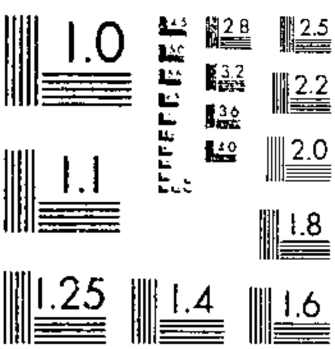
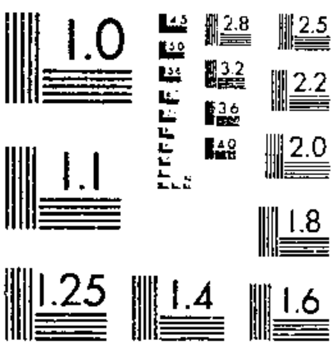
TB 1529 (1975) USDA TECHNICAL BULLETINS UPDATA
PREDICTING EMPLOYMENT IN FOUR REGIONS OF THE WESTERN U.S.
BENDER, L.D. 1 OF 1

START



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963 A



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

PREDICTING EMPLOYMENT IN FOUR REGIONS OF THE WESTERN UNITED STATES

LLOYD D. BENDER

UNITED STATES DEPARTMENT OF AGRICULTURE
ECONOMIC RESEARCH SERVICE
IN COOPERATION WITH MONTANA STATE UNIVERSITY
AGRICULTURAL EXPERIMENT STATION

TECHNICAL BULLETIN NO. 1529

PREDICTING EMPLOYMENT IN FOUR REGIONS OF THE WESTERN UNITED STATES. By Lloyd D. Bender, Economic Development Division, Economic Research Service, U.S. Department of Agriculture. Technical Bulletin No. 1529.

ABSTRACT

Local employment multipliers (ratio of the number of jobs in service and support sectors to the number in each basic industry) in rural communities vary from place to place and are different for each industry and scale of industry. The predicting model can be used to estimate local service employment as employment in basic industries changes in each community. The prediction model works best in regions which conform to the assumptions of central place theory. The model should be refined further and tested in other regions. The resultant statistical model can be applied as a planning tool in local communities. Rapid changes in employment, population, and infrastructure needs in rural communities accompany the trend in manufacturing decentralization and energy resource development.

The study tests a model designed to predict the local service employment associated with employment in basic industries within local communities of four regions of the western United States. Rationale for the model is economic base and central place theory. The hypotheses are that local multipliers for agriculture, mining, manufacturing, and transportation employment are a function of the industry, its size, and distance from a major trade center. The model is applied to data for nonmetropolitan counties in the Plains, Mountain, Intermountain, and Coast regions. Results of statistical tests conform to theoretical expectations in those regions where conditions correspond to the assumptions of central place theory.

Key words: Regional analysis, Central place, Economic base, Nonmetropolitan communities, Western regions, Employment prediction, Regression, Analysis of variance, Cross-section, Census data.

Washington, D.C. 20250 November 1975

CONTENTS

Acknowledgment is due Peggy Humphrey, Jeri Thieme, Stuart Townsend, and Robert Lott for assistance in analysis of data; Oscar Burt for statistical advice; and Robert Coltrane for aid in conceptualizing the model, guidance during the project, and critical review of the results.

SUMMARY

This study investigates procedures to estimate changes in local ancillary employment (service or support jobs) in response to changes in the basic sectors of the economy of a community. A regression model draws upon economic base and central place theory. Local multipliers for the basic industries of agriculture, mining, manufacturing, and transportation each are hypothesized as dependent upon the scale of the basic activity and distance from a regional trade center. The unit of observation is the nonmetropolitan (nonmetro) counties of the western United States divided into four regions: Plains, Mountain, Intermountain, and Coast.

The statistical model was designed to test economic base and central place theory for prediction. Both a priori expectations and standard statistical tests of parameter estimates were used as a basis for judging the efficacy of the model. In general, signs of the parameter estimates, size of the multipliers, and statistical significance of parameter estimates tend to support central place and economic base theories. In cases where signs of coefficients were illogical, the test of significance proved inconclusive; that is, these variables could be dropped from the model without affecting predictive capability.

The statistical model conformed least to expectations when applied to the Intermountain and Coast regions. In these regions, the types of economic activity and their spatial distribution conform less to assumptions underlying central place theory compared to the other two regions. Results of the model applied to the Plains region yielded the best statistical results and coincided with theoretical expectations.

Statistical tests were for the hypothesis that each industry, its size, and its distance from a trade center are important to prediction. In cases where statistical tests were inconclusive, an evaluation of results indicated conformity to theoretical expectations. Each industry clearly influenced the level of ancillary support employment differently. Effect of distance from a major trade center proved statistically conclusive in three regions. An evaluation of resulting multipliers showed dramatic interaction between industry and distance. Coefficients of variables measuring effect of industry size were significant in only one of the four regions. Ancillary employment per unit of basic employment did tend to decline as industry size increased in all but three cases.

Basic purpose of the model is to predict ancillary employment. Evaluations of the model in terms of multipliers were made for theoretical reasons. Subject to certain qualifications, the model appears to describe systematically the relationships between basic employment and ancillary employment. In cases where coefficients for basic industries are not significant, that finding is useful. It implies that little confidence can be placed in predictions of local employment change for these cases. Policy makers should know the estimating error associated with predictions.

PREDICTING EMPLOYMENT IN FOUR REGIONS OF THE WESTERN UNITED STATES

by Lloyd D. Bender, Economist
Economic Development Division
Economic Research Service

Events of the past decade have heightened the need for techniques to predict and evaluate anticipated local developments, especially in nonmetro areas. Examples are the long-term trend toward decentralized manufacturing throughout the country and the more recent energy resource development in the Northern Great Plains. Local evaluation and planning attempt to accommodate the infrastructure needs and fiscal requirements associated with new industry and the externalities it creates in a community.

Predicted impacts of changes in the economic base of a community are necessary in order to implement local planning. Communities may wish to control development. In other instances, they may foster development through purposive actions such as the issuance of revenue bonds for industry subsidies. In either case, added community costs and revenues generated by development are primary considerations. Accurate predictions of local employment impacts are the basis for estimating development infrastructure needs and revenue flows.

Changes in total employment in an area are viewed in this study as a result of changes in employment in basic economic activities. Basic activities are those bringing export revenue or transferring income into the community from outside. Thus, levels of basic activity in a locality are determined by forces outside the community.

Any change in a basic activity produces direct, indirect, and induced changes in total employment. The direct effect is defined as the change in employment in the basic activity. Indirect employment effects stem from basic activities which purchase supplies and services from other local firms and industries. Finally, the induced effect is the change in employment in retail and other service firms, which depends upon the level of both basic and indirect employment. The change in ancillary support employment (indirect and induced effects) per unit change in basic employment is called the multiplier in this study. The change in total employment is equal to the sum of changes in basic employment and ancillary employment.

This general theory has been made operational in impact analyses for both large and small regions. But current techniques, of which input-output is an example, are often expensive or lose accuracy when applied to local, rural economies. In this study, the theory is rationalized and tested as it applies to local economies.

OBJECTIVE AND PROCEDURES

This study develops and tests an empirical model for estimating ancillary support employment in nonmetro communities once basic employment is known. It uses multiple regression on cross-section employment and allied data to estimate parameters for nonmetro counties in the West. Regions analyzed are: Plains, Mountain, Intermountain, and Coast (fig. 1). Published secondary data were used in order to minimize costs, provide ease in updating, and give continuity between regions.

The hypothesis to be tested is that the employment multiplier computed for each basic industry will vary by community (location), type of industry, and scale of industry. Model specifications permit computation of a multiplier for each industry which takes on different values. The hypothesis assumes that a multiplier incorporating these elements is superior to any single-valued multiplier representing the relationship between basic and ancillary employment.

THE EMPLOYMENT ESTIMATION MODEL

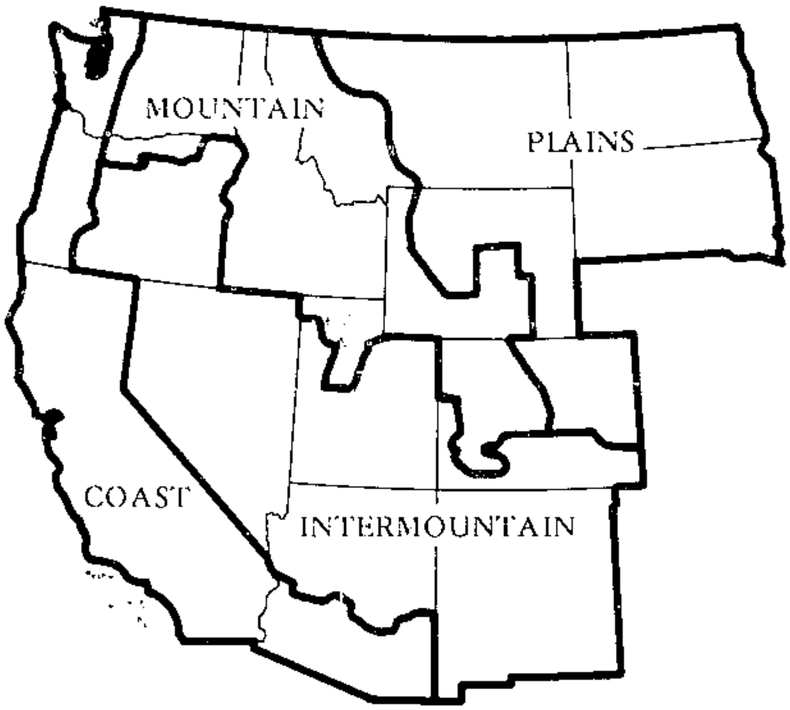
The statistical model draws upon inferences from economic base and central place theories. Economic base theory asserts that autonomous changes in basic activities result in a multiplied change in total employment. The changes in basic activities are influenced by supply and demand forces outside the local community. 1/ Central place theory implies that the multiplier may be different at each central or city location. The reason is that each type of central place or trade center performs different economic functions, which in turn in-

---

1/ Charles M. Tiebout, The Community Economic Base Study, Committee for Economic Development, Supplementary Paper 16, Dec. 1962, pp. 1-82.

Figure 1

Study Regions of the Western United States



fluence local spending patterns. In general, the number of different types of economic functions found in a central place increases with the population size of the place. 2/

Implications of these theoretical considerations for this study are:

1. Multipliers which vary with the location of the community in economic space are required, i.e., multipliers should reflect local economic conditions.

The propensity of firms and consumers to spend and respend locally is the fundamental multiplier concept. Spending and respending by people and firms in their own communities are determined by prices and the range of services and goods offered locally in relation to other locations. Small local service centers tend to cater to the everyday demands of consumers and firms in the immediate vicinity. But, the local service center is often by-passed by consumers for larger trade centers offering specialized and complex services and

2/ Brian J. L. Berry and William Garrison, "Recent Developments in Central Place Theory," papers and proceedings of the Regional Science Association, 4:107, 1958, pp. 109-120.

capital goods. The larger, regional trade center has a complex economy which fulfills local demands plus the specialized needs of firms and people in the whole region. The location of a place with respect to its regional trade center is expected to reflect the propensity of firms and individuals to spend locally.

2. Multipliers for each basic industry are desirable because anticipated impacts on a local community are likely to be different for each industry.

Each basic industry in a community is expected to generate different levels of ancillary support employment because of different derived demands. Each basic industry likely uses different quantities of factor inputs and marketing services. And, chances are that only part of the inputs and services can be purchased locally; the remainder must be purchased in other service centers. The proportion purchased locally is influenced by the range of services provided locally and the transportation cost of moving the services and inputs from the center in which they are available for purchase. The distance between the community and this center (assumed to be a regional trade center in this study) serves as an estimate for the availability and competitiveness of local services.

3. Multipliers will vary with the scale of industry.

The multiplier for each basic industry will depend in part on the size of that industry--the scale effect. In this case, the effect of additional basic industry employment on ancillary employment will depend upon the cost structure of the supportive activities. Prices of goods and services in small communities may be higher than in regional trade centers. The difference will be the cost of travel to a trade center and transfer costs of the goods or service. Service firms in small communities may not be large enough to attain low-cost operations. High levels of derived demand should result in lower costs and, eventually, lower supply prices. Local supply firms would become competitive with firms in regional trade centers; hence, business would tend to be conducted locally. Fewer resources per unit output would be required if this theory is valid.

The Model

The dependent variable--ancillary employment--is total employment minus agricultural, mining, manufacturing, and that portion of transportation employment considered basic. These sectors are exogenous elements of the local economy of primary interest in this study. The amount of ancillary employment associated with each industry theoretically is a function of type of basic activity, scale of industry, and location of the industry in relation to regional trade centers. As will be explained below, the specification of the model reflects these relationships. Total employment estimates associated with a given level of basic employment can be calculated by summing ancillary employment estimates and basic employment.

Additional independent variables in the statistical equation are linear expressions of variables which adjust for other autonomous influences on the economic bases of the local economies. These variables reflect activities which play a support role when considered a part of a regional economy but are basic to the local economy. A concentration of government activities, or the presence of universities, medical services, or institutions are examples. While they might serve a local function, a major portion of their clientele are from outside the local community and revenue flows into the local economy because of their location. Furthermore, travel and tourist-related activities constitute an important part of a local economic base.

The independent variables accounting for these activities are defined precisely below and are explained only briefly here. An inordinate number of local government employees indicates local support of activities which serve a regional clientele. In addition, a school may serve a State, region, or multicounty district. The number of students in group quarters and a dummy variable for presence of students in group quarters is designed to measure this basic activity. The influence of penal and other institutions such as old age homes is indicated by the number of persons in group quarters in these institutions. The importance of travel and related activities is measured by the number of hotel and motel employees in a county and a dummy variable where these data are not reported.

Finally, the variable median family income corrects for the effect of income level on the demand for local services. A higher income in a rural setting can initiate conflicting influences. On the one hand, it allows consumers to travel greater distances for goods and services. At the same time it tends to increase demand for the convenience of local goods and services.

The independent variables are defined as:

A = 1970 agricultural employment

DA = the cross product of distance and agricultural employment, A 3/

D^2A = the cross product of distance squared and agriculture 3/

A^2 = the square of agricultural employment, A 4/

M = 1970 mining employment

DM = the cross product of distance and mining employment, M 3/

D^2M = the cross product of distance squared and mining 3/

M^2 = the square of mining employment 4/

3/ Distance is miles from the largest city in a county to a regional trade center. Regional trade centers are listed in app. table 8.

4/ The squared employment terms in this model are hypothesized to measure effects of scale of industry.

F = 1970 manufacturing employment

DF = the cross product of distance and manufacturing employment, F 3/

D^2F = the cross product of distance squared and manufacturing 3/

F^2 = the square of manufacturing employment 4/

T = 1970 basic transport employment 5/

DT = the cross product of distance and basic transport employment, T 3/

D^2T = the cross product of distance squared and basic transportation 4/

T^2 = the square of basic transport employment 4/

V = basic local government employment in 1970 calculated in the same manner as the variable T above 6/

I = 1970 institutional population 7/

C = number of college students in group quarters, 1970

Cd = dummy for location of colleges where 1 = no students

H = number of motel and hotel employees in 1970

Hd = dummy for location of motels and hotels where 1 = no motel or hotel employees in 1970

Y = 1970 median family income

All data used in the study were taken from the 1970 U. S. Census of Population and County Business Patterns.

5/ The transportation variable was calculated by using a location quotient as follows: $T_j = T'_j - (X_j k)$; k = 1970 transportation employment in a region's counties divided by total employment in those counties, X_j = 1970 total employment in county j, and T'_j = transport employment in county j. The value of T_j was constrained to nonnegative values. The difference between total 1970 transportation employment in county j and T_j was added to ancillary employment in county j. This variable was meant to account for railroad support centers.

6/ The local government variable was calculated by using a location quotient in the same manner as the basic transportation variable was calculated. In this case, no adjustment was made in ancillary employment because the source of data was not a part of the census employment series.

7/ Institutional population is the sum of penal, old-age home, and other persons in group quarters in 1970.

Specifications for Agriculture, Mining, Manufacturing, and Transportation

Nonmetro communities tend to be specialized in one or more of four basic industries: agriculture, mining, manufacturing, and transportation. These four are assumed to be especially important in this study and are designated by the symbols A, M, F, and T, respectively.

The specification of each of these four in the prediction equation (in order to take into account type of activity, scale, and distance) is:

Equation 1

$$ANC_{ij} = b_1X_{ij} + b_2X_{ij}D_j + b_3X_{ij}D_j^2 + b_4X_{ij}^2$$

where: ANC_{ij} = ancillary support employment associated with basic industry i in county j;

X_{ij} = basic employment in industry i in county j;

i = agriculture, mining, manufacturing, or transportation;

D_j = distance county j is from a regional trade center;

b_k = estimated parameters; and

j = county observation.

The X_i variable assures that the multiplier for each industry can be different. The X_i^2 variable specifies that the multiplier for each industry can vary depending upon the scale of industry activity. The terms X_iD and X_iD^2 imply that distance from a major trade center has a nonlinear effect upon ancillary employment. Each of the four basic industries are entered in a full model in the above manner.

Theoretical Evaluations of the Multipliers

Since each basic industry is independent of the other, that portion of the full model applying to one industry (as expressed in equation 1) can be evaluated by itself. Theory suggests a graph of the relationship between ancillary support employment associated with an industry and employment in that industry as illustrated in figure 2 for a given distance (D).

Equation 1 specified as a polynomial with cross-product terms has several interpretive features. First, the intercept value is zero implying that no ancillary support employment is required without basic industry employment. Next, the sign of the term b_4 determines the inflection of the curve of ancillary employment associated with industry X_i; that is, a negative sign results in the curve increasing at a decreasing rate to some point a. *A priori* point a should be beyond the range of the relevant empirical data. Finally, the curve will be different at each distance as illustrated in figure 2 (D and D').

The average multiplier is the slope of a radius vector passing through the origin to a point on the curve as shown by r in figure 2. Its value at each level of X_i is the value of equation 1 divided by X_i:

Equation 2

$$\frac{ANC_{ij}}{X_{ij}} = \frac{b_1 X_{ij} + b_2 X_{ij} D_j + b_3 X_{ij} D_j^2 + b_4 X_{ij}^2}{X_{ij}}$$

The average multiplier has a different value at each level of X_i and for each distance as required by the hypothesis.

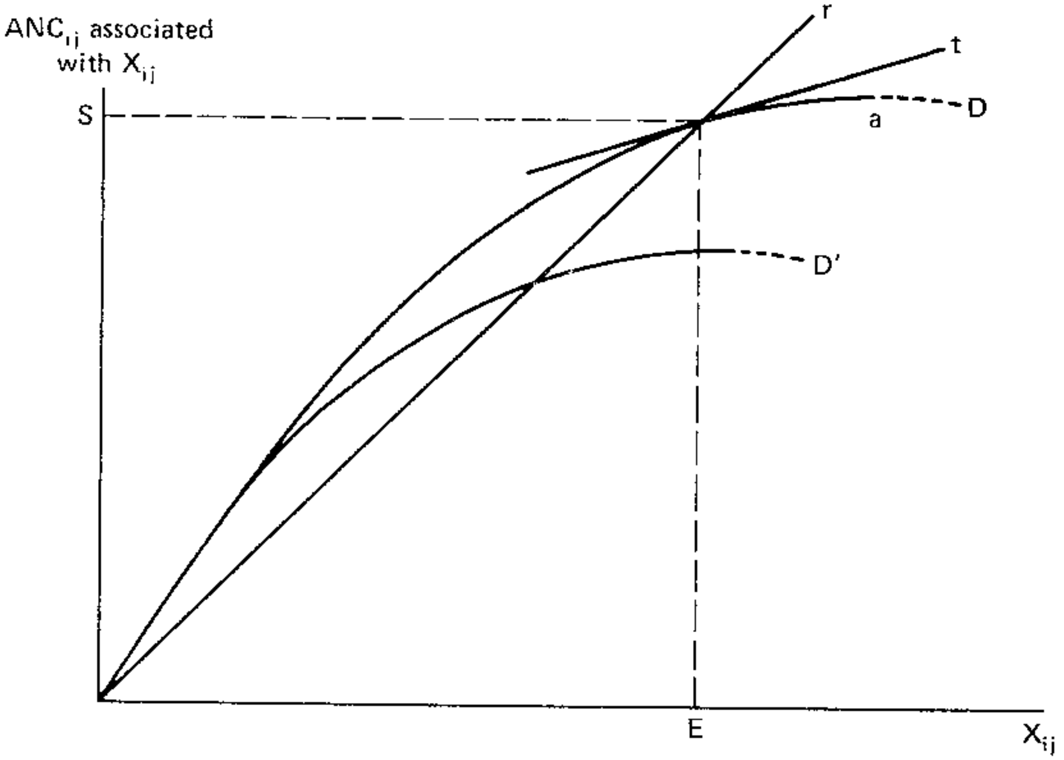
The value of the partial derivative of equation 1 for a given D value is the incremental multiplier. The partial with respect to X_{ij} is:

Equation 3

$$\frac{\partial ANC_{ij}}{\partial X_{ij}} = (b_1 + b_2 D_j + b_3 D_j^2) + 2b_4 X_{ij}$$

Figure 2

Theoretical Impact of Basic Industries on Ancillary Employment



In graphic form, it is the slope of the tangent to a point on the total curve as shown by t in figure 2. The incremental multiplier also has a different value at each level of X_{ij}.

Graphics of figure 2 demonstrate that a multiplier appropriate for one level of output cannot be applied to any other level. To do so would be equivalent to following ray r out from the zero point of X_{ij}. Therefore, the model cannot be used to generate multipliers at one level of X_{ij} which are then applied to estimate a higher level of X_{ij}. The utility of the model is that ancillary employment is estimated without explicit use of multipliers. Multipliers in this study are for expository purposes only. The concept is useful to frame a model which then can be used to test hypotheses.

EMPIRICAL TEST OF THE MODEL

The model was tested using data for the four regions. 8/ The regional division takes into account physiographic features (soil, climate, and temperature configurations), economic activities, and settlement patterns. 9/ Selected characteristics for the nonmetro counties in these regions are shown in table 1. 10/ Each region has unique characteristics which are the basis for the delineation. Counties in the Plains region are dominated by agriculture which is spread more or less evenly across the region. The region is predominately rural, and towns are small. Natural barriers to highway travel are minimal in contrast to the Mountain and Intermountain regions. Mountain and Intermountain regions are also predominately rural but are not as dependent on agriculture as the Plains. Mining, manufacturing, and transportation are also important, but each activity tends to be specialized in local economies with an uneven spatial distribution. The Coast region tends to be dominated by metro places. The nonmetro counties tend to be large both in size and population with location near metro centers. The economies of Coast counties are more diverse than those of the other three regions. Conditions in each region may not fit the assumptions underlying central place theory if spatial distribution of basic activities is uneven or natural travel barriers exist. Since conditions in the four regions are so different, they provide an excellent test for the model.

8/ All North and South Dakota counties were included for adequate representation in the Northern Great Plains.

9/ Suggestions of Calvin Beale were used extensively to make this delineation. For reference, see Donald J. Bogue and Calvin L. Beale, Economic Areas of the United States, New York: The Free Press of Glencoe, Inc., 1961.

10/ Nonmetro counties are defined as those which are not in or on the border of an SMSA (Standard Metropolitan Statistical Area) in 1970. This designation excludes the dominant effect of urban centers on the data. An SMSA is defined by the Census Bureau as a county or group of contiguous counties (except in New England) containing at least one central city or twin cities with at least 50,000 population.

Table 1--Number of counties and average nonmetro county characteristics, four regions, western U.S., 1970

Item	Region			
	Plains	Mountain	Inter-mountain	Coast
Number of counties	191	129	107	104
Metro 1/	10	11	6	37
Nonmetro	181	118	101	67
	Nonmetro county averages 2/			
Population of largest city	4,796	6,410	6,779	11,031
Miles to trade center	83	72	111	41
Total population	9,765	14,077	15,350	40,668
Urban	3,990	6,761	7,538	19,632
Rural nonfarm	3,484	5,767	6,720	18,920
Rural farm	2,292	1,550	1,091	2,116
Total employment	3,475	5,158	5,018	13,660
Agriculture	744	641	426	1,327
Mining	62	157	293	151
Manufacturing	212	635	425	2,403
Transportation	198	352	334	809
Wholesale & retail trade	700	1,032	961	2,645
Professional & related	682	956	965	2,367
All other	878	1,385	1,614	3,960
Median family income(dollars)	7,200	8,367	7,472	8,738

1/ Counties in or on the fringe of designated SMSA's (see SMSA definition in text footnote 10).

2/ Means of county data.

Source: U.S. Census of Population, Bureau of the Census.

Region Differences

A standard analysis of variance was used to test whether data for contiguous regions could be pooled. 11/ The results (app. table 1) show significant differences for all combinations except one, indicating that data for the four regions should not be pooled for purposes of analysis. The equation tends to have different explanatory properties for individual versus pooled regions. Thus, the results which follow are reported for each region.

Metro and Nonmetro Differences

Data also were tested for the purpose of determining whether metro and nonmetro counties within each region should be combined for analysis. 12/ The F ratios all indicate a much higher additional mean square error for the n+m equations where all observations are combined than can be attributed to chance at $F_{.01}$ level of probability. 13/ These tests indicate the observations for metro and nonmetro counties should not be pooled for analysis because the two sets of data do not obey the same relations. Thus, the remaining discussion and analyses will treat only the nonmetro counties of the four regions.

Multicollinearity

Parameter estimates of equations exhibiting high levels of interdependence among the independent variables often are unstable values; that is, a change in the model specification (for example, deleting or adding variables) will produce different parameter estimates and levels of confidence of those estimates. Two measures of multicollinearity were used in this study: The simple correlation matrix, and coefficient of determination (R^2) when each independent variable was regressed upon the others in the model.

For the most part, multicollinearity is not a severe problem. 14/ Of the ratios in app. table 2, only 16 are .70 or larger, the largest three being .88,

---

11/ J. Johnston, Econometric Methods New York: McGraw-Hill Book Co., 1972, p. 198. The estimating equation was fitted successively to the pooled data for each combination of contiguous regions and the resulting mean sum of squares residual was related to the sum of squares obtained when the same equation was applied to each region separately and then summed for comparable sets.

12/ The conventional test described in text footnote 11 was used for the Coast region. For the other three regions, the number of the m, SMSA counties, < k, the number of variables in the equation. In this special case, the method suggested by Johnston (p. 198) is used.

13/ The F ratios are: Plains 6.30; Mountain 16.32; Intermountain, 40.37; and Coast, 197.88.

14/ Since the cross-product and squared terms for each of the four basic industries, and dummy variables for colleges and motels and hotels are known to be interrelated with the respective base variable, and are to be interpreted as a set, they were treated as shown in app. table 2.

.87, and .85--all in the Plains region. The simple correlation coefficients (r) presented in app. tables 3-6 confirm that individual variables are not highly correlated with each of the other independent variables, except for the cross-product, squared, and dummy variables.

Interrelationships Between Distance and Industries

The distance is included in the estimating equation to reflect the hypothesis that distance affects availability and cost of local goods and services. This is part of the basic rationale for expecting a location effect on the multipliers. In order for the distance variable to serve as a surrogate measure of propensity to consume locally, it is a necessary but not sufficient condition that location with respect to a trade center be independent of the size and mix of basic sectors. If the distance variable were associated closely with the size of each industry in a systematic pattern, then little would be gained by including it in the estimating equation; that is, it would be a redundant variable reflecting only the relationship between size or mix of industry and ancillary employment. If distance were associated with the relative size of an industry, then it would make a contribution to the estimating equation but not as hypothesized, i.e., distance would merely reflect differences in the mix of industries in economic space.

The independence of distance from a major trade center and size of each of the basic industries was tested. Simple correlations between distance, and absolute and relative size of agriculture, mining, manufacturing, and basic transportation employment show that only very small relationships exist (app. table 7). The largest r was 0.28. It can be concluded that no systematic relationship exists between size and mix of basic industries, and distance to major trade centers in the western regions. Because of these results, the hypothesis which can be tested is that multipliers will vary from one community to another because of differences in the cost of obtaining goods and services from a designated major trade center. The assumed effect of distance entered as a quadratic in the estimating equation is that it does reflect propensity to consume locally.

Statistical Results

Regression results are shown in table 2. Each equation had extremely high overall F values. 15/ The closeness of fit is indicated by R^2's ranging from 90 to 97. The signs of the coefficients in all equations were reasonable except in three instances. In each of these cases, the coefficients of the variables were not significant. Unexpected signs in two cases were for the coefficient for the hotel dummy (Hd) in the Plains and Intermountain regions, implying that counties without hotel accommodations reported have higher levels of ancillary employment than counties with accommodations. Third, the transport coefficient in the Coast region has a negative sign on the first term,

15/ The F ratios are: Plains, 204; Mountain, 47; Intermountain, 30; and Coast, 34.

Table 2--Regression equations, nonmetro counties, four regions, western U.S., 1970

Variable	Region 1/			
	Plains	Mountain	Intermountain	Coast
	Regression coefficients			
Agriculture	**	**		**
A	.84520	2.42893	1.97798	2.68358
DA	.01264	-.02096	-.01548	-.02581
D^2A	-.00004	.00012	.00003	-.00010
A^2	-.00039	-.02096	-.00008	-.00009
Mining	*	**	**	
M	3.52746	1.60376	1.33238	1.16875
DM	-.02854	-.00438	-.00312	.06632
D^2M	.00018	-.00011	-.00003	-.00030
M^2	-.00152	.00007	.00001	-.00114
Manufacturing	**	**	**	**
F	5.57243	3.53614	3.47995	1.25475
DF	-.04394	-.03790	.00992	-.01163
D^2F	.00019	.00013	-.00005	.00023
F^2	-.00010	-.00025	-.00061	.00002
Transportation	**			
T	8.27339	1.32613	6.19336	-6.33368
DT	-.04770	.09378	-.02104	.02410
D^2T	-.00027	-.00045	-.00042	-.00054
T^2	-.00419	-.00018	-.00121	.01347
Local gov't. V	1.43666	2.26497	2.06611	3.00432
Institutions I	1.46407**	.40367	1.62298*	1.37151*
Colleges				
C	.78834**	.90484**	1.30778*	.86221
Cd	-832.61111**	-1,397.50301**	-557.00075	-3,427.90654**
Hotels				
H	11.42315**	.89831	12.91496**	6.43467
Hd	729.60372	-1,041.61287	476.81607	-626.14250
Income Y	.10357	.15230	.17253	.34749
Intercept	-751.13764	878.31682	-571.35482	1,540.27082
R^2	.97	.92	.90	.95

1/ Significant F values when variable or set of variables excluded from full models shown by $F_{.01}$** and $F_{.05}$*.

but neither the complete set nor any one taken singly is significant. The lack of significance for the coefficients of the variables with unexpected signs means that they can be excluded from the model without reducing estimating efficiency, and that the hypothesis with respect to that variable remains unconfirmed.

Tests of Hypotheses

The analysis is a statistical test of the effect of agriculture, mining, manufacturing, and basic transcontinental transportation activities on ancillary support employment. These effects are hypothesized to be a function of the size of each activity, and the distance from a regional trade center.

The statistical test of these hypotheses is a standard analysis of variance. 16/ Each set of the elements was dropped from the full equation for each region. The resulting increase in residual sum of squares was compared to that of the full models. For the industry test, the complete sets (X, XD, XD^2, and X^2), were deleted from the full models. The test for the effect of size of industry (that is, the scale effect) was accomplished by dropping all the X^2 terms for agriculture, mining, manufacturing, and transportation. By the same token, the test for the effect of distance was made by omitting the XD and XD^2 terms from the equations in each region.

Industry

The test for each basic industry (agriculture, mining, manufacturing, and transportation) is shown in table 3. Each set of industry coefficients was significant for the Plains region. Transportation proved not to be significant in the remaining three regions. Mining activity did not contribute significantly in the Coast region. Finally, coefficients reflecting agriculture were not significant within the Intermountain region.

The lack of significance of some of the industry variables may have been due to insufficient observations. Mining employment in the Coast region, for instance, was limited. Only 8 of the 67 counties had mining employment levels of 150 or more employees, and only 4 registered 550 or more. Only three counties in that region had mining activities constituting 10 percent or more of total employment. Agriculture in the Intermountain region is a different matter. Counties with agricultural employment over 500 numbered 33 out of 101. It may be that agricultural employment in an area of extensive rangeland agriculture typified by much of the Intermountain region simply does not require significant amounts of local support. This model does not permit a conclusion about the economic impact of these activities in the two regions.

16/ Johnston, op cit., p. 198.

Table 3--Statistical tests for effect of industry, nonmetro counties, four regions, western U.S., 1970

Variable set deleted	Region			
	Plains	Mountain	Intermountain	Coast
	F ratios 1/			
A, DA, D^2A, A^2	6.12 (4,157)**	5.93 (4,94)**	0.98 (4,77)	7.05 (4,43)**
M, DM, D^2M, M^2	3.08 (4,157)*	3.82 (4,94)**	4.33 (4,77)**	0.25 (4,43)
F, DF, D^2F, F^2	60.35 (4,157)**	19.92 (4,94)**	8.54 (4,77)**	18.49 (4,43)**
T, DT, D^2T, T^2	7.72 (4,157)**	1.86 (4,94)	1.84 (4,77)	0.81 (4,43)

1/ Bracketed numbers are degrees of freedom associated with each F statistic.

*Significant at $F_{.05}$.

**Significant at $F_{.01}$.

Size of Industry

The effect of size of industry on ancillary support employment is tested by deleting the squared terms (X^2) of each industry. The results reported in table 4 are clear. Industry size in each region appears important only in the Plains region. The hypothesis is confirmed for only one of the four regions. But, since it has theoretical validity, it is reasonable to expect that a scale effect does exist but that its influence has not been captured completely in this analysis. Additional research is needed to test the hypothesis in the remaining three regions.

Table 4--Statistical tests for effect of industry size, nonmetro counties, four regions, western U.S., 1970

Variables deleted	Region			
	Plains	Mountain	Inter-mountain	Coast
	F ratios			
A^2, M^2, F^2, T^2	12.53**	1.48	0.69	1.36

**Significant at $F_{.01}$.

Distance

The distance hypothesis is confirmed for three of the four regions as shown in table 5. In the Coast region, large metro areas are within commuting distance of most nonmetro counties. The average distance is only 41 miles.

Table 5--Statistical tests for effect of distance, nonmetro counties, four regions, western U.S., 1970

Variables deleted	Region			
	Plains	Mountain	Inter-mountain	Coast
	F ratios			
AD, AD^2, MD, MD^2, FD, FD^2, TD, TD^2	4.15**	4.25**	8.05**	0.89

**Significant at $F_{.01}$.

Discussion of Results

This study tested the model for prediction of local ancillary employment. Since the degree of multicollinearity is low among the independent variables, the equations also allow estimates to be made of changes in ancillary employment which are associated with the set of variables for each major industry. This capability is tempered by three considerations. First, some variables do not exhibit statistical significance. The data used in the analysis do not allow estimates with a high degree of reliability to be made in some regions. The second reason relates to the sign associated with the coefficients. In each case where the signs are opposite those expected from theoretical reasoning, the coefficients are not significant.

Finally, special care should be exercised in using the parameter estimates for prediction when large changes are expected in basic employment. The average 1970 mining employment in counties of the Plains region was only 62. Attempting to estimate the ancillary employment impact of a new mine employing 500 workers may be extending the coefficients for mining beyond the limits of the data. Estimates from large changes will have a low degree of reliability.

The analysis does imply that a procedure of using a single multiplier coefficient for all basic industries in every locality is inappropriate The results indicate that local multipliers can vary by industry and locality. Use of a regional or State coefficient to make predictions of local impacts could prove misleading for planners when this is true.

In order to use these questions for prediction, the full model should be used in the following steps:

1. Estimate ancillary employment in period t.

2. Estimate ancillary employment in period t+1.

3. Subtract ancillary employment in period t and t+1 to get change in ancillary employment predicted.

4. Add the anticipated changes in basic employment to the predicted changes in ancillary employment.

5. Estimated total employment in period t+1 will be actual employment in a county in period t plus the predicted change in employment estimated.

Average multipliers calculated for mean values in each region show the variability of results by industry and region (table 6). The mean multiplier reflects the added local ancillary employment per unit of basic employment when all variables are set equal to their mean values. The values for agriculture seem reasonable and are similar in those regions where this variable set was significant. These values range from 1.33 to 1.41. The values for mining in the Plains (2.31) and the Mountain (0.73) regions are substantially different. Differences in the type of mining activity could be reasons for these large differences. Theory does not indicate the magnitude to be expected. The remaining most important industry is manufacturing. Average multiplier values range from 1.21 to 3.71. The latter value seems high on an intuitive basis. Finally, the average multipliers calculated for the Intermountain and Coast regions demonstrate illogical results. In each case, the multiplier has a negative sign. In both instances, however, the variable set is not significant.

The influence of industry, distance, and scale were all statistically significant in the Plains region. For this reason, graphs of ancillary employment associated with each industry are presented for the Plains region in figure 3. Parenthetically, these evaluations also demonstrate the range of data for which predictions are reasonable. Graphed are the results for each major industry up to one standard deviation from the mean values for the Plains counties. The reader is cautioned that the mean of the distance variable for the Plains region is 83.4 miles.

The effect of distance is shown clearly. For agriculture, for instance, a given level of employment will require greater associated ancillary employment as distance increases up to some theoretical limit. The theoretical limit represents an interstice intermediate between two trade centers. But the effect of distance is not the same for each industry. For manufacturing at any given level, local ancillary impacts are lower at 100 miles than at either 50 or 150 miles.

Finally, the effect of size of industry on the average multiplier is presented in table 7. Shown are the average multipliers for industry employment up to approximately one standard deviation above the mean employment level. The variable representing industry size (the X_i term in the estimating equation) was significant for the Plains region only. However, multipliers are shown also for the remaining regions. In most cases the patterns are the same.

Table 6--Average multipliers at mean values, nonmetro counties, four regions, western U.S., 1970

Variable and region	Values of variables 1/		Average multiplier
	Mean	Standard deviation	
Agriculture			
Plains	744	477	1.33
Mountain	641	610	1.41
Intermountain	426	377	2/ .59
Coast	1,327	2,058	1.34
Mining			
Plains	62	193	2.31
Mountain	157	458	.73
Intermountain	293	697	.62
Coast	151	496	2/ 3.21
Manufacturing			
Plains	212	565	3.21
Mountain	635	783	2.04
Intermountain	424	573	3.71
Coast	2,403	2,546	1.21
Transportation			
Plains	36	133	2.27
Mountain	55	192	2/ 5.74
Intermountain	45	118	2/ 3/ -1.38
Coast	69	130	2/ 3/ -5.32

1/ Mean distances are: Plains, 83.4; Mountain, 72.1; Intermountain, 111.1; and Coast, 41.0.
2/ Variable sets were not significant at $F_{.05}$ (see table 3).
3/ Results do not conform to theoretical expectations.

Figure 3

Ancillary Employment Associated with Each Major Industry, Plains Region, 1970

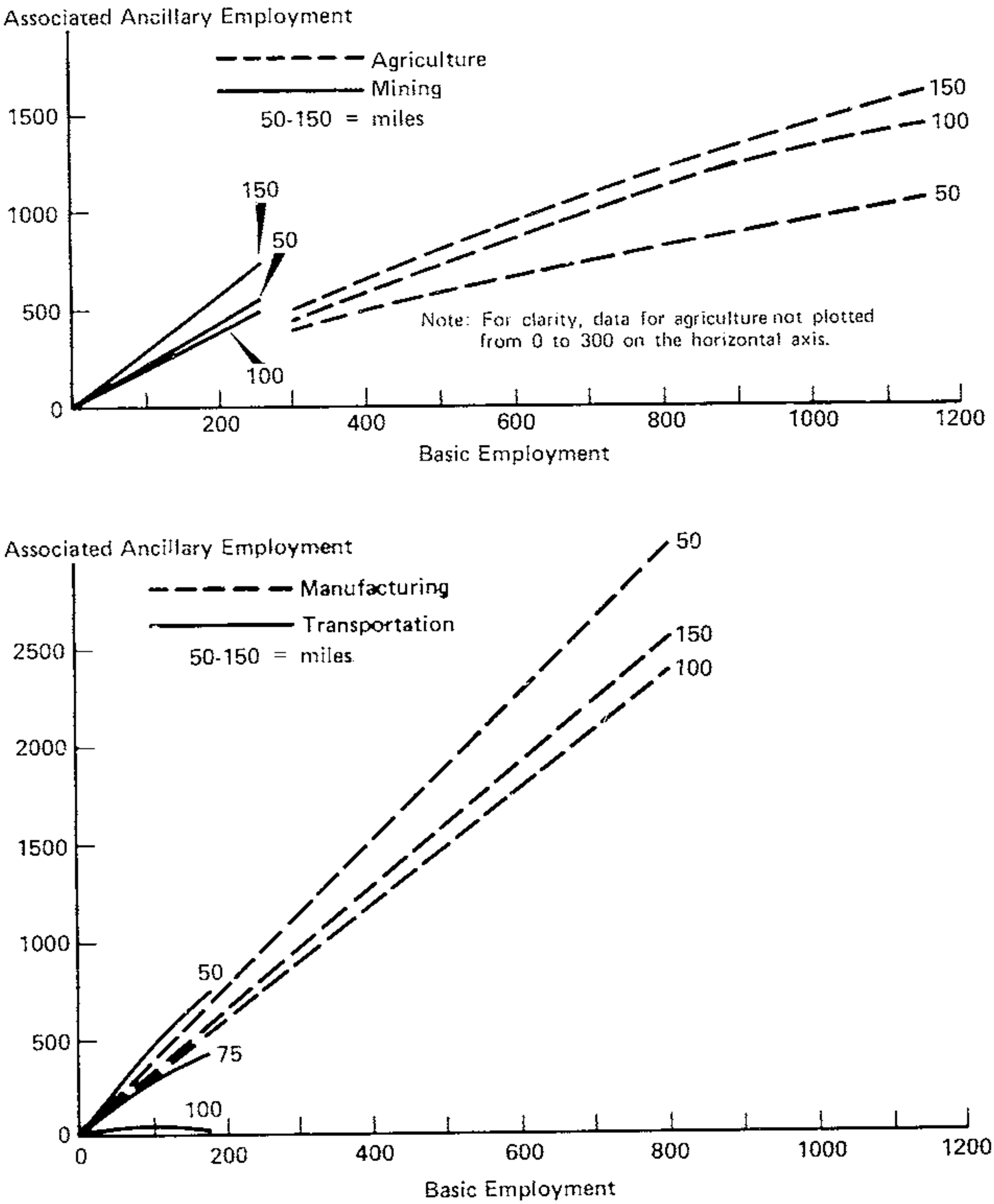


Table 7--Effect of industry size on average multipliers, nonmetro counties, four regions, western U.S., 1970 1/

Region	Number of employees												
	50	100	150	200	250	500	750	1,000	1,250	2,000	3,000	4,000	5,000
							Agriculture						
Plains	1.60	1.58	1.56	1.54	1.52	1.43	1.33	1.23	1.13				
Mountain	1.53	1.52	1.51	1.50	1.49	1.44	1.39	1.34	1.29				
Intermountain	.62	.62	.62	.61	.61	.59	.57						
Coast	1.45	1.45	1.44	1.44	1.43	1.41	1.39	1.37	1.34	1.28	1.19		
							Mining						
Plains	2.32	2.24	2.17	2.10	2.02								
Mountain	.72	.72	.73	.73	.73	.75							
Intermountain	.62	.62	.62	.62	.62	.62	.62	.63					
Coast	3.33	3.27	3.21	3.16	3.10	2.81	2.53						
							Manufacturing						
Plains	3.22	3.21	3.21	3.20	3.20	3.18	3.15						
Mountain	2.18	2.18	2.16	2.15	2.14	2.08	1.95	1.89	1.70				
Intermountain	3.93	3.90	3.87	3.84	3.81	3.66	3.51	3.35					
Coast	1.17	1.17	1.17	1.17	1.17	1.17	1.18	1.18	1.19	1.20	1.22	1.24	1.26
							Transportation						
Plains	2.21	2.00	1.79										
Mountain	5.74	5.73	5.72	5.71	5.70								
Intermountain	2/	2/	2/	2/	2/								
Coast	2/	2/	2/	2/	2/								

1/ The mean distance in each region is assumed.
2/ Values have illogical signs.

In all but three cases, the average multiplier declines as industry size increases. In general, this fact indicates that economies of scale are captured in the support sectors as derived demand expands. The instances exhibiting increased average multipliers are mining in the Mountain and Intermountain regions and manufacturing in the Coast region. However, the increases are nominal.

Overall, the statistical model proved most effective in the Plains region. In this region, tests of significance and the signs associated with the variables coincided with expectations implied by economic base and central place theory. The model applied to the remaining regions yielded some parameter estimates which are not in accordance with theoretical expectations, or are not statistically significant. It appears that the shortcomings of the model can be corrected by minor modifications which will not reduce estimating efficiency.

RESEARCH CONSIDERATIONS

Additional theoretical and empirical research is needed on employment prediction models. It should be recognized that no single model may describe systematically the diverse conditions and trade patterns in local economies of every region. However, further refinements in this statistical model should be tested. The following discussion reflects improvements which could be made.

Theoretical constructs are needed on designation of trade centers and their measurement, the effect of distance from a trade center on the trade linkages for various economic activities, and the trade interrelationships among isolated hinterland communities. This study used a fixed set of predetermined trade centers as a point of reference. From these, the distance measure was expected to be indicative of linkages.

The distance variable and its point of reference need refinement. Incorrect trade center designations may have been responsible for the poor results obtained in all regions except the Plains. Travel barriers, weather uncertainty, and the quality of highways all influence the extent, timing, and cost of travel. Better methods to identify trade centers and their sphere of influence could improve statistical results materially.

The effect of distance from a trade center cannot be anticipated from theory. Distance proved to be important in the Plains region model. But the effect on the multiplier was different for each industry. Thus, theory provided no a priori rationale against which results could be compared.

Finally, a complete theoretical and empirical model would specify the linkages of one community with another and the direction of that linkage. The model in this study abstracted from the fact that one hinterland county could service some of the support service demands of adjacent counties. Only basic sectors within the county in relation to a major trade center were hypothesized to influence support activities in this study. However, a model should hypothesize that support activities intermediate between the regional trade center and the hinterland county are important. In order to incorporate this influence into the model, a directional measure of trade would have to be rationalized and quantified. Then, size of activity in adjacent counties could have been incorporated into the model.

Some of the problems of specifying a predicting model are due to data limitations. The secondary data used in this study are highly aggregated. The difference between extensive and intensive farming operations will affect interindustry linkages but are not reflected in the aggregated measure of agricultural employment. Another data problem is illustrated by construction employment. It is probable that construction is supported for the most part by a flow of funds into a community. As such it should be a basic activity. But census data often are distorted because large projects depending upon temporary personnel dominate the observation matrix.

Finally, a model using cross-section data should be updated periodically. The model should be adapted to employment and income series which are available annually for small geographic areas such as counties. Certain problems are evident in all series and special care must be exercised when such data are used. Income data, for instance, have problems concerning imputed income, transfers, and spending leads and lags.

APPENDIX TABLES

Appendix table 1--Test of homogeneity for contiguous regions 1/

Regions 2/	"F" values	Degrees of freedom
P, M, I, C	6.545**	69,375
P, M, I	4.756**	46,331
M, I, C	2.815**	46,217
P, M	4.892**	23,253
M, I	2.440**	23,173
I, C	2.342**	23,122

1/ Values shown are for nonmetro counties.
2/ Abbreviations are: P, Plains; M, Mountain; I, Intermountain; and C, Coast.

**Significant at $F_{.01}$.

Appendix table 2--Multicollinearity among independent variables, nonmetro counties, four regions, western U.S., 1970

Variables tested	Variables excluded 1/	Region			
		Plains	Mountain	Inter-mountain	Coast
		R^2			
A	DA, D^2A, A^2	.65	.60	.55	.74
DA	A, D^2A, A^2	.58	.67	.55	.66
D^2A	A, DA, A^2	.76	.70	.76	.77
A^2	A, DA, D^2A	.87	.59	.55	.59
M	DM, D^2M, M^2	.52	.49	.41	.35
DM	M, D^2M, M^2	.51	.27	.17	.30
D^2M	M, DM, M^2	.52	.46	.21	.30
M^2	M, DM, D^2M	.48	.45	.41	.43
F	DF, D^2F, F^2	.85	.65	.59	.50
DF	F, D^2F, F^2	.72	.51	.46	.45
D^2F	F, DF, D^2F	.88	.47	.66	.49
F^2	F, DF, D^2F	.81	.66	.49	.41
T	DT, D^2T, T^2	.69	.30	.58	.42
DT	T, D^2T, T^2	.40	.49	.31	.50
D^2T	T, DT, T^2	.77	.53	.45	.51
T^2	T, DT, D^2T	.67	.23	.55	.43
V		.27	.35	.16	.74
I		.44	.43	.33	.72
C	Cd	.70	.35	.55	.38
Cd	C	.44	.47	.52	.58
H	Hd	.82	.48	.67	.71
Hd	H	.69	.52	.40	.40
Y		.41	.21	.21	.57

1/ Cross-product and related dummy variables were excluded from equations as indicated.

Appendix table 3--Simple r's for variables of nonmetro counties, Plains region, western U.S., 1970 1/

	A	AD	AD^2	A^2	M	MD	MD^2	M^2	F	FD	FD^2	F^2	T	TD	TD^2	T^2	V	I	C	Cd	H	Hd	Y	ANC
A				.87					.67			.60							.54					.65
AD			.84								.59													
AD^2											.59								.50					
A^2									.72			.72												.60
M						.89	.68	.95																
MD							.89	.59																
MD^2																								
M^2																								
F										.57		.93							.77		.58	-.61		.90
FD											.73	.57							.51					
FD^2															.78									.74
F^2																			.74					
T																.91					.69			
TD															.82									
TD^2																					.63			
T^2																								
V																								
I																								.53
C																				-.56				.77
Cd																								-.59
H																						-.84		.76
Hd																								-.71
Y																								
ANC																								

1/ Only values larger than .50 are recorded.

Appendix table 4--Simple r's for variables of nonmetro counties, Mountain region, western U.S., 1970 1/

	A	AD	AD^2	A^2	M	MD	MD^2	M^2	F	FD	FD^2	F^2	T	TD	TD^2	T^2	V	I	C	Cd	H	Hd	Y	ANC
A				.93					.64			.59												.63
AD			.89							.63	.59													
AD^2										.54	.64			.51										
A^2									.61			.65												.51
M						.55		.96																
MD							.93																	
MD^2																								
M^2																								
F												.92												.80
FD											.90													
FD^2																								.70
F^2																								
T																.92								
TD															.96									
TD^2																								
T^2																								
V																								
I																								
C																				-.56				
Cd																								-.64
H																						-.83		
Hd																								
Y																								
ANC																								

1/ Only values larger than .50 are recorded.

Appendix table 5--Simple r's for variables of nonmetro counties, Intermountain region, western U.S., 1970 1/

	A	AD	AD^2	A^2	M	MD	MD^2	M^2	F	FD	FD^2	F^2	T	TD	TD^2	T^2	V	I	C	Cd	H	Hd	Y	ANC
A				.93					.56											-.51				.53
AD			.88							.52	.59													
AD^2										.51	.72													
A									.56			.51												.50
M						.54		.92																
MD							.91																	
MD^2																								
M																								
F												.89												.72
FD											.89													
FD^2																								
F^2																								.51
T														.56		.94								
TD															.89									
TD^2																								
T^2																								
V																								
I																								
C																				.55	.52			.55
Cd																								-.58
H																						-.66		.57
Hd																								
Y																								
ANC																								

1/ Only values larger than .50 are recorded.

Appendix table 6--Simple r's for variables of nonmetro counties, Coast region, western U.S., 1970 1/

	A	AD	AD²	A²	M	MD	MD²	M²	F	FD	FD²	F²	T	TD	TD²	T²	V	I	C	Cd	H	Hd	Y	ANC
A				.92													.64	.53						.73
AD			.78							.56	.54													
AD²											.55													
A²																	.52							.58
M						.67		.95																
MD							.96																	
MD²																								
M²						.50																		
F												.95												.65
FD											.85													
FD²																								
F²																								.53
T																.93								
TD															.95									
TD²																								
T²																								
V																								
I																					.55			.64
C																								
Cd																					-.57			-.53
H																						-.65		.60
Hd																								
Y																								
ANC																								

1/ Only values larger than .50 are recorded.

Appendix table 7--Association between distance 1/ and size and mix of major basic industries, nonmetro counties, four regions, western U.S., 1970

Variable	Region			
	Plains	Mountain	Inter-mountain	Coast
	Correlation coefficient, r			
Agriculture				
Number employed	.08	.11	.07	.12
Percent employed	.01	2/	.12	.01
Mining				
Number employed	2/	2/	.04	.04
Percent employed	2/	.04	2/	.28
Manufacturing				
Number employed	.05	.18	.10	.23
Percent employed	.05	.07	.02	.05
Transportation				
Number employed	.04	.02	.07	.02
Percent employed	.01	2/	.01	.03

1/ Distance from largest city in county to Rand McNally trade center in miles.
2/ Less than .01.

Appendix table 8--Designated regional trade centers, western U.S., 1970

Trade center	Trade center	Trade center
Wyoming	Longview	New Mexico
Cheyenne (Scotts Bluff, Nebr.)	Oregon	Santa Fe
Casper	Walla Walla, Wash.	Clovis
Sheridan	Portland	Albuquerque
Rock Springs	Salem	Gallup
Colorado	Eugene	Roswell
Greeley	Roseburg	Carlsbad
Colorado Springs	Coos Bay	Hobbs
Grand Junction	Medford	Texas
Denver	Klamath Falls	El Paso
Pueblo	California	Amarillo
Durango	Klamath Falls, Ore.	Utah
Montana	Eureka	Salt Lake City
Miles City	Redding	Ogden
Billings	Reno, Nev.	Provo
Great Falls	Santa Rosa	Logan
Havre	Sacramento	Nevada
Bozeman	Chico/Oroville	Las Vegas
Helena	San Francisco/Oakland/ San Jose	Reno
Kalispell	Santa Cruz/Watsonville	North Dakota
Missoula	Salinas/Monterey	Williston
Butte	Stockton	Dickinson
Arizona	Modesto	Bismarck
Flagstaff	Merced	Minot
Douglas	Fresno	Grand Forks
Nogales	Visalia/Hanford	Fargo
Tucson	San Luis Obispo	South Dakota
Yuma	Santa Barbara	Rapid City
Phoenix	Bakersfield	Aberdeen
Washington	Oxnard/Ventura	Huron
Spokane	Los Angeles	Mitchell (Sioux City, Iowa)
Walla Walla	San Diego	Watertown
Wenatchee	El Centro/Calexico	Sioux Falls
Bellingham	San Bernadino/Riverside	
Seattle	Marysville	
Olympia	Idaho	
Yakima	Twin Falls	
Tacoma	Lewiston	
Port Angeles	Boise	
Bremerton	Pocatello	
Aberdeen	Idaho Falls	

Source: 1970 Rand McNally Commercial Atlas and Marketing Guide, Rand McNally and Co., Chicago.

END