentrates or disperses depends upon the net effects of all forces and can be directly measured by comparing the 1947 and 1958 distribution. Of interest in this study is the analytic value of determining whether or not internal forces are causing concentration or dispersal.

The results of the analysis are presented later in a summary table for the 53 industries analyzed. The transformations gave consistent answers for 43 of these industries. Internal forces tended to create spatial concentration in 10 of these industries, balanced growth in nine, and relative dispersion in 24. Thus, internal forces caused a majority of these industries to seek new locations or at least to grow fastest in the areas with the least amount of employment in the base year. For the industries analyzed herein, at least, this finding is not consistent with the hypothesis that industry tends to grow most in the areas where it is already located.

Measuring the Area Size Effects

In the ideal case, the size effects would be measured in the absolute form of the regression by the coefficient of the variable representing the distribution of total employment (in thousands). A coefficient of 0.5 would mean that each 1,000 workers in an area are associated with an expected increase of 0.5 worker in industry u. Unfortunately, the regressions do not permit such a pure interpretation of the coefficient, for two reasons:

11/ These difficulties of interpretation could have been reduced if a third form of the regression using the base-year employment in industry u in area i for normalization had been estimated. To run this third set of regressions would have required computations of completely new cross-product matrices by the Bureau of the Census.

(1) Where there are large size effects on the change in employment, there is usually high intercorrelation between the 1947 distribution of industry u and total employment, and (2) several of the variables representing other area characteristics have size effects. (Principal examples are distribution of military contracts--variable 35, and presence of a large city--variable 29.) Thus, the presence of multicolinearity means that size effects are captured in several of the coefficients, and not just in the coefficient of total employment.

The most direct evidence of size effect is provided by the drop in R^2 between the absolute and normalized versions of the regression. Since the normalized version is created from the absolute version by eliminating the size effect, the difference between the correlation of the absolute version and that of the normalized version is a measure of the size effects. This measure of size effects, however, is neither sufficient nor necessary. It is not necessary because the R^2 may not decline if the size effects coincide with the structural effects. For example, if the change in employment in an industry is correlated with the percent of employment in fabricated steel manufacturing (SIC 34) and per capita incomes, and these in turn are highly correlated with size, then the normalized version will not show a decline in R^2. In this case, size effects would be determined by the significance of the size variable a_2 in the absolute form of the regression. A drop in R^2 is usually sufficient evidence of size effects but it may also reflect statistical difficulties with the regression--that is, improperly specified variables. Such a drop is best verified by the presence of some significant variables that include size effects.

The results of the analysis of size effects for each industry are shown later in a summary table. A condensation of these results is given in the following tabulation:

Size effects	Industries
Strong	19
Moderate (positive)	16
Negligible	14
Moderate (negative)	3
Undetermined	1
	53

The above tabulation indicates that location of growth in most industries was affected by the size of areas. A strong size effect indicates that most of the growth in the industry tends to go to large areas. In many cases where size effects are strong, there is orientation toward final consumer or military markets. Of the 19 industries that show strong size effects, six can be classified as being oriented to consumer markets (SIC 2661, 2711, 2741, 3271, 3441, and 3442); four others (SIC 3559, 3613, 3661, and 3811) can be classified as being oriented toward military markets, as represented by the distribution of military contracts--a factor which is highly intercorrelated with size of area. The strong size effects for the other nine industries have explanations not apparent from the regressions, although some of these may also

be market-oriented.

Selection of Area Variables to Represent External Forces of Location

Previous studies of industry location suggest various area characteristics that might be expected to influence industry location. Essentially, these characteristics must reflect one or more of the three major causes of location listed by E. M. Hoover: 12/ (1) Transportation costs of final products and material inputs, (2) production costs that vary geographically, and (3) agglomeration factors. The agglomeration factors are a host of external economies and complimentarities that cause establishments of one or more industries to cluster in a limited number of places.

Transportation costs cannot easily be incorporated into the regression analysis because such costs are not part of the generally available data on areas. Nevertheless, an association between the distribution of a given industry and the distributions of industries that are major purchasers or suppliers indicates the interindustry transport linkage; and an association between the distribution of an industry and the total population or income indicates the orientation of that industry to consumer markets.

Whether or not a transport-oriented industry has a distribution of employment that correlates highly with the distribution of its market depends in large measure on the size of the establishments, the geographic extent of the market for its products, and the nature of competition in the industry. It is not inconsistent for all the firms selling a competitive product in a national market to cluster in one cost-minimizing location. For such an industry, there would be low correlation between industry distribution and population distribution. A high correlation between industry distribution and market distribution in our model indicates that the products of the industry travel relatively short distances. Transport orientation of industries not located close to their markets or sources of supply could be determined only by highly complex models using large quantities of marketing and transport cost data. The independent variables that suggest transport orientation to final consumer markets are variables 26-29, 33, and 43 in table 2. The last variable is the direct size variable, discussed above. The others are also size-related.

There are other variables that describe the characteristics of the area's population and that may suggest refinements of the market orientation in terms of income and education levels, or alternatively may suggest some of the characteristics of populations associated with agglomeration. These population characteristics are incorporated in independent variables 2-8.

12/ E. M. Hoover, Location of Economic Activity, McGraw-Hill, New York, 1948.

Many industries may be oriented toward rural populations or farm markets. In addition to the absolute size variables, the wealth and degree of commercialization of the farm community may be factors influencing location. These characteristics are captured in variables 9-14, 22, and 23.

Transport orientation toward intermediate markets or toward suppliers of major material inputs can be fully captured only by use of industry specific variables which were not employed in this study. Several aggregative variables were employed that can suggest this type of orientation. In constructing the resource and industry variables, it was decided to normalize for size because of the expected high correlation between the absolute distribution of major industry categories and total size. Thus, the agricultural variables are in terms of percent of farmland in different major categories of farms. The mining and manufacturing industry variables are in terms of the percent of employment in the relevant two-digit industry. In many cases, orientation to a particular industry structure may suggest inter-industry transport relations, or may suggest other factors. For example, the orientation of the apparel industry to coal mining suggests the shifts in the industry that took place to take advantage of low cost and unemployed labor available in the coal mining areas, and not the need of the apparel industry for coal. On the other hand, the orientation of changes in employment in machine shops (SIC 3599) to areas with high proportions of employment in manufacturing of transportation equipment (SIC 37) and instruments (SIC 38) does suggest the market orientation of the machine shops industry. The variables capturing the resource and industry structure of areas are as follows:

For agriculture, variables 15-21
For mining, variables 50-54
For manufacturing, variables 56-75

Geographic Variation in Production Costs

Although concern for total profits leads an enterprise to consider the entire mix of factor inputs in selecting a location, some industries may show an orientation to a particular set of factor variables because these are very important in total costs (such as the emphasis of aluminum production on cheap power), and/or because the factor price tends to vary considerably among areas. The most important factor input that has affected industry location is labor. Several variables, in addition to total employment, which can mean either market or labor orientation, are introduced to capture the variations in quantity and quality of the available labor force, and to capture the variations in the level of hourly wage rates. These are variables 5, 6, 36, and 49.

The only other factor whose variation in price can be expected to affect location significantly is power. Differences in power costs are captured in variable 37.

For industries in which investment in plant is very large relative to variable costs, it is possible that differences in construction costs could be a factor in determining location. These differences are captured in variable 38.

Some industries have minimum requirements that must be met if the industry is to exist in the area. These take the form of either/or propositions, and are best represented by dummy variables. 13/ Two of the most important are the presence of process water in sufficient quantity and access to a port facility for water shipment. These two factors are captured as dummy variables 30 and 31. Variables for railroad and highway access are not included because they are sufficiently ubiquitous to provide little basis for differentiation by SEA.

Agglomeration Factors

There has been considerable discussion in the literature concerning the tendency for certain industries to cluster in a limited number of places when this clustering cannot be explained in terms of transportation or factor costs. Certain area features, either singly or in combination, may serve as the impetus for clustering. Some of these have been captured in this study, but others undoubtedly have not. Urbanization, or metropolitanization, represents a way of life and method of organizing human activity that in itself can be an agglomerating force not directly associated with the size of the population in the area. On the other hand, for some industries urbanization may be a deterrent to location. Besides urbanization, several other factors possibly influencing location have been captured in this study. These are: the weather, the presence of a university or college, and State inducements. Since it is possible that location is affected by certain areas possessing a desirable combination of features, an attempt was made to capture some likely combination in the form of interacting variables. The single and interacting variables not related to transportation or production costs that may influence location are 24, 25, 32, 34, 40-45, 47, and 48.

A variable was introduced to capture the effect of the distribution of prime military contracts (variable 35). This is regarded as both cause and effect of many location decisions. If an industry is heavily weighted by firms with prime military contracts, then the correlation is spurious. If the industry is heavily weighted with firms having military subcontracts, then the correlation suggests the tendency for military subcontractors to locate near the source of prime contracts. If the industry does not sell directly or indirectly to the military, then a high correlation suggests that the presence of defense contracts also creates an environment conducive to other industries--i.e., agglomeration effects. Thus, a high correlation with prime defense contracts would need to be further analyzed to determine the nature of the markets served by the industry before conclusions as to the

---

13/ Dummy variables take discrete values, and are used to represent qualitative differences. Their use is discussed in appendix C.

meaning of such relationships could be drawn.

Before we turn to the investigation of the multiple correlations, attention should be given to how the various independent variables are related. Since most of the variables relate to aspects of the socioeconomic environment of the area, the inherent interdependence of socioeconomic conditions is reflected (creating problems of multicolinearity.) 14/

The Intercorrelation Matrix and Correlation of Socioeconomic Phenomena

Table 3 is a list of those pairs of independent variables that are intercorrelated, having a correlation coefficient of 0.6 or greater. The table indicates that in several respects high intercorrelations restrict our ability to determine the impact of variables. For example, urbanization (variable 2 in table 2) is highly correlated with the median income variables (7 and 8). Thus, it is impossible to differentiate between the general affects of urbanization and the more specific effects of income levels on industry location. Further, median income of males (7) is highly correlated with wage rates in manufacturing (49), making it difficult to differentiate between the effects of income as a measure of consumer markets and the effect of wages as a measure of labor costs.

The variables representing the conditions and wealth of the farm community are also highly intercorrelated. A high proportion of Class I farms is associated with a high value of farm lands and buildings; a high proportion of Class II farms is associated with a low proportion of Class V and VI farms; a high proportion of Class V farms is associated with low values of farm lands and buildings; and a high proportion of Class VI farms is associated with a low percent of farms owning tractors.

These associations are logical and indicate that one or two variables may stand as proxy for the status of the farm community in the area.

Table 3 also shows that with interacting variables the whole is usually highly correlated with one of its parts. Thus, the interaction of percent urban population and January temperature (47) is highly correlated with percent urban population (2) alone; and urbanization-large city (40) or urbanization-port facility (42) are highly correlated with presence of large city (29) or port facility (31) separately; and so on. The interacting variables are often stronger than the separate variables in explaining industry location.

14/ See J. Johnston, op. cit., pages 201-07.

Table 3.--Independent variable pairs with high correlation

Independent variable pairs	Correlation of pairs
2 & 3	-0.9070
2 & 7	0.6915
2 & 8	0.7080
2 & 47	0.7831
3 & 47	-0.6788
4 & 5	0.6515
4 & 7	0.7588
7 & 8	0.7864
7 & 49	0.8420
9 & 22	0.7795
10 & 13	-0.7185
10 & 14	-0.7233
10 & 23	0.6681
11 & 23	0.7489
13 & 14	0.8082
13 & 22	-0.6620
14 & 23	-0.7072
23 & 24	-0.6554
26 & 35	0.6775
29 & 39	0.9960
29 & 40	0.8141
29 & 41	0.8050
31 & 42	0.8815
31 & 43	0.8658
32 & 45	0.9353
34 & 48	0.8847
39 & 40	0.7598
39 & 41	0.8014
40 & 41	0.6617

Additional insight into the nature of the intercorrelation matrix can be gained by examining the linear association of a variable with a set of other variables in the matrix. Of particular interest here is the extent to which the industrial factors of an area are interrelated. Ten of the 19 two-digit manufacturing industries show a significant degree of interdependence--i.e., the percent of employment in each of these 10 industries can be explained to a significant degree by a set of other variables in the matrix. (Note: For significance in this test, we select a set of three independent variables with a coefficient of multiple correlation of at least 0.5 which tests as significant at the 1 percent level.)

The 10 manufacturing industries whose relative strength in an area is significantly related to other variables in the matrix are SIC 23, 24, 27, 33, 34, 35 36, 37, 38, and 39. Table 4 shows the best three independent variables for each of these industries. The results are informative for their own sake. The interdependence of the metalworking industries is demonstrated in this table; that is for each of the industries from SIC 33 to SIC 39, one or more of the best three independent variables is the percent of employment in another metalworking industry. For example, a high percent of employment in metal fabricating (SIC 34) is highly associated with a high percent of employment in primary metals (SIC 33), nonelectric machinery (SIC 35), and miscellaneous manufacturing (SIC 39). A hig percent of employment in instrument industries (SIC 38) is associated with a hig percent of employment in electrical machinery (SIC 36) and transportation equipment (SIC 37).

A high percent of employment in the apparel industry (SIC 23) has an unexpected set of associations: A high percent of area farms in poultry raising (20 presence of a large city (29), and a high percent of employment in anthracite mining (51). These associations may reflect both the historic association of th apparel industry with New York and other big cities and the newer tendency for t industry to move to areas of lower wages and excess labor. The wood products industry (SIC 24) is associated with areas that are predominantly rural (3), tha have low-cost power (37), and that have a high proportion of employment in the paper and paper products industry (62). The publishing industry (SIC 27) is strongly associated with big cities, urbanization, and the nonelectric machinery industry (SIC 35).

Many other sets of interrelationships exist. Table 4 shows three others of particular interest. A high rate of population growth in the decade of the 1950 (28) was closely associated with areas having the following characteristics: Hi level of adult education (4), high median income of males (7), and high January temperature (24). A high level of male income (7) was associated with areas having the following characteristics: Highly urban (2), high level of adult education (4), and a high percent of employment in metal fabricating industries (70 Thus, we see a general association between high levels of growth, high incomes, urbanization, and the metalworking industries. In addition, the emphasis on hig winter temperature indicates the substantial growth that occurred in the southwe crescent of the United States from California to Texas. A further complexity is introduced by showing that the distribution of prime military contracts is interrelated with the same system. The percent distribution of prime military is ass ciated with the following: Size of the central city (26),* and the interaction urbanization, January temperature, and the presence of a university (48).†

---

* This variable is highly correlated with the distribution of total employment indicating a strong tendency for military contracts to be distributed in accordance with the distribution of population.

† The two independent variables have a correlation coefficient of 0.70. However, this is raised to 0.72 when the presence of a university is added as a separate third variable. The sign of this variable is negative indicating that the interaction assumed in our variable is somewhat inappropriate. More weight should have been given to urbanization and weather and less weight to the presence of a university.

Table 4.--Selected linear combinations from intercorrelation matrix 1/

Selected dependent variable	Three best independent variables 2/	R^2	F-value
59 (SIC 23)	20, 26, 51	.34	83.6
60 (SIC 24)	3, 37 (-), 62	.25	55.0
63 (SIC 27)	2, 26, 71	.46	142.5
69 (SIC 33)	6, 65, 70	.30	71.3
70 (SIC 34)	69, 71, 75	.49	159.8
71 (SIC 35)	10, 63, 70	.44	126.8
72 (SIC 36)	71, 73, 75	.25	55.0
73 (SIC 37)	49, 70, 74	.29	67.3
74 (SIC 38)	5, 72, 73	.40	108.4
75 (SIC 39)	8, 19, 20	.34	83.6
7	2, 4, 70	.74	477.4
28	4, 7, 24	.58	226.0
35	26, 34 (-), 48	.52	183.3

1/ See table 2 for description of variables.
2/ (-) indicates that regression coefficients are negative.

We will refer again to these complex sets of interrelationships between major socio-economic variables in explaining the growth determinants for individual industries discussed in the following chapter and appendix A and comprising the central issue of this report.

Selection of Independent Variables and Construction of Final Regression for Each Industry

As mentioned above, the external forces affecting location are a set of area characteristics that were found to significantly affect the value of the dependent variable. These characteristics were selected from a set of 74 variables described in detail above. It is obvious that a meaningful regression could not be constructed using all of the 74 variables available for the purpose.

Thus, it was essential to select a set of feasible independent variables to test for significance. The selection was accomplished as follows: (1) A correlation matrix was calculated which provided the simple correlations between the dependent variable and each of the 74 independent variables. (2) Any independent variable that had a correlation coefficient of 0.2 or higher with the dependent variable was tentatively selected. (3) Additional independent variables were selected which had low simple correlations with the dependent variable but which a priori knowledge suggested might be a factor in the location of the industry. (4) The intercorrelation matrix was examined so that no two independent variables would be selected which had intercorrelations greater than 0.6. In this manner, a list of independent variables ranging from eight to 18 for the various industries was selected. Multivariate equations were calculated for this set of variables.

In addition, the computer program calculated equations using all possible sets of three of the independent variables to find that set of three independent variables which provided the highest multiple correlation. In most cases, the additional explanation provided by the additional variables is quite small.

After testing several sets of regressions, we selected a final set that included all of the statistically significant area variables 15/ with obviously spurious correlation removed if possible, and with certain variables added which lacked statistical significance but which added substantially to the value of the correlation coefficient for the regression as a whole.

These selections were made three times since the final output comprises three regressions for each four-digit industry: (1) A regression with absolute change as the dependent variable but omitting the variables representing internal forces of concentration and area size affects, (2) a regression with absolute change as the dependent variable and including variables representing concentration and size effects, and (3) a regression with normalized change as the dependent variable and including the normalized variable representing concentration. The final set of regressions is presented in Appendix A. Appendix B gives a brief description and interpretation of the results for each industry. The results are summarized in table 5.

15/ In this study, a statistically significant variable is one that passes the t-test at the 5 percent level of confidence; that is, the t-value is greater than 1.96.

The primary measures of the overall significance of the regression equation used in this study were the coefficient of determination (R^2) and the b values and their significance. The coefficient of determination (R^2) is the square of the coefficient of correlation. It is the proportion of the total variance in the dependent variable that is statistically "explained" by the combination of independent variables. The b values indicate the importance of an independent variable as it affects the dependent variable.

IV. USE OF THE ANALYTICAL RESULTS

Potential for Planning

Multivariate analysis of location can provide information for regional development planning at two levels: (1) for the local development planner--preliminary evaluation of prospects for acquiring new industries, or expansion of existing industries; and (2) for the regional analyst--prediction of the spatial distribution of growth.

The determination of factors influencing location of industry can provide the local planner with a scheme for grouping industries according to their prospects for being developed or expanded in his area. The theory underlying the construction of the multiple regressions in this study indentifies three major influences: (1) internal forces of concentration; (2) effects of area size (in terms of employment); and (3) external forces.

Table 6 shows the significant external variables influencing location patterns for the 53 industries which are indentified and described in appendix A. An area development planner attempting to devise a list of industries most suitable for his area might refer to table 2 to identify attributes that would be favorable or unfavorable for industry location in his area. These include presence or absence of a characteristic distinguished by a dummy (0, 1) variable, or a continuous attribute having a value in his area significantly different from the national averages. (See appendix B for a discussion on the use of "dummy" variables.)

Table 6 can be used to identify industries which may be suitable for an area with given attributes. Local policymakers could list the attributes their area has and scan table 6 for industries. Also, the level of some attributes can be changed by local policy. Examples are attribute 36--the relative extent of unionization (by State), attribute 33--miles of surfaced roads per 1,000 population (by State), and attribute 32--State has inducement for industry location. Local policymakers can use table 6 in a number of ways.

An area might have a low percent of urban population (2), a high percent of rural nonfarm population (3), low median incomes (7, 8) and wages (49), a high percent of Class I and II farms (9, 10), a low percent of Class V and VI farms (13, 14), a high percent of dairying activities (19), a high average value of land and buildings per farm (22), high January and July temperatures (24, 25), low population density (27), low growth (28), State inducements (32), and a high percent of industries in SIC 20, 28, and 29 (56, 64, 65).


ECONOMI
MILES
0
200
400

FIGURE 1

GIONS AND STATE ECONOMIC AREAS: 1960

LEGEND

- ECONOMIC SUBREGION BOUNDARY
- STATE BOUNDARY WHERE NOT PART OF ECONOMIC SUBREGION BOUNDARY
- STATE ECONOMIC AREA BOUNDARY. ALL ECONOMIC SUBREGION BOUNDARIES AND STATE BOUNDARIES ARE ALSO STATE ECONOMIC AREA BOUNDARIES
- 57 ECONOMIC SUBREGIONS - LARGE NUMBERS
- 8 A-N STATE ECONOMIC AREAS - SMALL NUMBERS AND LETTERS

DEPARTMENT OF COMMERCE BUREAU OF THE CENSUS

Table 5.--Summary results of multiple regression analysis

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
		1947	1958				Orientation		Absolute		Normalized		Predict-ability	
SIC Number	No. of SEAs	Employ-ment (000)	Employ-ment (000)	Growth Rate	Size Effects	Concen-tration	Urban-Rural	Market-Resource	R^2	Signif-icance	R^2	Signif-icance	Large Areas	Small Areas
2015	378	21.6	55.6	1.58	O	C	R	R	0.23	Yes	0.26	Yes	Fair*	Fair*
2025	139	7.4	6.9	-0.07	M	D	R	•	.70	Yes	.39	Yes	Good	Fair
2037	173	13.9	40.5	1.90	O	•	R	R	.26	Yes	.22	Yes	Fair*	Fair*
2099	421	59.6	97.9	0.64	M	•	U	M	.30	Yes	.27	Yes	Fair	Poor
2253	109	34.5	60.5	0.76	S	C	•	•	.88	Yes	.28	Yes	Good	Poor
2256	86	13.5	18.4	0.36	M	D	•	•	.34	Yes	.55	Yes	Fair	Fair
2322	91	7.3	12.5	0.70	O	D	R	•	.48	Yes	.42	Yes	Fair	Fair
2329	267	74.8	91.7	0.23	M	C	•	•	.30	Yes	.22	Yes	Fair	Poor
2331	118	32.9	46.3	0.40	O	C	•	R	.37	Yes	.38	Yes	Fair	Fair
2339	137	12.4	34.6	1.80	S	B	•	•	.83	Yes	.26	Yes	Good	Poor
2361	121	22.5	34.8	0.55	S	C	•	•	.65	Yes	.30	Yes	Good	Fair
2391	135	8.0	16.3	1.04	M	C	•	R	.59	Yes	.41	Yes	Fair	Fair
2422	140	10.4	13.0	0.25	O	D	R	R	.35	Yes	.36	Yes	Fair	Fair
2432	124	27.5	47.7	0.73	O	D	•	R	.34	Yes	.46	Yes	Fair	Fair
2512	283	44.8	67.7	0.51	M	C	R	•	.42	Yes	.23	Yes	Fair	Poor
2514	169	19.7	31.2	0.60	S	•	•	•	.52	Yes	.24	Yes	Fair	Poor
2522	71	14.3	17.5	0.22	O	•	•	•	.43	Yes	.44	Yes	Fair	Fair
2531	184	10.2	16.0	0.57	O	D	•	•	.32	Yes	.34	Yes	Poor†	Poor†
2661	138	22.2	38.5	0.73	S	•	U	M	.37	Yes	.13	Yes	Fair	Poor
2671	276	109.7	140.4	0.28	M	D	U	M	.32	Yes	.08	Yes	Fair	Poor
2691	99	9.3	12.7	0.40	M	C	U	•	.41	Yes	.38	Yes	Fair	Fair
2711	503	234.4	294.2	0.26	S	D	U	M	.77	Yes	.49	Yes	Good	Good*
2732	216	12.0	28.5	1.38	S	D	U	•	.25	Yes	.13	Yes	Poor	Poor
2741	226	12.0	19.8	0.65	S	D	U	M	.76	Yes	.21	Yes	Good	Poor
2819	215	42.6	88.4	1.07	O	•	•	R	.13	Yes	.16	Yes	Fair*	Fair*
2823	123	28.6	50.8	0.77	M	D	R	•	.23	Yes	.39	Yes	Fair	Fair
2829	154	85.5	104.5	0.22	O	D	•	R	.24	Yes	.48	Yes	Good*	Good*
2899	273	22.2	26.3	0.18	M	D	U	M	.27	Yes	.22	Yes	Poor	Poor
3253	49	6.8	12.4	0.83	M	B	•	•	.46	Yes	.24	No	Fair	Poor
3271	497	46.8	69.6	0.49	S	D	U	M	.45	Yes	.15	Yes	Good*	Fair*
3281	281	9.9	20.4	1.08	M	B	•	•	.72	Yes	.49	Yes	Good	Fair
3334	21	8.9	17.4	0.95	M	B	•	•	.90	Yes	.64	Yes	Good	Good
3352	84	27.4	42.4	0.55	O	D	U	•	.45	Yes	.31	Yes	Fair	Poor†
3359	63	7.6	16.6	1.19	M	B	•	R	.36	Yes	.52	Yes	Good*	Good*
3393	67	15.7	23.9	0.52	•	•	•	M	.58	Yes	.28	Yes	Fair	Fair*
3441	376	79.1	137.9	0.68	S	D	U	M	.68	Yes	.15	Yes	Good	Poor

Table 5 (continued)

(1) SIC Number	(2) No. of SEAs	(3) 1947 Employment (000)	(4) 1958 Employment (000)	(5) Growth Rate	(6) Size Effects	(7) Concentration	(8) Orientation Urban-Rural	(9) Orientation Market-Resource	(10) Absolute R^2	(11) Absolute Significance	(12) Normalized R^2	(13) Normalized Significance	(14) Predictability Large Areas	(15) Predictability Small Areas
3442	264	21.4	54.3	1.54	S	D	U	M	.58	Yes	.10	Yes	Fair	Poor
3468	245	28.6	35.4	0.27	M	D	•	M	.63	Yes	.55	Yes	Good	Good*
3499	230	7.6	19.7	1.59	S	D	•	•	.58	Yes	.17	Yes	Fair	Poor
3511	22	22.2	44.0	0.98	-M	D	U	•	.89	Yes	.64	Yes	Good	Good
3559	281	68.0	73.9	0.09	S	D	•	•	.26	Yes	.12	Yes	Fair*	Poor
3571	56	45.6	82.4	0.81	O	•	•	•	.16	No	.20	No	Poor	Poor
3599	493	57.1	122.8	1.15	S	B	U	M	.76	Yes	.55	Yes	Good	Good*
3612	28	7.8	9.0	0.14	S‡	C	•	•	.44	Yes	.71	Yes	Fair†	Fair†
3613	110	20.9	45.1	1.15	S	D	•	M	.72	Yes	.18	Yes	Good	Poor
3617	70	7.3	9.9	0.37	O	B	•	•	.45	Yes	.38	Yes	Fair	Fair
3619	125	10.9	17.6	0.61	M	D	•	•	.27	Yes	.28	Yes	Poor	Poor
3661	238	178.6	297.6	0.67	S	B	•	M	.55	Yes	.19	Yes	Good	Poor
3662	88	27.7	92.9	2.35	M	D	U	•	.22	Yes	.24	Yes	Poor	Poor
3716	197	9.3	20.7	1.23	-M	C	•	•	.51	Yes	.50	Yes	Fair	Fair
3721	79	146.0	438.8	2.00	M	B	U	•	.73	Yes	.48	Yes	Good	Fair
3811	135	18.2	63.4	2.46	S	•	U	M	.37	Yes	.18	Yes	Poor	Poor
3941	216	27.1	43.0	0.59	S	•	U	•	.69	Yes	.09	Yes	Good	Poor

* Raised one level because variable has strong economic significance and high F value.
† Reduced one level because variables lack economic significance or excessive constant.
‡ Located in a small number of large areas in base year, with unsubstantial changes in the period.

Key to Table 5

Column	Remarks
1	Source: U.S. Bureau of the Budget, Standard Industrial Classification Manual, Vol. 1, Part 1, Washington, D.C., November 1945.
2	The number of State economic areas that had some employment in the industry in either 1947 or 1958.
3	1947 employment in the industry in thousands.
4	1958 employment in the industry in thousands.
5	$\frac{\text{1958 employment - 1947 employment}}{\text{1947 employment}}$
6	S = Strong 1. R^2 of normalized regression is 50 percent or less of R^2 of absolute regression, and variables contain evidence of size effects; or 2. The bivariate correlation between change in industry employment and total 1947 employment is greater than 0.5, and total employment is a significant variable in the multiple regression with coefficient of 0.3 or higher. M = Moderate 1. R^2 of normalized is between ½ and 2/3 of R^2 of absolute regression, and some evidence of size effect; or 2. Bivariate 0.5 or greater, and size variable significant in multiple with coefficient of 0.1, or high intercorrelation with 1947 industry total, if the industry shows concentration. O = Otherwise or negligible. • = Unknown.
7	C = Concentration. Concentration is evidenced if the weighted value of the constant plus the coefficient of the industry distribution variable is greater than the growth rate of the industry. B = Balanced growth. Balanced growth exists if the weighted value of the constant plus the coefficient of industry distribution is approximately equal to the growth rate of the industry.

Key to Table 5 (continued)

Column	Remarks
	D = Dispersion. Dispersion is evidenced by the weighted value of the constant plus the coefficient of industry distribution being less than the growth rate of the industry.
	• = Inconsistent results. The two forms of the regression fail to give consistent answers.
8	U = Urban. The significant independent variables indicate strong preference for urban area locations.
	R = Rural. The significant variables indicate strong preference for rural area locations.
	• = Unknown. No obvious preference indicated by the independent variables.
9	M = Market 1. Specific variables representing interindustry or final markets, such as military contracts; or 2. Strong size orientation for industries known to have high transportation costs.
	R = Resource. Specific variables representing interindustry, agricultural, or mining activities.
	• = Unknown. No definite indication of market or resource orientation.
10	Absolute regression.
11	The regression is, or is not, statistically significant at the 1 percent confidence level.
12	Normalized regression.
13	The regression is, or is not, statistically significant at the 1 percent confidence level.
14	Good = R^2 for best of the three regressions ≥ 0.6, or between 0.3 and 0.6, with variables indicating particularly strong economic significance.

Column	Remarks
	Fair = R^2 for best of the three regressions $\geq$ 0.3, or less than 0.3 with variables indicating particularly strong economic significance; or $R^2 \geq$ 0.6, with variables economically weak.
	Poor = R^2 for best of the three regressions less than 0.3, or between 0.3 and 0.6, with variables economically weak.
15	Same conditions as for 14, except that conditions must hold for the normalized version of the regression.

Table 6.--Significant external forces of concentration*

Attributes	INDUSTRIES																									
	2015	2025	2037	2099	2253	2256	2324	2329	2331	2339	2361	2391	2422	2432	2512	2514	2522	2531	2661	2671	2691	2711	2732	2741	2819	2823	2829
2																											
3			+																								
4												-										+					
5								-	-																		
6				-																							
7			-					-																			
8	-										-																
9			+																								
10																	+										
11															-												
14		+													+												
15																											
16																				+							
18			+										+														
19																			-								
20	+			+																						-	
21																											
22				-																							
23					-						-																
24			+																+								+
25																											
27				+							+									+						-	
28																						+					
29																											
30		-																							+		
31						+							-														
32								+																			
33							+		+	+								+				+					
34																											
35										+		+			+			+		+	+	+		+			
36							-																				
37								-		-															-		
38	-														-		-										
41																											
42						-													+								
44																				+	-					+	
45					+						+					+											
47																				+		+					
48																						+					
49													+	+							-						
50										+																	
52							+																				
54			+				+																		+		+
56		+		+																							
57											+													+			
58					+				+	+		+												-	+		
59																							+				
60						+							+	+													
61																+	+										
62										+			-														
63																											
64																							+			+	
65																											+
67					+																						
68											-															+	
69																-											
70																											
72																											
73																							+				
74																											
75												+						-									

* Variables statistically significant at the 5 percent level of significance are shown by a mark for the relevant industry. A plus sign indicates an attribute positively associated with growth of the industry in an area and a minus sign indicates negative association.

Table 6.--Significant external forces of concentration (continued)

Attributes	INDUSTRIES																									
	2899	3253	3271	3281	3334	3352	3359	3393	3441	3442	3468	3499	3511	3559	3571	3599	3612	3613	3617	3619	3661	3662	3716	3721	3811	3941
2									+																	
3															+							+				
4																										
5														+		-										
6																										
7																										
8				-	-					-																
9																										
10																						+				
11																										
14																										
15																								+		
16																										
18			+																							
19				+																						
20																										
21						+			+																	
22																										
23	-																						+			
24										+																
25	+																									
27																										
28			+										+			+										
29									+												-					-
30														+												
31															+											
32		+																								
33																				+						
34														-												
35						+		-	+		+	-				+		+			+		+			+
36						+											-							+		
37				+																						
38																							-	-		
41		+																-								
42															+				+							
44					+			+	+				+													
45																			+							
47										+																
48									+	+			+													
49																										-
50							+																			
52																										
54		+										+														
56	+																									
57																					+					
58																					+					
59																										
60																		+								
61				+													+			+					+	
62					-													-								
63																							-			
64																				-			+			
65								+																		
67																										+
68										-																
69							+												-							
70									+		+								-							+
72							+											+								
73																+					-					
74																+										
75											+	+														

The area, however, might be poor in water and port facility availability (30, 31), have few military contracts (35) and lack a major university facility (34).

A list of industries for which the relevant area attributes are statistically significant in determining location can be made by reference to table 6. Cases of conflict (one attribute is positive and another is negative for an industry) would result in the industry being omitted from the list.

Suitable industries are:

2015	2829
2025	2899
2037	3253
2329	3281
2522	3393
2661	3442
2732	3662
2823	

The list may be further reduced by use of table 5, which will form the basis to eliminate industries not consistent with area size characteristics, or with the concentration tendencies of industries. For example, if the area is small, industries with strong size effects are eliminated, i.e., 2661, 2732, and 3442. If it is desired to have a shorter list to investigate, industries with moderate size effects might also be eliminated.

If the area presently has employment in industries 2025, 2329, 2829, and 2899, we would check to see if dispersal is to be expected. Since dispersal is occurring for industries 2025 and 2829, these would also be eliminated from the list of candidates, since further growth is not to be expected in areas presently possessing these industries. The remaining industries on the list would be checked for concentration tendencies. Concentration means that industry growth would not be expected to occur in new areas. Such a tendency is noted for industry 2015, and it is eliminated from the list. The remaining 9 industries on the list given below provide the best prospects for attraction in the total of 53 industries whose growth was statistically analyzed:

2037	2823
2329	2899
2522	3253

3281 3662

3393

This list provides only the first step in any investigation, hopefully providing a means for separating industries worthy of further investigation from those that are not. The premise is that industries not meeting the conditions set forth above with regard to area size orientation, concentration tendencies, etc., will be unlikely to provide new growth opportunities. Of course, this is only a probabilistic statement. With only limited funds available for investigation, the chances of finding suitable location candidates are higher if the effort is concentrated on industries in the list of good prospects.

Another example of the use of the results is to scan table 6 for those attributes that either positively or negatively affect the industries studied. Attribute 35--percent distribution of prime military contracts in the United States, fiscal 1960--is positively associated with growth in 16 industries and negatively associated with growth in two. The two industries which had negative associates were 3393--Welded and Heavy-Riveted Pipe, and 3499--Fabricated Metal Products.

The 16 industries with positive associations were:

2339--Women's and Misses' Outerwear, n.e.c. (not elsewhere classified)

2391--Curtains and Draperies

2512--Wood Household Furniture, Upholstered

2531--Public-Building and Related Furniture

2671--Paperboard Boxes; Folded, Set-up and Corrugated

2691--Die-Cut Paper and Paperboard, and Cardboard

2711--Newspapers

2741--Miscellaneous Publishing

3352--Rolling, Drawing, and Alloying of Aluminum

3441--Fabricated Structural Steel and Ornamental Metal Work

3468--Electroplating, Plating, and Polishing

3599--Machine Shops

3613--Instruments for Indicating, Measuring, and Recording Electrical Quantities and Characteristics

3661--Radios, Radio and Television Equipment (except Tubes), Radar and Related Detection Apparatus, and Phonographs

3716--Automobile Trailers (for Passenger Cars)

3941--Games and Toys (except Dolls and Children's Vehicles)

One attribute--median income of females with income in 1960, attribute 8--was negatively associated with the five industries with which it was significantly related. The industries are:

2015--Poultry and Small Game Dressing and Packing--Wholesale

2361--Children's and Infants' Dresses

3281--Cut-Stone and Stone Products

3334--Primary Refining of Aluminum

3442--Metal Doors, Sash, Frames, Molding, and Trim

The negative association means that as median female income increased the total employment in the listed industries decreased.

Care should be used in interpreting the results presented in table 6 since high correlation among some independent variables was found to exist. Some variables were dropped on this account. Thus a variable may appear significant in the table but it is in fact a proxy for a group of variables.

SIC 2015 - Poultry and Small Game Dressing and Packing--Wholesale

Establishments in this industry are primarily engaged in killing, dressing, packing, and canning poultry, rabbits, and other small game for the trade. Employment in this industry increased 158 percent from 21,600 in 1947 to 55,600 in 1958. The industry was widely dispersed, being located in 378 SEA's.

The high value of the constant in the normalized version of the regression shows that the industry tended to concentrate spatially in the period. Size of area had negligible effect on location. Resource orientation of this industry is shown by the high correlation of growth with the percent of farms in poultry raising (20). 16/ Building costs (38) were highly significant variables in all forms of the equation. Building costs are probably a proxy for the tendency for the industry to expand in the South. (Building costs are lowest in the South and highest in the Great Lakes and North Atlantic regions.) The significance of the income and wage variables further emphasizes the concentration of this industry in the low income and low wage areas of the South. Although the regressions explain less than 30 percent of the variance, the strong resource orientation increases the regressions' predictive value.

SIC 2025 - Special Dairy Products

Establishments in this industry are primarily engaged in manufacturing special dairy products. Important products of this industry include processed and ladle butter, malted milk, and fermented milk; and processed cheese, cheese pastes, and spreads made from purchased natural cheese. Employment in this industry declined about 7 percent from 7,400 in 1947 to 6,900 in 1958. The industry was located in 139 SEA's.

The regressions show relative dispersion of this industry. Since employment in this industry actually declined in the period, these relative shifts mean that absolute shifts in industry location were occurring away from the areas strong in the industry in 1947 and toward areas with little of this industry employment in 1947. The areas receiving growth had distinguishing characteristics: They were moderately large (43), they had a relatively high proportion of employment in other food processing industries (56), and Class VI farms were a relatively high proportion of the area's farmlands (14). In general, the industry shifted toward rural areas hospitable to food processing industries but with poor farmland. The industry specific and general size variables give statistical strength to the equations. The regressions explain almost 70 percent of the variation in employment change among areas.

16/ Numbers in parentheses throughout this appendix designate the variable (see table 2).

SIC 2037 - Frozen Fruits, Vegetables, and Sea Foods

Establishments in this industry are primarily engaged in quick-freezing and cold-packing (freezing) vegetables, fruits, fish, and shellfish. Employment in this industry increased 190 percent from 13,900 in 1947 to 40,500 in 1958. The industry was located in 173 SEA's in 1947 or 1958.

Resource orientation is indicated by the positive correlation with areas having a high proportion of farmland in fruits and nuts. The significance of warm winter temperatures is consistent with the growing conditions for fruits and vegetables. The growth areas in this industry also tend to be rural areas (3), with low per capita incomes (7), and tend to have a relatively high proportion of employment in quarrying (54). The evidence on concentration is contradictory and size effects are negligible. Thus, employment in the base year for the industry was of no help in predicting the distribution of growth. Because of the relatively strong interpretive value of the independent variables, these regressions are regarded as "fair," even though they explain less than 30 percent of the variance.

SIC 2099 - Food Preparations, n.e.c.

Establishments in this industry are primarily engaged in manufacturing prepared foods and miscellaneous food specialties, n.e.c. (not elsewhere classified), such as bakers', confectioners', and household supplies (including chili powder, shredded coconut, pie and cake fillings, and powdered sugar); bouillon cubes; instant chocolate; ready-mix desserts; eggs (dehydrated, processed, etc.); "health foods"; "native" preparations (Italian, Mexican, etc.); peanut products; potato chips; sandwich spreads (except cheese); spices; sweetening syrups; and roast coffee (excluding roasting done by wholesale grocers). Employment in this industry increased 64 percent from 59,600 in 1947 to 97,900 in 1958. Located in 421 SEA's, the industry was widely dispersed.

There are moderate size effects but there are no strong indications as to whether the industry was concentrated or dispersed. Area characteristics relevant to the location of growth are high population density (27) and specialization in food processing industries (56). The significant variables indicate a preference for urban locations and a tendency to be market oriented. The regressions are useful only for estimation in relatively larger areas.

SIC 2253 - Knit Outerwear Mills

Establishments in this industry are primarily engaged in knitting outerwear or in manufacturing outerwear from knit fabrics produced in the same establishment. Important products include sweaters, bathing suits, dresses, and headwear. Employment in this industry increased 76 percent from 34,500 in 1947 to 60,500 in 1958. The industry was located in 109 SEA's in 1947 or 1958.

Internal forces tended to cause considerable concentration in the areas with high employment in the industry in 1947. High intercorrelation between size of area and size of industry means that there was strong size effects. External

forces or characteristics of the area's attracting growth in this industry were a high proportion of employment in the two-digit textile industry (58), relatively high male incomes in States that offered inducement for industry locatio (45), and the tendency for the farms in the area to be low in tractor ownership (23). The strong size effects result in a regression that explains almost 90 percent of the variation in growth among areas. The normalized regression explains 25 percent of the variance, indicating poor predictability for small areas.

SIC 2256 - Knit-Fabric Mills

Establishments in this industry are primarily engaged in knitting tubular or flat fabric and in dyeing or finishing knit fabric. Employment in this industry increased about 35 percent from 13,500 in 1947 to 18,400 in 1958. The industry was located in a relatively few areas: 86 in 1947 or 1958.

Employment in this industry tended to disperse from areas that had high employment in the industry in 1947. There were moderate size effects. When size was discounted (in the normalized version), high relative growth was associated with a high proportion of employment in lumber and wood products industries (60). Since no apparent interindustry relations were involved, this relationship may reflect association with a common set of area characteristics. No particular set of characteristics appears to be closely associated with the 1947 distribution of the lumber and wood products industries. This industry was particularly strong in New England, the South, and the Pacific Northwest. The normalized version was most significant, explaining more than half the variance in location of growth and thereby indicating a fair ability to predict location of this industry in small as well as large areas.

SIC 2322 - Men's, Youths', and Boys' Underwear

Establishments in this industry are primarily engaged in manufacturing men's, youths', and boys' underwear cut and sewed from purchased woven or knit fabric. Employment in this industry increased 70 percent from 7,300 in 1947 to 12,500 in 1958. The industry was located in 91 SEA's.

This industry tended to disperse from its areas of concentration in 1947. The areas attracting growth were relatively nonunionized and had high proportions of employment in bituminous coal mining (52) and quarrying (54). Probably this attraction was due to the labor force characteristics associated with prevalence of these industries rather than to interindustry relations. The attraction to mining areas probably indicates a tendency to locate in rural areas where mining is most prevalent. The regressions explain better than 40 percent of the variance in the location of growth.

SIC 2329 - Men's, Youths', and Boys' Work, Sport, and Other Clothing, n.e.c.

Establishments in this industry are primarily engaged in manufacturing men's, youths', and boys' work, sport, and other clothing not elsewhere

classified. Important products of this industry include overalls, work pants, industrial garments, mackinaws, lumberjackets, ski and snow suits, and hunting and riding garments (excluding separate trousers). Employment increased about 23 percent from 74,800 in 1947 to 91,700 in 1958. The industry was located in 267 SEA's in 1947 or 1958.

The high constant in the normalized version and the high intercorrelation with size (43) show that this industry tended to concentrate and that there were moderate to strong size effects. External characteristics of areas receiving growth are low per capita incomes (7), low percent of employment in professional or managerial capacities (5), and low power costs (37). The normalized version shows also a tendency for employment to grow in States providing location inducements (32). The equations explain a relatively small proportion of the variance--20 to 30 percent for both normalized and absolute versions.

SIC 2331 - Women's and Misses' Blouses and Waists

Establishments in this industry are primarily engaged in manufacture of women's and misses' blouses and waists. Employment in this industry increased 40 percent from 32,900 in 1947 to 46,300 in 1958. The industry was located in 118 SEA's in 1947 or 1958.

The absolute version of the regression does not appear to be a good equation, and is ignored in this analysis. (Note lack of significant variables.) The normalized version shows that this industry had a strong tendency to concentrate. Concentration, however, does not appear to be related to total size of the area. Areas with a high percent of employment in the two-digit textile industry (58) and areas in States with a high density of road mileage per capita (33) were attractive to the industry. These areas also have a low percent of the work force in professional and managerial occupations (5). The normalized regression is rated as "fair" in explaining location of growth.

SIC 2339 - Women's and Misses' Outerwear, n.e.c.

Establishments in this industry are primarily engaged in manufacturing women's and misses' outerwear not elsewhere classified such as bathing suits, beachwear, slacks, riding habits, and ski suits cut and sewed from purchased woven or knit fabrics. Employment in this industry increased 180 percent from 12,400 in 1947 to 34,600 in 1958. The industry was located in 137 SEA's.

Despite tremendous growth, the industry displays spatial balance in its location pattern, and is strongly affected by the size of the area. Significant variables representing external forces are distribution of defense contracts (35) and density of roads per capita in the State (33). Since this industry is in no way related to defense spending, the distribution of defense contracts stands as proxy for a set of related area characteristics: mainly size of area, degree of urbanization, and January temperatures. The road density variable indicates that the areas attracting growth tend somewhat to be in States that have low densities of population. The regression expressing absolute distribution of growth explains a very high proportion of variance--above 80 percent. When normalized,

however, the regression explains only a small part of the variance. This regression explains well the location pattern of the industry but is unreliable for prediction of employment growth in small areas.

SIC 2361 - Children's and Infants' Dresses

Establishments in this industry are primarily engaged in manufacturing children's and infants' dresses, cut and sewed from purchased woven or knit fabric. Employment in this industry increased about 55 percent from 22,500 in 1947 to 34,800 in 1958. The industry was located in 121 SEA's.

The industry displayed a tendency to concentrate in large areas. Other characteristics of the areas attracting growth were high population density (27 low per capita incomes (8), and a relatively high proportion of employment in tobacco manufactures (57). This latter variable obviously stands as a proxy for a set of area characteristics common to both the tobacco (SIC 21) and children's dress industries. The regression expressing absolute change explains better than 60 percent of the variance in location of growth. Prediction in small districts will be fair.

SIC 2391 - Curtains and Draperies

Establishments in this industry are primarily engaged in manufacturing curtains and draperies from purchased fabric. Employment in this industry increased 104 percent from 8,000 in 1947 to 16,300 in 1958. The industry was located in 135 SEA's.

The industry displayed a tendency to concentrate somewhat in large areas. Other characteristics of areas receiving growth in this industry were a high percent of prime defense contracts (35), a high percent of employment in miscellaneous manufacturing (75), and a high percent of employment in textile manufacturing (58). The correlation with military contracts probably reflects the size effects and the area characteristics associated with the distribution of these contracts. The third variable may indicate some resource orientation since this industry purchases from the textile industry. In the normalized version, percent growth was negatively associated with median years of adult schooling (4), which is contrary to a positive association with military contracts in the absolute version of the regression and indicates that the size effects are nonlinear. Despite nonlinear associations, these regressions are regarded as fair for prediction purposes.

SIC 2422 - Veneer Mills

Establishments in this industry are primarily engaged in producing commercial veneer, either face or technical for sale to the trade. Employment in this industry increased 25 percent from 10,400 in 1947 to 13,000 in 1958. Located in 140 SEA's, the industry was widespread considering the relatively small amount of employment.

The industry tended to shift locations with the areas dominant in 1947 tending to lose both absolutely and relatively to other areas. The growth, however, did remain in areas with high proportions of employment in lumber and wood product industries (60). In absolute terms, low summer temperatures (25) and absence of port facilities (probably representing inland areas not on navigable waterways) (31) were also favorable characteristics. The normalized regression showed a different set of relevant characteristics. The areas were low in paper production (62), had low capital value per farm (22), had a high proportion of farms in fruit and nuts (18), and had high wage rates (49). The regressions are regarded as fair for prediction.

SIC 2432 - Plywood Plants

Establishments in this industry are primarily engaged in manufacturing commercial plywood, including technical and industrial grades, from veneer produced in the same establishment or from purchased veneer. Employment in this industry increased more than 70 percent from 27,500 in 1947 to 47,700 in 1958. The industry was located in 124 SEA's.

As with the veneer mills (SIC 2422), this industry tended to shift away from areas of previous concentration but to remain, generally, in areas with high proportions of employment in lumber and wood products (60). The only other significant characteristic of attractive areas was that they were high wage areas (49). Predictability for this industry should be fair for both large and small areas.

SIC 2512 - Wood Household Furniture, Upholstered

Establishments in this industry are primarily engaged in manufacturing upholstered furniture on wood frames or manufacturing wood frames for upholstered furniture. Employment in this industry increased 50 percent from 44,800 in 1947 to 67,700 in 1958. The industry was located in 283 SEA's.

The industry exhibited a strong tendency to concentrate in those areas with high absolute employment in the industry and with a high proportion of total employment in the industry in 1947. There were moderate size effects, indicated by the positive association with military contracts and by the high intercorrelation between (39) and (43). Areas of growth also tended to be rural areas with poor farms (negative 11 in equation 2, and positive 14 in equation 3). Only the absolute version of the regression provides satisfactory prediction.

SIC 2514 - Metal Household Furniture

Establishments in this industry are primarily engaged in manufacturing metal household furniture of a type commonly used in dwellings. Employment in this industry increased about 60 percent from 19,700 in 1947 to 31,200 in 1958. The industry was located in 169 SEA's.

The distribution of growth in this industry was strongly correlated with the size distribution of areas. Other characteristics of the growth areas were relatively high per capita income in States offering location inducements (45) and relatively low proportion of employment in primary metal production (69). The regression displayed some evidence of dispersal to areas that had a low percent of employment in metal household furniture but a high percent in the two-digit furniture manufacturing industry (61), thus indicating some tendency for the industry to seek out new areas for growth that are hospitable to furniture manufacturing. The regression representing absolute distributions explains about 50 percent of the variance. The low value of R^2 for the normalized version, however, indicates that prediction in small areas is likely to be poor.

SIC 2522 - Metal Office Furniture

Establishments in this industry are primarily engaged in manufacturing metal office furniture. Employment in this industry increased more than 20 percent from 14,300 in 1947 to 17,500 in 1958. The industry was located in 71 SEA'

The regressions are not consistent in showing concentration; i.e., concentration is indicated by the high constant in the normalized regression but is countered by the negative coefficient in the absolute version. Although this industry was located in a small number of areas with high average employment, the growth in the period was not associated with the size of the area. In the absolute form of the regression, there was a tendency to seek out locations with characteristics similar to those sought by the metal household furniture industry (SIC 2514), as noted by the retreat of variables 45 and 69 (the latter with negative sign). Other characteristics of growth areas are that they have wealthy farm lands (10), low building costs (38), and a high proportion of employment in tobacco manufactures (57). The latter two variables indicate shift to Southeastern, Midwestern, or Appalachian locations. The regressions explain over 40 percent of the variance.

SIC 2531 - Public-Building and Related Furniture

Establishments in this industry are primarily engaged in manufacturing furniture for schools, theaters, assembly halls, churches, and libraries, and in manufacturing seats for public conveyances. Employment in this industry increased about 55 percent from 10,200 in 1947 to 16,000 in 1958. This industry was located in 184 SEA's.

The industry showed a strong tendency to shift location. External characteristics of areas attracting growth are a high percent of prime military contracts (35), States with high density of surfaced roads per capita (33), and a low percent of employment in miscellaneous manufacturing (75). These variables would appear to be proxies for the true locational determinants. The military contracts orientation probably reflects the shift to areas of high urbanization and warm climate with an added pull to the south exerted through the road density by State. The regressions explain about one-third of the variance. Since, however, the variables appear to be proxies for the variables that would have true economic significance, it can be expected that the regressions for this industry would not have predictive power.

SIC 2661 - Paper Bags

Establishments in this industry are primarily engaged in manufacturing paper bags from purchased paper and similar materials. Important products of this industry include kraft paper bags for groceries, heavy duty bags for cement, etc., and glassine, cellophane, and miscellaneous merchandise bags. Employment in this industry increased 73 percent from 22,200 in 1947 to 38,500 in 1958. This industry was located in 138 SEA's.

The regressions were unable to show the internal location trends of this industry. There was an orientation toward growth in large areas (43). Other characteristics of growth areas are: they have a low percent of farms in dairying (probably a spurious correlation) (19) and they tend to be urbanized areas with port facilities (42), probably indicating the market orientation of this industry. The absolute version provides fair predictability for large areas. However, the very low R^2 value for the normalized version indicates that prediction for small areas is unreliable.

SIC 2671 - Paperboard Boxes: Folded, Set-up and Corrugated

Establishments in this industry are primarily engaged in manufacturing paperboard containers and boxes from purchased paperboard or fiber stock. Employment in this industry increased 28 percent from 109,700 in 1947 to 140,400 in 1958. The industry was located in 276 SEA's.

The industry displayed some tendency to disperse from areas in which the industry was previously concentrated. There were moderate size effects despite the negative correlation with variable 43. In absolute terms, the growth in this industry was oriented to densely populated areas (27), to areas that had a high proportion of prime military contracts (35), and to four areas that had large cities (500,000 population) and were located in States providing inducements (44). Six metropolitan areas met this requirement: two in Pennsylvania, two in Maryland, one in Kentucky, and one in Louisiana. In general, the significant variables indicate a market orientation for this industry. The absolute version of the regression explains about one-third of the variance. Nevertheless, the very low R^2 for the normalized version means that prediction for small areas will be very unreliable.

SIC 2691 - Die-Cut Paper and Paperboard, and Cardboard

Establishments in this industry are primarily engaged in die-cutting purchased paper and paperboard, and in manufacturing cardboard by laminating, lining, and surface-coating paperboard. Important products of this industry are paper bottle caps and tops; index, library, and other cut cards; cutouts; photograph folders; stencil cards; and cardboard panels and pasted chipboard and newsboard. Employment in this industry increased 40 percent from 9,300 in 1947 to 12,700 in 1958. This industry was located in 99 SEA's.

Growth in the industry tended to concentrate in areas with high industry employment and moderately high total employment in 1947. Contrary size effects emerged as two areas that lost employment in this industry had large cities and were located in States providing location inducements (44). Low wages also appear to be an attractive force for this industry (49). Both the absolute and normalized versions of the regressions have a fair degree of predictability.

SIC 2711 - Newspapers

Establishments in this industry are primarily engaged in publishing newspapers or in publishing and printing newspapers. Employment in this industry increased about 25 percent from 234,400 in 1947 to 294,200 in 1958. The industry was located in all but three of the 506 SEA's.

The strong orientation of this industry to population as a final market is indicated in the normalized version of the equation which shows the very strong correlation between normalized change in employment in this industry and percentage change in population. Additional characteristics of areas receiving high growth in newspaper employment irrespective of area size are a combination of high percent urban and warm winter temperatures (47), high level of median adult education (4), and high density of roads (33).

The absolute version of the regression explains about two-thirds and the normalized version about one-half of the variance. Predictability should be good for large and fair for small areas.

SIC 2732 - Book Printing

Establishments in this industry are primarily engaged in printing only or in printing and binding books and pamphlets but are not engaged in publishing. Employment in this industry increased 138 percent from 12,000 in 1947 to 28,500 in 1958. This industry was located in 216 SEA's.

This industry displayed a tendency to shift location away from areas of previous concentration toward large areas (43) having a relatively high percent of employment in the apparel industry (59), in the chemical industry (64), and in transportation equipment (73). The variables representing industry structure of the area are probably proxies for other area characteristics since there is no obvious connection between book printing and these industries.

SIC 2741 - Miscellaneous Publishing

Establishments in the industry are primarily engaged in publishing such products as maps, atlases, sheet music, and directories, or are engaged in miscellaneous publishing activities not elsewhere classified, whether or not they are engaged in printing. Employment in this industry increased 65 percent from 12,000 in 1947 to 19,800 in 1958. The industry was widely spread in 226 SEA's.

The industry showed a tendency to disperse from areas of previous concentration toward large areas. External characteristics of the areas receiving growth are high proportion of defense contracts (35), high proportion of employment in tobacco manufactures (57), and low proportion of employment in textile manufactures (58). This represents some nonlinear association of growth with large urbanized areas in the South Atlantic States plus shifts to new growth areas in the West. The absolute version of the regression explains about 75 percent of the variance. The normalized version, however, explains only 20 percent of the variance and indicates that predictability for smaller areas is poor.

SIC 2819 - Industrial Inorganic Chemicals, n.e.c.; SIC 2811 - Sulfuric Acid

Establishments in this industry are primarily engaged in manufacturing sodium hydroxide, potassium hydroxide, potassium and sodium carbonates, sodium bicarbonate, sal soda, soda ash, and chlorine; also included are establishments primarily engaged in manufacturing sulfuric acid by the contact or chamber process. Employment in this industry increased 107 percent from 42,600 in 1947 to 88,400 in 1958. The industry was located in 215 SEA's.

Growth in this industry was not particularly related to the 1947 distribution of the industry or to the size of the area. Growth was related to the presence of adequate processing water (30), to the textile industry (58), and to low power costs (37). In the normalized version, the presence of quarrying and mineral mining (54) replaced water availability as a significant variable in explaining growth. In general, the variables indicate the importance of resources (water and minerals) in location, and perhaps low wage labor, as indicated by the presence of the textile industry. Although the variables explain less than 30 percent of the variance, indicating a low level of reliance on the regressions for prediction, the economic strength of the variables indicates that these regressions may be fair in predicting location.

SIC 2823 - Plastic Materials and Elastomers, except Synthetic Rubber

Establishments in this industry are primarily engaged in manufacturing plastic materials in the form of sheets, rods, tubes, granules, powders, or liquids for use in further manufacturing. Important products of this industry include (1) condensation plastics such as alkyd resins, (2) protein plastics such as casein, (3) cellulose plastics and vulcanized fiber, (4) carbohydrate plastics, and (5) synthetic elastomers such as plasticized polyvinyl alchohol. Employment in this industry increased 77 percent from 28,600 in 1947 to 50,800 in 1958. The industry was located in 123 SEA's.

The industry showed some tendency to disperse. There is moderate positive correlation with area size (43) yet this relationship appears to be complex and nonlinear, making the absolute version of the regression useless. In the normalized version, the area characteristics are low percent of poultry farms (20); low population density (27); and high percent of employment in chemicals (64) and stone, clay, and glass manufacturing (68). The normalized version of the regression has fair predictive value.

SIC 2829 - Industrial Organic Chemicals, n.e.c.;
SIC 2822 - Intermediates, Dyes, Color Lakes, and Toners

Establishments in the industry are primarily engaged in manufacturing industry organic chemicals not elsewhere classified. Important products of this industry include (1) noncyclic organic chemicals such as acetic and tartaric acids and their metalic salts; (2) solvents such as amyl, butyl, and ethy alcohols; acetone and carbon tetrachloride; (3) synthetic perfume and flavoring materials such as saccharin and terpineol; (4) rubber chemicals such as antioxidants; (5) plasticizers such as synthetic camphor; (6) synthetic tanning materials; and (7) chemical-warfare gases. In addition, some establishments ar primarily engaged in manufacturing cyclic intermediates, dyes, color lakes, and toners. Employment in these industries increased 22 percent from 85,500 in 195 to 104,500 in 1958. The two SIC classes of industry were located in 154 SEA's.

The industry showed a tendency to disperse toward areas that were proportionally strong in the petroleum products industry (65) and that had a warm winter climate (24). This is indicative of (a) the dominance of petrochemicals in the growth sector of the organic chemical industry, and (b) the resource orientation of this industry. In the normalized version, orientation to minera mining (54) is also evidenced. The normalized version explains almost half of the variance. On the basis of the economic strength of the variables, this regression is rated as a good predictor of location.

SIC 2899 - Chemicals and Chemical Products, n.e.c.

Establishments in this industry are primarily engaged in manufacturing miscellaneous chemical products not elsewhere classified such as bluing, laundry sours, mucilage, writing and stamp pad inks, and dextrine sizing; agricultural and industrial and household disinfectants and deodorants; household insecticides and repellants; household adhesives and cements; and industrial compounds--boiler and insulating compounds, metal, oil and water-treating compounds, waterproofing compounds, and chemical supplies for foundries. Employment in this industry increased about 20 percent from 22,200 in 1947 to 26,300 in 1958. The industry was located in 273 SEA's.

The industry showed a strong tendency to disperse with size of area having a moderate effect on the location of new employment (43). In the normalized version, areas attracting growth had high percentages of employment in food processing (56), high summer temperatures (25), and a low percent of farms owning tractors (23). This would indicate a shift to the South with some indications of consumer market orientation. The regressions are considered poor for prediction.

SIC 3253 - Floor and Wall Tile, except Quarry Tile

Establishments in this industry are primarily engaged in manufacturing ceramic tile, mosaic tile, and glazed and enameled tile including faience. Employment in this industry increased 83 percent from 6,800 in 1947 to 12,400 in 1958. The industry was confined to 49 SEA's.

The industry was reasonably stable in its location pattern, remaining in large areas. Other characteristics of attracting areas were the presence of a large city interacting with a high rate of population growth (41) and warm winter temperatures (24). In the normalized version, quarrying or mineral mining (54) and location in States with location inducements (32) were also significant. Predictability is fair for large areas and poor for small areas.

SIC 3271 - Concrete Products

Establishments in this industry are primarily engaged in manufacturing concrete building blocks and shapes, pipe and conduit, and similar products such as poles, piling, septic tanks, incinerators, and vaults made from a combination of stone or gravel, sand, and cement. Employment in this industry increased 49 percent from 46,800 in 1947 to 69,600 in 1958. The industry was spread widely throughout the Nation, being located in 497 SEA's.

Growth in this industry tended to disperse, being strongly affected by size of area (43). In addition, the area characteristics for growth were rapidly growing population (28) and high proportions of area farms in fruit and nut production (18). This industry is market oriented; the major market is the construction industry which in turn is oriented toward large areas growing in urbanization. The strong economic significance of the independent variables led to upgrading the prediction rating of these regressions. Prediction is considered to be good for large areas and fair for small areas.

SIC 3281 - Cut-Stone and Stone Products

Establishments in this industry are primarily engaged in cutting, shaping, and finishing marble, granite, slate, and other stone for building and miscellaneous uses. Employment in this industry increased 108 percent from 9,900 in 1947 to 20,400 in 1958. The industry was located in 281 SEA's.

Despite large growth, the location pattern of this industry remained very stable with growth predominately related to the 1947 distribution of the industry. This growth also demonstrated an inverse relationship with size; that is, employment tended to grow more in areas with low total employment (43). In the normalized version, other characteristics emerge: high power costs (37), high percent of employment in the furniture industry (61), and low per capita incomes (8). The attraction to high power costs probably represents a geographic proxy; high power costs are typical in the Great Lakes and New England areas. The Great Lakes area is also the older center of the furniture manufacturing industry.

Because of the location stability of this industry, the regressions explain a large proportion of the variance--over 70 percent for the absolute version and almost half for the normalized version.

SIC 3334 - Primary Refining of Aluminum

Establishments in this industry are primarily engaged in producing aluminu from the ore or from alumina and in refining aluminum by any process. Employment in this industry increased 95 percent from 8,900 in 1947 to 17,400 in 1958 The industry is concentrated in 21 SEA's.

The industry tended to remain concentrated in a very small number of areas each with high employment in the industry. These were not large in total employment. Within this group, however, there was some shifting. Large growth occurred in the area that had a large metropolitan city and was in a State that offered location inducements (44). Growth was also attracted to areas that tended to have low per capita incomes (8) and a low percent of employment in paper products industries (62). The regressions explain a large proportion of the variance. The variables in the regression are highly specific to a very few areas with employment in this industry and cannot be extrapolated to predict growth in other areas. They should, however, provide good prediction in the 21 SEA's with employment in this industry.

SIC 3352 - Rolling, Drawing, and Alloying of Aluminum

Establishments in this industry are primarily engaged in rolling, drawing, and extruding aluminum and aluminum-base-alloy basic shapes such as plates, sheets, rods, wire, and tubing. Employment in this industry increased 55 percer from 27,400 in 1947 to 42,400 in 1958. The industry was located in 84 SEA's.

The industry strongly tended to shift location, being heavily drawn to areas with high percentages of military prime contracts. Since this industry is not necessarily size-oriented, the heavy pull to defense contract areas is indicative of the market orientation of this industry. Inclusion of the livestock variable (21) in the normalized version unfortunately does not make too much sense. The apparent attraction of areas high in unionization probably reflects the need for skilled labor in this industry.

The absolute and normalized versions of the regression explain 45 percent and 30 percent of the variance, respectively. The absolute version is the more defensible in terms of economic logic although the large constant with a high variance can be expected to create difficulty in making reliable predictions.

SIC 3359 - Rolling, Drawing, and Alloying of Nonferrous Metals, n.e.c.

Establishments in this industry are primarily engaged in rolling, drawing, and alloying nonferrous metals other than copper and aluminum. Employment in this industry increased 119 percent from 7,600 in 1947 to 16,600 in 1958. The industry was located in 63 SEA's.

The absolute version of the equation is statistically weak and the normalized version suggests balanced growth. The only significant external effect was a clear tendency for growth to locate in metal mining areas (50). This implies a heavy resource orientation of this industry, rather than the market orientation evidenced in the rolling of aluminum (SIC 3352). The normalized version can be expected to provide good predictions for those areas with some employment in the industry in the base year.

SIC 3393 - Welded and Heavy-Riveted Pipe

Establishments in this industry are primarily engaged in manufacturing welded pipe and tubes and lock-joint and heavy-riveted pipe from purchased skelp. Employment in this industry increased 52 percent from 15,700 in 1947 to 23,900 in 1958. The industry was concentrated in 67 SEA's.

The most significant characteristic of areas attracting growth in this industry is the proportion of employment in the petroleum industry (65). This probably reflects the fact that the growth sector of SIC 3392 is heavily engaged in providing pipe for the petroleum industry and that this sector is market oriented in its location. The interaction of a large (500,000 population) city and State inducements is a strong variable; the two SEA's which met these qualifications each received about 1,500 additional employees in this industry. The industry growth also appears to be oriented toward areas that are low in defense contracts (35), probably indicating a Midwestern and Southern orientation. The strong orientation to petroleum refining led to considering prediction in small areas to be fair despite the low R^2 for the normalized version of the regression.

SIC 3441 - Fabricated Structural Steel and Ornamental Metal Work

Establishments in this industry are primarily engaged in manufacturing (a) fabricated iron and steel or other metal for structural purposes and (b) architectural ornamental work of ferrous and nonferrous metals. Employment in this industry increased about 75 percent from 79,100 in 1947 to 137,900 in 1958. The industry was widespread, being located in 376 SEA's.

This industry underwent structural shifts. Growth was greatest in large areas which were low in employment in this industry in the base year. The growth areas were metropolitan areas that were heavy in military contracts and that had warm winters and institutions of higher education. Six of the growth areas had large cities and were in States with industry inducements. On the average, these six areas received an extra 245 employees in this industry. Of these areas, two were located in Pennsylvania, two in Maryland, one in Kentucky, and one in Louisiana. The heavy size effects indicate high predictability in the large areas but poor predictability in small areas.

SIC 3442 - Metal Doors, Sash, Frames, Molding, and Trim

Establishments in this industry are primarily engaged in manufacturing ferrous and nonferrous metal and metal-covered doors, sash, window and door frames, store fronts, molding, and trim. Employment in this industry increased 154 percent from 21,400 in 1947 to 54,300 in 1958. The industry was located in 264 SEA's.

This industry dispersed from previous areas of concentration, showing strong size effects with a close correlation between growth in the industry and distribution of total employment. Other than size, the growth areas tended to be highly urbanized areas in the warm winter climates with some tendency to be strong in higher education (47 and 48). Growth was also attracted to areas in which per capita incomes tended to be low (8). In general, growth characteristics of this industry were similar to those for SIC 3441. The strong size orientation of growth means that the absolute version of the equation explains a high proportion, almost 60 percent, of the variance. However, the lack of strong structural orientation results in a very low R^2 for the normalized version indicating very poor predictability in small areas.

SIC 3468 - Electroplating, Plating, and Polishing

Establishments in this industry are primarily engaged in all types of electroplating, plating, and metal-polishing work. Employment in this industry increased about 25 percent from 28,600 in 1947 to 36,400 in 1958. This industry was located in 245 SEA's.

This industry demonstrated a strong tendency to disperse with moderate size effects. The new growth areas were large in total employment (43), had concentrations of prime military contracts (35), and were heavy in miscellaneous manufacturing (75). The high correlation of the normalized version shows that the structural shifts were consistent with the size effects--i.e., small areas lost and big areas gained in both absolute and percentage terms. The gaining areas also tended to have high percentages of employment in metal fabricating (the two-digit industry to which electroplating belongs). In general, the industry demonstrated market orientation; its dispersion was probably indicative of a shifting market structure. The consistency between the size effects and structural changes in the industry gave both the absolute and normalized versions of the regression similar and quite high R^2 with both over 50 percent.

SIC 3499 - Fabricated Metal Products, n.e.c.

Establishments in this industry are primarily engaged in manufacturing metal products not elsewhere classified, including metal novelties and specialties. Employment in this industry increased 159 percent from 7,600 in 1947 to 19,700 in 1958. Considering the small employment in the industry, it was widespread, being located in 230 SEA's.

The industry shifted location, with growth dominated by size effects. Attracting areas tended to be large, relatively low in prime military contracts (35), and high in percent of employment in mineral mining and quarrying (54). These tendencies indicate that the areas may be large in total employment and population but relatively nonurbanized for their size. Since the shifts in the industry were dominated by size effects, the normalized version had a very low R^2 and the absolute version of the regression a high R^2, providing the basis for fairly reliable prediction only in larger areas.

SIC 3511 - Steam Engines, Turbines, and Water Wheels

Establishments in this industry are primarily engaged in manufacturing steam engines, steam turbines, water wheels, and water turbines. Employment in this industry increased 98 percent from 22,200 in 1947 to 44,000 in 1958. The industry was concentrated in 22 SEA's.

On the average, the 22 areas in which the industry was located were large, with mean employment of 408,000 in 1947 compared with a mean of about 112,000 for all SEA's. Within these 22 areas, growth was negatively associated with size and tended to shift from the areas with higher percents of employment in this industry in the base year. A sizeable portion of the growth occurred in two areas that met the conditions of variable 44 (discussed elsewhere). Each of these two areas tended to receive about 3,000 additional workers in this industry. The normalized version of the regression demonstrated the strong tendency for this industry to grow in areas experiencing large proportional growth (28) and in areas having attributes associated with growth during the period--i.e., urbanization, warm winters, and higher education (48). Prediction should be good when confined to the 22 areas presently containing the industry.

SIC 3559 - Special-Industry Machinery, n.e.c.

Establishments in this industry are primarily engaged in manufacturing special-industry machinery not elsewhere classified, such as smelting and refining equipment; machinery for cement, paint, cigar-cigarette, and shoe manufacturing; and machinery for the working of leather, rubber, tobacco, and stone. Employment in this industry increased less than 10 percent from 68,000 in 1947 to 73,900 in 1958. The industry was located in 281 SEA's.

The industry was size-oriented and tended to shift location. The shift was to large areas not prominent in the industry in the base year. Other characteristics of areas attracting growth in this industry are a high proportion of professional and managerial personnel in the work force (5), a good supply of processing water (30), and absence of major centers of higher education (34). The combination of a positive correlation with (5) and a negative correlation with (34) implies location in areas that must attract college graduates from other areas.

Economic strength in the variables leads to rating predictability for large areas as fair.

SIC 3571 - Computing Machines and Cash Registers

Establishments in this industry are primarily engaged in the manufacture of adding, calculating, tabulating, and bookkeeping machines; cardpunching, sorting, and tabulating machines; and cash registers (and electronic computers Employment in this industry increased 81 percent from 45,600 in 1947 to 82,400 in 1958. The industry was heavily clustered in 56 SEA's.

There was no evidence of concentration or size effects. The only area characteristics of importance were presence of a port facility (31), low level of urbanization (3), and the combination of these (42). The relationship between these characteristics was complex--i.e., areas with a port facility and with urban population less than 73 percent of total population received positive growth, areas with a port and with urban population greater than 73 percent suffered declines, and areas without port facilities regardless of percent urban population had no change in the industry employment. This resulted in a discontinuous function that cannot be the basis for prediction. The regressions are not valid for prediction.

SIC 3599 - Machine Shops

Establishments in this industry are primarily engaged in (a) producing machines and equipment parts not elsewhere classified from materials owned by them, or (b) performing machine operations on materials owned by others. Machi shops classified in this industry are characterized by their method of operation and may be primarily though not exclusively engaged in repair work. Usually they operate on a job or order basis and are equipped with power-driver metalworking machinery. Employment in this industry increased 115 percent from 57,100 in 1947 to 122,800 in 1958. The industry was spread throughout the country, being located in 493 SEA's.

The industry demonstrates strong size effects and balanced growth. The industry was drawn to areas with a high proportion of prime defense contracts (35). The size effects are clearly nonlinear as evidenced by the negative effects of having a large city in the area. Growth was also less in areas with a high proportion of employment in professions and managerial capacities. In the normalized version of the regression, the percent of employment in SIC 37, which included aircraft, substitutes for the size-related percent of military contracts.

The regressions are rated as being good predictive equations. The strong economic significance of the variables in the normalized version led us to raise the level of expectation for prediction in small areas.

SIC 3612 - Carbon and Graphic Products for Use in the Electrical Industry

Establishments in this industry are primarily engaged in manufacturing carbons; carbon, graphite, and metal-graphite brushes and brush stock; carbon or graphite electrodes for thermal and electrolytic uses; and other carbon, graphite, and metal graphite products. Employment in this industry increased 14 percent from 7,800 in 1947 to 9,000 in 1958. The industry was concentrated in 28 SEA's.

The industry was located in a small number of large areas in the base year and continued a tendency to concentrate in these areas although within this group of areas growth was not directly correlative with size. The growing portion of the industry appeared to search out areas weak in unionization (36). The areas of growth tend to be weak in unionization and strong in furniture manufacturing (61) although this is probably a proxy for other area characteristics. (The furniture industry was strong in the Great Lakes region, New York, Virginia, North Carolina, Kentucky, and Tennessee.)

The absolute version of the regression is weak and can be rejected. The normalized version is also weak although statistically significant.

SIC 3613 - Instruments for Indicating, Measuring, and Recording Electrical Quantities and Characteristics

Establishments in this industry are primarily engaged in manufacturing pocket, portable, panelboard, and graphic recording instruments for measuring electricity--voltmeters, ammeters, wattmeters, watthour meters, demand meters, and other meters and indicating instruments. This industry also includes establishments primarily engaged in manufacturing meter transformers and analyzers for testing the electrical characteristics of internal combustion engines, radio apparatus, etc. Employment in this industry increased 115 percent from 20,900 in 1947 to 45,100 in 1958. This industry was located in 110 SEA's.

Although the industry's growth was not directly related to the distribution of total employment, it was strongly affected by the distribution of defense contracts, another size-related variable. Defense is the principal market for the growth sector of this industry. This sector has different location determinants than the older parts of the industry as indicated by the strong shift that occurred in location patterns. Nonlinearity in size relationships is indicated by the fact that areas heavy in military contracts had somewhat less than a proportionate share of the growth if they were fast-growing areas with a large city (500,000 population) (41). The absolute version of the regression is rated as good for prediction in large areas. The normalized version is rated as poor for prediction.

SIC 3617 - Electrical Welding Apparatus

Establishments in this industry are primarily engaged in manufacturing electrical welding apparatus, electrode holders, and other welding apparatus. Employment in this industry increased 37 percent from 7,300 in 1947 to 9,900 in 1958. The industry was located in 70 SEA's.

The distribution of the industry remained reasonably stable. The industry was not particularly size oriented. The area characteristics that attracted change in this industry were a high percent of urban population in areas with port facilities (42); high per capita income in States offering location inducements (45); and a relatively high percent of employment in furniture manufacturing (61). A relatively low percent of employment in primary metals (69) and a low percent of employment in metal fabricating (70) were also characteristic of

areas attracting growth in this industry. These characteristics demonstrate a spatial pattern of growth rather than the pattern that is typical of metal and machinery industries in general. The regressions are rated as providing fair predictions.

SIC 3619 - Electrical Equipment for Industrial Use, n.e.c.

Establishments in this industry are primarily engaged in manufacturing industrial and commercial electrical apparatus and appliances not elsewhere classified, including electric heating units for furnaces and ovens. Employment in this industry increased 61 percent from 10,900 in 1947 to 17,600 in 1958. The industry was located in 125 SEA's.

The industry tended to shift location relative to the 1947 distribution. Growth was moderately related to size of area. The value of R^2 did not decline in the normalized version because size effects were consistent with the structural changes. Other characteristics of areas attracting growth in this industry were high density of roads per capita (33) and a high proportion of employment in furniture manufacturing (61). These variables may suggest rapid growth in Virginia and North Carolina. Predictability is poor for both large and small areas.

SIC 3661 - Radios, Radio and Television Equipment (except Tubes), Radar and Related Detection Apparatus, and Phonographs

Establishments in this industry are primarily engaged in manufacturing radio and television receiving and transmitting equipment, electrical and magnetic field detection apparatus, light and heat emission detection apparatus, object detection apparatus, other products associated with radio equipment including parts, and phonographs and accessories (except records). Employment in this industry increased about 65 percent from 178,600 in 1947 to 297,600 in 1958 The industry was located in 238 SEA's.

Growth in the industry was spatially balanced and was correlated with size of area. Growth in the period was first dominated by radio and television manufacturing for final consumption and later by production of military electronic apparatus. There was a strong relationship between growth in this industry and the distribution of prime defense contracts. Since this sector of the industry comprises mostly military subcontractors rather than prime contractors, the high correlation reflects the market orientation of the industry. The negative correlation with nearness to a big city (29) and employment in the transportation equipment industry (73) indicates that the relationship with military contracts is not fully linear, tending to be somewhat less than proportionate to percent of military contracts in the very large areas with previous heavy commitment in the aircraft industry. The normalized version for the regression is weak and includes variables such as percent of employment in tobacco manufactures that are either proxies or spurious. Only the absolute version is satisfactory for prediction. Because of economically strong variables, the rating for predictability has been raised to "good."

SIC 3662 - Radio Tubes

Establishments in this industry are primarily engaged in manufacturing radio receiving and transmitting tubes (and television tubes). Employment in this industry increased 235 percent from 27,700 in 1947 to 92,900 in 1958. The industry was concentrated in 88 SEA's.

Although growth occurred in a small number of areas, there was a tendency for growth to shift to new areas. Twenty of the 88 areas had large cities (population 500,000 or over) and on the average these areas had an additional 867 workers in the industry. In the normalized version, certain rural characteristics appeared to dominate with proportionate growth being associated with the percent of rural nonfarm population (3) and the percent of Class I (relatively wealthy) farms (10). This indicates a structural shift to well-to-do rural areas although absolute changes are still dominated by growth in big cities (or their suburbs) and by areas previously dominant in the industry. The regressions are poor for prediction.

SIC 3716 - Automobile Trailers (for Passenger Cars)

Establishments in this industry are primarily engaged in manufacturing trailers for attachment to passenger cars. Employment in this industry increased 123 percent from 9,300 in 1947 to 20,700 in 1958. The industry was located in 197 SEA's.

The industry demonstrated a strong tendency to concentrate further in areas of previous dominance. The size relationships are somewhat negative with growth occurring more in smaller areas. The area attributes that relate to growth in the industry are a high percent of military contracts (35), low construction costs (38), low percent of employment in printing and publishing (63), a high percent in chemicals (64), and a high percent of farms in the area having tractors (23). Some of these variables suggest a tendency for this industry to grow in less urbanized areas. The apparent conflict between distribution of military contracts and total employment, which are positively intercorrelated, suggests a geographic shift to the coastal areas and the Southwest.

SIC 3721 - Aircraft

Establishments in this industry are primarily engaged in manufacturing or assembling complete aircraft such as airplanes, gliders, dirigibles, and balloons but may also manufacture aircraft parts. Employment in this industry increased threefold from 146,600 in 1947 to 438,800 in 1958. The industry was heavily concentrated in 79 SEA's.

The industry showed spatial stability despite rapid growth. High unionization and low construction costs may indicate the requirements of the industry for unionized labor. To the extent that the industry sought new locations, construction costs appeared to be an important consideration. Distribution of prime military contracts has been omitted from the regression because growth in aircraft employment in the period was oriented to the military market and the

aircraft companies were the principal recipients of the prime contracts. Thus, the relationship would be tautological. Predictability is good for large areas and fair for small areas.

SIC 3811 - Laboratory, Scientific, and Engineering Instruments

Establishments in this industry are primarily engaged in manufacturing laboratory, scientific, and engineering instruments such as nautical, navigational, aeronautical, surveying, drafting, and instruments for laboratory work and scientific research--except surgical, medical, and dental instruments and microscopes and telescopes. Employment in this industry increased about 250 percent from 18,200 in 1947 to 63,400 in 1958. The industry was located in 135 SEA's.

The industry was strongly influenced by size effects. An expected correlation with distribution of military contracts (35) appears in the first version of the regression. Industry specific and size variables replace military contracts and population density in the second version although the variables do not suppress the economic significance of military contracts which provided the major market for products of the growth sector of this industry. Variable 61, percent of employment in furniture manufacturing in 1947, is undoubtedly a prox for certain area characteristics or geographic distributions. The regressions are poor for prediction.

SIC 3941 - Games and Toys (except Dolls and Children's Vehicles)

Establishments in this industry are primarily engaged in manufacturing indoor games and game sets and mechanical and nonmechanical toys including trains, toy guns, toy furniture, etc. Employment in this industry increased about 60 percent from 27,100 in 1947 to 43,000 in 1958. The industry was located in 216 SEA's.

The industry is overwhelmingly size oriented. The size relationship is highly nonlinear, however, as indicated by the fact that two size variables--military contracts (35) and total employment (43)--have positive signs; and two size variables--big city (29) and the interacting variable 48--have negative signs. Thus, the biggest areas did not grow proportionately to the moderate size areas. Factors other than size do not appear to emerge strongly; thus predictability is good for large areas but poor for small areas.

APPENDIX B. NOTES ON THE USE OF DUMMY VARIABLES

Several of the independent variables are dummy variables (29, 30, 31, 32, 34, and 44) and several are variables in which a dummy variable interacts with a continuous variable (40, 41, 42, 45, 46, and 48). Dummy variables are discrete variables taking a value of 1 or zero. They may be used to represent qualitative differences such as having or not having a particular resource; they may be used to represent quantitative variables such as income where it is thought that only broad income groupings are relevant. They may also be used to represent temporal shifts as between wartime and peacetime. Essentially, the dummy variable serves to partition the linear equation, shifting the value of the intercept for the two conditions represented by the dummy variable.

As an illustration, suppose change in employment in industry u is linearly related to the total employment in the area except that there is a downward shift in the function for areas that do not have port facilities (variable 31). Using the dummy variable as the shifter, the equation would be as follows:

$$Y = a_o + a_1X + a_2Z$$

where X is 1 if the area has a port and zero if not, and Z is the continuous variable representing total employment. Thus, if the area has a port, the equation becomes:

$$Y = (a_o + a_1) + a_2Z$$

whereas if the area does not have a port, the equation becomes:

$$Y = a_o + a_2Z$$

These two equations have the same slope but differ in Y-axis intercept by the size of the coefficient of the dummy variable. Thus, each area with a port receives more new employees in industry u by the constant amount a_1.

When a dummy variable interacts with a continuous variable, the slope as well as the intercept is shifted. If there is only an interaction variable, then only the slope is affected. For example, if change in employment in industry u is related to total employment only when the area has a port, then the equation is as follows:

$$Y = a_o + a_1(XZ)$$

If X = 1,

$$Y = a_o + a_1Z$$

and if X = 0,

$$Y = a_o$$

Thus, all areas receive a_o additional employment in industry u but only in area with ports is additional employment in industry u related to total employment i the area.

The dummy variable may also enter the equation as a separate term in combination with the interacting variable. In this case, both the Y-axis intercept and the slope of the regression are affected.

Thus,

$$Y = a_o + a_1X + a_2(XZ)$$

If $X = 1$,

$$Y = (a_o + a_1) + a_2Z$$

and if $X = 0$,

$$Y = a_o$$

* U.S. GOVERNMENT PRINTING OFFICE : 1968 O—311-829